JMIR Serious Games

Journal Impact Factor (JIF) (2023): 3.8
Volume 9 (2021), Issue 4 ISSN 2291-9279 Editor-in-Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

A Telehealth System Incorporating a Serious Game Intervention to Aid Occupational Therapists in Identifying and Treating Children With Difficulty Crossing the Body's Midline: Key Informant Interviews Among Occupational Therapists (e27761)
Jonathan Jacobs, Reolyn Heymann, Jacob Greeff. 4

Feasibility of a Sensor-Controlled Digital Game for Heart Failure Self-management: Randomized Controlled Trial (e29044)
Kavita Radhakrishnan, Christine Julien, Tom Baranowski, Matthew O'Hair, Grace Lee, Atami Sagna De Main, Catherine Allen, Bindu Viswanathan, Edison Thomaz, Miyong Kim. 21

Development and Validation of a Mobile Game for Culturally Sensitive Child Sexual Abuse Prevention Education in Tanzania: Mixed Methods Study (e30350)
Maria Malamsha, Elingarami Sauli, Edith Luhanga. 43

Effects of an Immersive Virtual Reality Exergame on University Students' Anxiety, Depression, and Perceived Stress: Pilot Feasibility and Usability Study (e29330)
Wenge Xu, Hai-Ning Liang, Nilufar Baghaei, Xiaoyue Ma, Kangyou Yu, Xuanru Meng, Shaoyue Wen. 58

Nationwide Deployment of a Serious Game Designed to Improve COVID-19 Infection Prevention Practices in Switzerland: Prospective Web-Based Study (e33003)
Melanie Suppan, Loric Stuby, Stephan Harbarth, Christophe Fehlmann, Sophia Achab, Mohamed Abbas, Laurent Suppan. 67

Virtual Reality in Health Care: Bibliometric Analysis (e32721)
Christian Pawassar, Victor Tiberius. 84

Using the Behaviour Change Wheel Program Planning Model to Design Games for Health: Development Study (e29964)
Michael Robertson, Tom Baranowski, Debbe Thompson, Karen Basen-Engquist, Maria Swartz, Elizabeth Lyons. 103

Design of a Virtual Reality Interactive Training System for Public Health Emergency Preparedness for Major Emerging Infectious Diseases: Theory and Framework (e29956)
Yue Luo, Mei Li, Jian Tang, JianLan Ren, Yu Zheng, XingLi Yu, LinRui Jiang, DingLin Fan, YanHua Chen. 116

Investigating the Use of a Serious Game to Improve Opioid Safety Awareness Among Adolescents: Quantitative Study (e33975)
Olufunmilola Abraham, Claire Rosenberger, Kathleen Tierney, Jen Birstler. 146

XSL•FO
RenderX

The Effect of Video Game–Based Interventions on Performance and Cognitive Function in Older Adults: Bayesian Network Meta-analysis (e27058)
Chao Yang, Xiaolei Han, Mingxue Jin, Jianhui Xu, Yiren Wang, Yajun Zhang, Chonglong Xu, Yingshi Zhang, Enshi Jin, Chengzhe Piao. 163

Gamifying Sexual Education for Adolescents in a Low-Tech Setting: Quasi-Experimental Design Study (e19614)
Hussein Haruna, Kingsley Okoye, Zamzami Zainuddin, Xiao Hu, Samuel Chu, Samira Hosseini. 183

Evidence of Construct Validity of Computer-Based Tests for Clinical Reasoning: Instrument Validation Study (e17670)
Tianming Zuo, Baozhi Sun, Xu Guan, Bin Zheng, Bo Qu. 198

Using a 360° Virtual Reality or 2D Video to Learn History Taking and Physical Examination Skills for Undergraduate Medical Students: Pilot Randomized Controlled Trial (e13124)
Yi-Ping Chao, Hai-Hua Chuang, Li-Jen Hsin, Chung-Jan Kang, Tuan-Jen Fang, Hsueh-Yu Li, Chung-Guei Huang, Terry Kuo, Cheryl Yang, Hsin-Yih Shyu, Shu-Ling Wang, Liang-Yu Shyu, Li-Ang Lee. 209

An Interactive Physical-Cognitive Game-Based Training System Using Kinect for Older Adults: Development and Usability Study (e27848)
Teerawat Kamnardsiri, Kochaphan Phirom, Sirinun Boripuntakul, Somporn Sungkarat. 247

Immersion Experiences in a Tablet-Based Markerless Augmented Reality Working Memory Game: Randomized Controlled Trial and User Experience Study (e27036)
Bo Zhang, Nigel Robb. 263

Long-term Effectiveness and Adoption of a Cellphone Augmented Reality System on Patients with Stroke: Randomized Controlled Trial (e30184)
Chong Li, Xinyu Song, Shugeng Chen, Chuankai Wang, Jieying He, Yongli Zhang, Shuo Xu, Zhijie Yan, Jie Jia, Peter Shull. 314

Dissecting Digital Card Games to Yield Digital Biomarkers for the Assessment of Mild Cognitive Impairment: Methodological Approach and Exploratory Study (e18359)
Karsten Gielis, Marie-Elena Vanden Abeele, Robin De Croon, Paul Dierick, Filipa Ferreira-Brito, Lies Van Assche, Katrien Verbert, Jos Tournoy, Vero Vanden Abeele. 348

A Virtual Supermarket Program for the Screening of Mild Cognitive Impairment in Older Adults: Diagnostic Accuracy Study (e30919)
Mingli Yan, Huiru Yin, Qiuyan Meng, Shuo Wang, Yiwen Ding, Guichen Li, Chunyan Wang, Li Chen. 395

Effectiveness and Utility of Virtual Reality Simulation as an Educational Tool for Safe Performance of COVID-19 Diagnostics: Prospective, Randomized Pilot Trial (e29586)
Tanja Birrenbach, Josua Zbinden, George Papagiannakis, Aristomenis Exadaktylos, Martin Müller, Wolf Hautz, Thomas Sauter. 407

Contribution of the co.LAB Framework to the Collaborative Design of Serious Games: Mixed Methods Validation Study (e33144)
Dominique Jaccard, Laurent Suppan, Félicia Bielser. 438

Team Building Through Team Video Games: Randomized Controlled Trial (e28896)
Mark Keith, Douglas Dean, James Gaskin, Greg Anderson. 454

Changes in Self-Reported Web-Based Gambling Activity During the COVID-19 Pandemic: Cross-sectional Study (e30747)
Emma Claesdotter-Knutsson, Anders Håkansson. 470

XSL•FO
RenderX

Predicting Video Game Addiction Through the Dimensions of Consumer Video Game Engagement: Quantitative and Cross-sectional Study (e30310)
Amir Abbasi, Umair Rehman, Zahra Afaq, Mir Rafeh, Helmut Hlavacs, Mohammed Mamun, Muhammad Shah. 481

Relationship Between Illness Representations and Symptoms of Internet Gaming Disorder Among Young People: Cross-Lagged Model (e28117)
Xue Yang, Kei Wong, Rui She, Chengjia Zhao, Nani Ding, Huihui Xu, Xiaolian Tu, Xinyi Lai, Guohua Zhang. 501

Reviews

Awareness, Prevention, Detection, and Therapy Applications for Depression and Anxiety in Serious Games for Children and Adolescents: Systematic Review (e30482)
Kim Martinez, Maria Menéndez-Menéndez, Andres Bustillo. 127

Effects of Active Video Games on Health-Related Physical Fitness and Motor Competence in Children and Adolescents With Overweight or Obesity: Systematic Review and Meta-Analysis (e29981)
Cristina Comeras-Chueca, Jorge Marin-Puyalto, Angel Matute-Llorente, German Vicente-Rodriguez, Jose Casajus, Alex Gonzalez-Aguero. .
2 3 0

Augmented Reality in Physical Therapy: Systematic Review and Meta-analysis (e30985)
Maria Vinolo Gil, Gloria Gonzalez-Medina, David Lucena-Anton, Veronica Perez-Cabezas, María Ruiz-Molinero, Rocío Martín-Valero. 328

Examining the Academic Trends in Neuropsychological Tests for Executive Functions Using Virtual Reality: Systematic Literature Review (e30249)
Euisung Kim, Jieun Han, Hojin Choi, Yannick Prié, Toinon Vigier, Samuel Bulteau, Gyu Kwon. 373

Software Engineering Frameworks Used for Serious Games Development in Physical Rehabilitation: Systematic Review (e25831)
Jorge Ambros-Antemate, María Beristain-Colorado, Marciano Vargas-Treviño, Jaime Gutiérrez-Gutiérrez, Pedro Hernández-Cruz, Itandehui Gallegos-Velasco, Adriana Moreno-Rodríguez. 421

Viewpoints

Game Design in Mental Health Care: Case Study–Based Framework for Integrating Game Design Into Therapeutic Content (e27953)
Panote Siriaraya, Valentijn Visch, Marilisa Boffo, Renske Spijkerman, Reinout Wiers, Kees Korrelboom, Vincent Hendriks, Elske Salemink, Marierose van Dooren, Michael Bas, Richard Goossens. 277

Serious Games Are Not Serious Enough for Myoelectric Prosthetics (e28079)
Christian Garske, Matthew Dyson, Sigrid Dupan, Graham Morgan, Kianoush Nazarpour. 302

Corrigenda and Addenda

Correction: Feasibility of Virtual Reality Audiological Testing: Prospective Study (e34994)
Hye Seol, Soojin Kang, Jihyun Lim, Sung Hong, Il Moon. 513

XSL•FO
RenderX

Original Paper

A Telehealth System Incorporating a Serious Game Intervention to Aid Occupational Therapists in Identifying and Treating Children With Difficulty Crossing the Body's Midline: Key Informant Interviews Among Occupational Therapists

Jonathan Jacobs[1*], BEng, MEng; Reolyn Heymann[1*], BEng, MEng, DEng; Jacob Jacobus Greeff[2], NDip, BTech, MTech

[1]Centre for Collaborative Digital Networks, Department of Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg, South Africa

[2]School of Computer Science and Information Systems, Faculty of Natural and Agricultural Sciences, University of North-West, North-West, South Africa

[*]these authors contributed equally

Corresponding Author:
Jonathan Jacobs, BEng, MEng
Centre for Collaborative Digital Networks
Department of Electrical and Electronic Engineering Science
University of Johannesburg
Engineering Building
Corner of Kingsway and University Road, Auckland Park
Johannesburg
South Africa
Phone: 27 741550219
Email: jonty.jacobs@gmail.com

Abstract

Background: The midline is an imaginary line that isolates the left and right parts of the body. Crossing the midline infers that a body part (eg, hand or foot) can spontaneously move over to the opposite side of the body to perform an action. A child who has difficulty crossing the midline can physically perform actions that cross the center of the body; however, they do not intuitively cross the midline when challenged with a task that requires this movement, as their perceptual components prevent them from engaging on the contralateral side. This requires treatment from an occupational therapist. Owing to the recent COVID-19 pandemic, access to therapeutic sessions was not possible or reduced, putting the responsibility for treatment on caretakers at home. Caretakers do not have the knowledge and skills to provide treatment, and occupational therapists do not receive adequate feedback from caretakers on the child's progress.

Objective: The first objective is to adapt a simple serious game, or applied game, into a telehealth solution. Children will play the game at home under the supervision of a caretaker, and the results will be stored on the web. Occupational therapists can monitor progress via a web-based dashboard, receive additional valuable feedback about the child's behavior during treatment, and easily adapt the game to target specific needs. The second objective is to evaluate whether the implemented telehealth solution is feasible as a treatment option for midline crossing difficulties and thus fit for purpose.

Methods: To meet the first objective, engineering and game development stakeholders formed a team with an occupational therapist, and through a collaborative design process combined with an agile programming approach, a telehealth solution was designed to assist remote monitoring of the serious gameplay. For the second objective, 6 different occupational therapists were introduced to the game, had the opportunity to play the game, and then provided feedback regarding the feasibility, benefits, and applicability of the system during structured interviews.

Results: A telehealth system was designed aimed to address this problem. All results are saved on the web and accessed by occupational therapists via a dashboard. In addition, observed behavioral information is also saved. During the interviews, occupational therapists indicated that the dashboard would support their treatment plan and was indeed a feasible solution.

XSL•FO
RenderX

Conclusions: The feedback from the occupational therapists for this telehealth solution suggests a feasible method to treat midline crossing problems remotely. The therapists commented on the convenience of integrating both assessment and treatment into the same application, as it assists them when grading a child. The therapists collectively agreed that the quantitative aspect the serious game creates by providing measurable and standardized data proves advantageous when compared with traditional methods of assessment and treatment.

(JMIR Serious Games 2021;9(4):e27761) doi:10.2196/27761

KEYWORDS

serious games; input device; telehealth; occupational therapy; midline crossing

Introduction

Background

Owing to the COVID-19 pandemic, numerous health care practitioners were unexpectedly required to transition their standard in-person treatment to telehealth options, often without advance preparation or training [1]. Telehealth is becoming an increasingly used service delivery model in rehabilitation services. Telehealth has the potential to alleviate provider shortages, decrease costs associated with providing therapy, and allow for treatment within a client's natural environment [2]. Telehealth can be used by occupational therapists for evaluation, intervention, education, and prevention of injury or exacerbation of conditions [3]. Telehealth facilitates collaboration and consultation with other professionals, which facilitates coordination of care [4].

The midline is an imaginary line that isolates the left and right parts of the body. Crossing the midline infers that a body part—for example, a hand or foot—can spontaneously move over to the opposite side of the body to perform an action. To clarify, a child who has difficulty crossing the midline can physically perform actions that cross the center of the body; however, they do not intuitively cross the midline when challenged with a task. The diagnosis of children with midline crossing difficulties requires a cluster of clinical observations that indicate bilateral integration dysfunction [4]. A study was performed in which 10% of the sample was considered to constitute a possible deficit range, and a further 10% was considered to be in the suspect range. The remaining 80% was considered to fall in the normal range [4].

Midline crossing is a difficulty that needs the intervention of an occupational therapist. Midline crossing difficulties may affect the child's physical well-being as well as their future development if it is not addressed in a timely manner. If a child does not get the needed intervention due to the COVID-19 lockdown, it can affect the child's future. Being able to cross the midline is a developmental milestone. By the age of 5 years, a child is expected to be able to cross the midline, that is, use both sides of the body simultaneously [5]. When the midline is crossed spontaneously, supporting neural networks and pathways for specific activities are developed. This is a prerequisite skill essential for the development and maintenance of motor and cognitive demands associated with specific activities. Consequently, children who have trouble crossing the body's midline also frequently experience difficulty with reading, writing, tying their shoelaces, brushing their teeth, and participating in physical activities [6].

Current treatment approaches used by therapists are arts and crafts in which actions such as threading beads, cutting, pasting, and folding paper are used as these actions require the midline to be crossed. Finger puppets or stickers are also used by placing or sticking the puppets or stickers on one of the child's hands and then encouraging the child to remove the puppet or sticker with the opposite hand. Other common methods used by therapists include building blocks and playing Twister and marching games using arms and legs. Occupational therapists use standardized assessment tools such as the Movement Assessment Battery for Children, the Developmental Test of Visual Motor Integration, and the Draw a Person Test to measure changes in function and occupational. These standardized tests have procedures for administration and scoring [7]. However, when measuring the outcome of an intervention, there is still a lack of tools available for objective measurement, particularly among children with perceptuomotor or attention deficit disorder [7].

At present, occupational therapists observe if the child exhibits the actions given in Textbox 1.

Textbox 1. How to identify whether a child has difficulties crossing the midline.

How to identify whether a child has difficulties crossing the midline

- Swaps hands midway through a task when writing, drawing, painting, or coloring
- Uses the left hand for activities on the left side of the body and right hand for activities on the right-hand side
- Rotates their trunk to the opposite side when reaching across the body (to avoid crossing the body midline)
- Has difficulty visually tracking an object from one side of the body to the other, such as following text when reading
- Has poor pencil skills (pencil grip)
- Uses different feet to kick a ball (mixed dominance)
- Has difficulty coordinating gross motor patterns (eg, crawling, skipping, and star-jumps)

XSL•FO
RenderX

On the basis of their observations and discretion, they conclude whether the child has difficulty crossing the midline. Assessment is an integral part of the occupational therapy process and is a necessity for evidence-based practice. Without appropriate measurement, therapists cannot provide evidence for the offered interventions [8]. The assessment is subjective, and there are issues with this subjectivity, as the same level of difficulty may be rated differently by different assessors. When therapists write reports for the child, it is challenging to justify the child's improvement as there is no criterion to compare it with. In addition, numerous children up to the age of 5 years living in underdeveloped countries, including South African rural areas, face exposure to multiple risks affecting their early childhood development [8]. These children and their parents are often not aware of the functional difficulties they may have, and even if they are aware, they need to travel long distances to receive the therapy they require. Moreover, because of the recent COVID-19 pandemic and resulting lockdowns, many children who would normally receive occupational therapy could not access occupational therapists or had to decrease their sessions.

If a child has a pathology of midline crossing, frustration in the behavior of the child will be noticeable, as the child will become angry when trying to engage in fine motor activities because of less refined hand skills. The coordination of both sides of the body will be less refined, leading to difficulties experienced when playing sports or doing any physical activity. In addition, when children have difficulty with midline crossing, they may additionally have trouble visually tracking an object from left to right. Thus, when the object reaches the midline, they often blink and have to refocus, and this results in their losing their place while reading. In addition, when drawing horizontal and diagonal lines as well as writing letters such as an *x*, they may segment these lines rather than overlap them because of midline crossing difficulties [9]. Some children may struggle to cross the body's midline easily. When a child shows hesitancy in reaching, stepping, or looking across the midline of the body, it is known as midline crossing inhibition. Sometimes, this delay can be seen when a child hesitates or is clumsy during gross motor tasks that require the arm or leg to cross over to the other side. Some children with delayed midline crossing skills may display some *compensatory mechanisms* in school that make writing awkward for them. Crossing the midline is a treatable affliction. If a child struggles to cross the body's midline and is treated, milestones such as developing a dominant hand will occur [6]. Pencil skills and fine motor tasks will be refined, easing the transition to an academic environment where those skills are expected to be grasped [6]. The child will be able to complete self-care tasks, for example, brushing their teeth and getting dressed. The child will be able to kick and hit balls as well as run as their gross and fine motor skills will be improved. Finally, their ability to visually track across a page effectively will be better and, therefore, will result in fluent reading.

At the forefront of technological advancements in occupational therapy are serious games [10]. For the purpose of this research, serious games are described, among other things, as digital games, virtual environments, simulations, and a mixed reality that engage the player. These serious games form encounters and experiences that convey meaning [11]. Serious games are introduced as a need to meet objectives that go beyond entertainment and benefit the user in the area that needs to be mitigated. The applications created under the terminology of serious games induce motivation and engage the user [12].

Objective

Before proposing how serious games can be integrated, a major obstacle to treating children in the conventional sense is boredom because of the intensive and repetitive practice required. The advantage of incorporating input devices and the use of serious games into treatment methods is that it combats the boredom factor. The child is placed into a game environment that is similar to the real world in terms of the perceptual stimuli it exhibits, which then puts the child at ease [13]. The child then has the ability to manipulate and control some of the stimuli and see the outcome of their actions in real time and adjust them accordingly. This aspect, being the interactive component, creates an engagement with the environment, allowing the child to feel and be in control of their movement. This idea is explained as *the perceptual illusion of nonmeditation* [14]. The sensations that are familiar and present when playing the game and the ability of the child to control and manipulate the stimuli that surround them generate the psychological effects of enjoyment and, particularly, involvement [14]. The unique merging of purpose and pleasure develops intrinsic motivation in the child.

Microsoft's Kinect sensor has been highly investigated and used for the development of new complements that help improve or optimize rehabilitation processes worldwide. In 2012, Ruiz and Cantos [15] designed a therapeutic tool using the Kinect for neurorehabilitation using games to stimulate patients, cognitive functions, perceptions, and gross and visual motor skills through play. This tool verified that patients had fun while being treated in this manner; intrinsic motivation is achieved by merging purpose and pleasure.

Moreover, using the Kinect and combining purpose and pleasure, Chang et al [16] performed a study in Taiwan in 2013 that proposed the possibility of rehabilitating two 14-year-old adolescents with cerebral palsy through therapies personalized to their condition. Data showed that the 2 participants had significantly increased motivation for upper limb rehabilitation, thus improving exercise performance during the intervention phases.

Although Kinect is a useful device, it can, however, encounter the problem of misdetection when it comes to extremity angles or overlapping extremities [17]. In 2016, the Taiwan University proposed a new rehabilitation gaming system, which focuses on the upper part of the body with wireless inertial measurement units (IMUs) and a Kinect device. The Kinect was used as a base tracking system by the gaming system. Multiple sets of IMUs were integrated into the extremity of the subject to calculate the angles through algorithms. Wireless IMUs were also added to compensate for the error in the calculation of angles in the Kinect device [17]. This study exposed that the use of the Kinect by itself may not be sufficient and, therefore, additional sensors were needed to ensure correct motion capture.

XSL•FO
RenderX

Nintendo's Wii Fit is a commercial product used for both fitness and fun and was created to encourage people to exercise as well as improve balance. The Wii Fit uses the Nintendo Wii console and a balance board. A study conducted by the University of Naples found that the Wii Fit showed better improvements regarding physical therapy in terms of balance and self-confidence than conventional treatment [18]. It was concluded that the Wii Fit is acceptable as an adjunct to virtual rehabilitation interventions and provides an exciting new therapy device [18]. The Wii console does not provide a platform or offer a game in which more specific impositions, such as difficulty crossing the midline, can be treated. If one wanted to use the Wii for treatment or training purposes, one would need to purchase not only the balance board but also the console and remote, thus making it too expensive for the communities that this study is aimed at.

Ultraleap's Leap Motion Controller is a computer hardware sensor device analogous to a mouse [19]. Using motion capture technology, Leap Motion is able to process the input. It does not require direct hand contact with the device as input; instead, hand and finger motions are tracked. Despite the small size of Leap Motion, it is capable of capturing smaller details, such as finger movement. Leap Motion is smaller and cheaper than Kinect; however, Kinect is more precise in capturing movement. Sourial and Reichardt [20] proposed implementing a virtual therapist (VT) to help patients do their exercises at home in an engaging gamified environment. The VT artificial intelligence used a hierarchical finite-state machine architecture. Hand therapy helps the patient regain the hand's full functionality after a certain injury or surgery. Hand therapy could be a very tedious process that implies physical exhaustion [20]. In addition, finding appointments with the therapist frequently enough for an efficient healing process is difficult and costly. To test the efficiency of the VT, a web-based hand therapy exercise was implemented using the Unity platform to build the exercise environment. Leap Motion technology was used to detect the information of the hand movement. This exercise was tested on 19 participants. The idea of being coached by a VT was welcomed by the participants, as the exercise was fun and motivating to them. VT guidance and assessment were helpful and easy to follow. However, some modifications are needed in the pain detection part to form a more efficient exercise [20].

Another system using Leap Motion and capable of improving fine motor skills in children, was proposed by Hidalgo et al [21] through a serious game and 3D environment. The proposed system allowed the therapist to choose among different levels of serious games according to the child's needs. The game excited the children; however, the children took time to adapt to the game because of the inaccurate readings of Leap Motion. Therefore, it can be noted that including different levels creates excitement for the children as they experience different challenges; however, when technology interferes with the game or is difficult to use, the child takes time to become familiar with the system.

Sony's motion capture system Intel RealSense enables fine-motoric gesture recognition, and its small form factor allows for preintegration into notebooks and tablets, substituting conventional cameras [22]. This setup enables new methods of therapy in the form of serious games that are engaging and easy to set up. Chhor et al [22] developed and evaluated a serious game prototype for rehabilitation using Intel's RealSense (called *Breakout*) based on a commercial game framework. Despite the fact that RealSense can easily integrate with the applications mentioned above, it is an expensive app.

Neuroplasticity refers to the ability of the brain to adapt structurally and functionally and is enhanced by training and experience. It is known that neuroplasticity is at its maximum in a critical period, which corresponds to the first 7 years of a child's life [23]. Therefore, with the development of a serious game and given the neuroplasticity of the child's brain at this age, the child may be able to develop the required pathways necessary to conduct movements that once seemed impossible. With the integration of technology and the innovation and creativity of this approach, a beneficial method of mitigating impositions such as midline crossing could be developed to assist occupational therapists in the treatment of children. A serious game with an input device was initially developed to assist in midline crossing therapy [24]. An improved wireless input device that accompanied the serious game was developed [25]. The aim of this study is twofold. The first objective focuses on the design of a web-based telehealth system that consists of a serious game, an input device, and a web-based dashboard that displays relevant data to occupational therapists. The second objective focuses on whether the system is fit for its purpose and can be adopted by occupational therapists to treat patients remotely.

Methods

Objective One: Designing a Remote Monitoring System for a Midline Crossing Serious Game

Overview

An agile software development life cycle approach was taken to develop the end-to-end solution. This methodology was chosen because of the frequent feedback needed during the design process, allowing the design team to give recommendations through collaborative design at the end of each iteration. The design team comprised game designers, electronic engineers, and an occupational therapist. Figure 1 shows the iterations of the agile software development life cycle. It should be mentioned that each component of the solution (serious game, input device, and telehealth system) was developed separately through iterations; however, the components were integrated and tested at the end of each iteration. Furthermore, the initial serious game from the previous project was used as a foundation to produce the web-based telehealth solution presented in this study [24].

A high-level solution design is shown in Figure 2. The sequence of events is illustrated in Figure 3.

XSL•FO
RenderX

Figure 1. Iterations in the agile software development life cycle. Req: requirements.

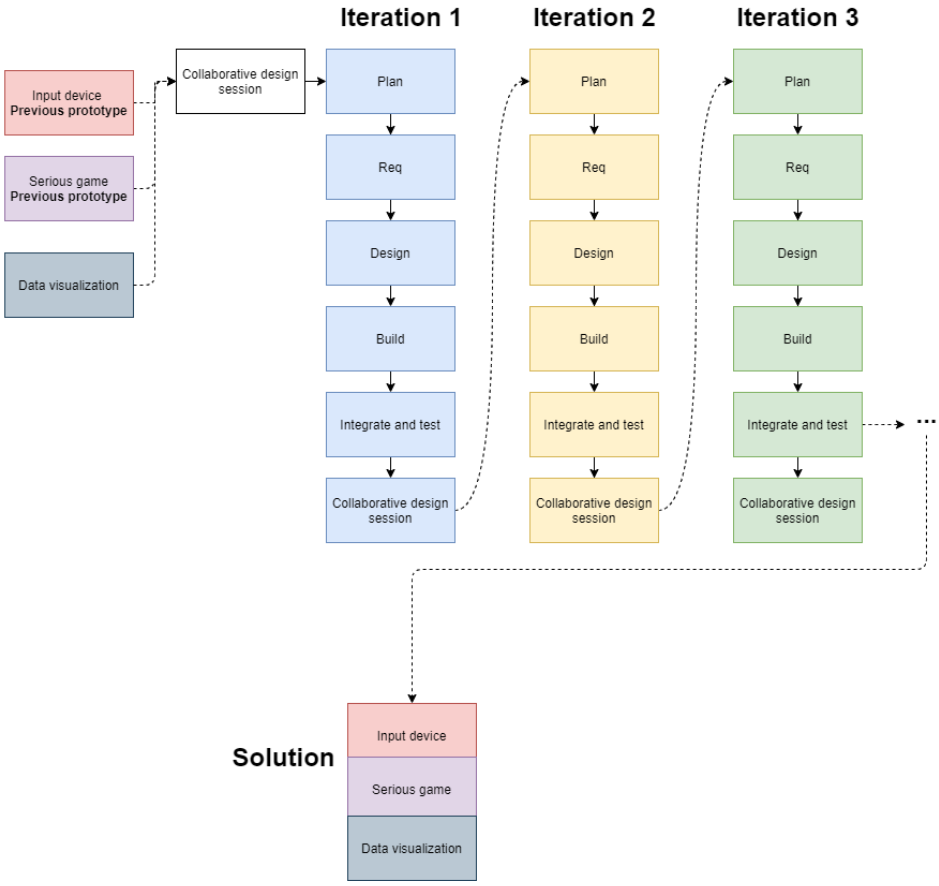
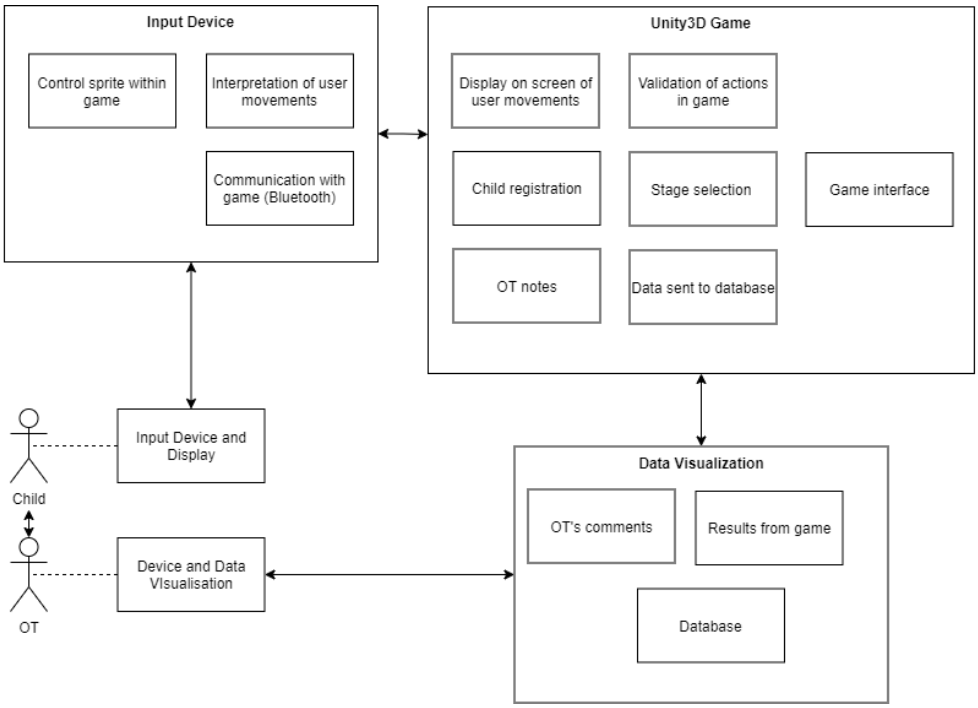
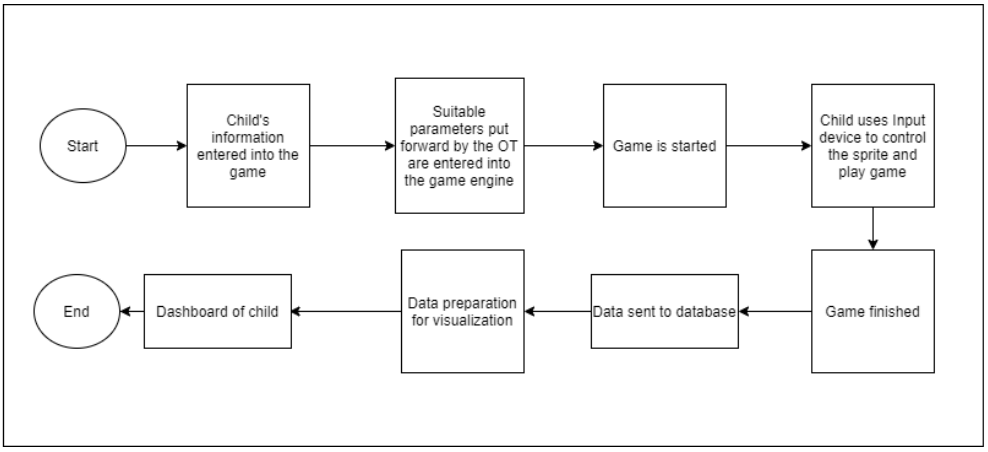


Figure 2. High-level design of solution. OT: occupational therapist.



XSL•FO
RenderX

Figure 3. Sequence of events. OT: occupational therapist.



The innovative low-cost input device shown in Figure 4 was developed previously and used so that the movement of the child could act as an input for the game [25]. The input device is designed in such a way that when the device is turned on and the game started, a connection is established. There are no complicated installations or setup steps. The role of the caretaker is to help the child set up the device with the game. Additionally, the caretaker will explain to the child how to play and then observe the child's movements according to the guidelines designed by the occupational therapists, which will be discussed further in the *Results* section. The design of the data visualization component and the new serious game is presented in this study.

Figure 4. Prototype of input device.

Initial Serious Game

The game genre chosen was a casual game because of its low system requirements, accessibility in most devices, and, at the same time, not demanding high levels of concentration from the player [26]. The game was not connected to the internet, and players had to install it to play it. The fact that the game does not take much concentration is an important factor in ensuring that the child does not become fatigued and frustrated. Moreover, the duration of the game was purposely designed to be short to retain the child's concentration. The developed game comprised a sprite controlled by the child [24]. Sprite is the terminology used to describe an avatar, shape, or character that the child has control over. The device was attached to the hand of the child, and the child moved the device in an arc shape with their arm extended in front of them while standing. The aim of the game was to collect faces falling at a calculated position from the sky. The movements in the game ensured that the child crossed the midline. The game intended to prompt the child to perform particular movements that occupational therapists would invoke when using traditional methods of treatment.

Input Devices

The problems encountered with the initial design of the input device and the detailed new design have been previously published [25]. For the sake of completeness, a short overview of the input devices is given in this section. In the initial input device, an accelerometer and magnetometer were used to obtain the movement of the child [25]. Tilting of the device was not the only issue that affected the feasibility of the design; before every game, a calibration was required to obtain the reference of the sensor. It did not seem to be the most suitable solution for therapists to expect the child to perform the calibration each time they wanted to play the game, ensuring that they do not tilt the sensor. Therefore, a more child-friendly design was needed that would be able to track the movement of the hand from side to side (crossing the body's midline) and transform the position into the game. In addition to accurately transforming the movement, a device that would not require calibrating the sensor was required. When discussions were held with an occupational therapist during the design phase, it was suggested that an input device that would be placed on a surface be designed. The idea behind this was to allow the child to move their hand from side to side and exhibit the same motion required when performing tasks that cross the midline, such as writing. Through extensive research into suitable sensors or electrical devices that could be used, a clean, innovative solution was reached. The design entailed the addition of a linear potentiometer; the input device is shown in Figure 4 [25].

XSL•FO
RenderX

Objective Two: Evaluate Whether the Online Game Is Fit for Purpose

After the prototype was developed, it was crucial to evaluate whether this solution was a feasible option for occupational therapists and whether it was fit for purpose. This was achieved through semistructured interviews with occupational therapists. These occupational therapists all had between 5 to 10 years of experience. In addition, as many children have difficulty crossing the midline, these occupational therapists were competent to give feedback regarding the identification and treatment of midline crossing difficulties. Note that the occupational therapist who was part of the participatory design sessions was not included in the evaluation of the system.

The researcher took each of the participants through a presentation detailing the purpose of this research. The researcher conducted a demonstration of the full solution, taking the occupational therapist through the game and explaining how each stage of the game was played. The occupational therapists then had the opportunity to play the game and evaluate each stage on their own. A semistructured interview took place, allowing the researcher to further explore the insights provided by

the therapist. Each session was conducted individually to ensure unbiased and objective responses. The interview aimed to highlight the strengths and weaknesses of the solution as well as the suitability of administering the solution remotely. The interview provided the therapists with the opportunity to suggest changes and offer recommendations. The questions were purposefully left open-ended so as not to limit the therapist's responses. The main themes that needed to be established from the interviews can be seen in Textbox 2. The questions contained in each section can be seen in Textbox 3.

Textbox 2. Main themes established from interviews.

Main themes

- The feasibility of using a serious game and input device to assist occupational therapists in identifying and treating children with midline crossing difficulties
- The benefits and applicability of the dashboard
- The applicability of the solution in areas where the availability of occupational therapy resources is limited

Textbox 3. Questions asked to occupational therapists during the interviews.

Questions

- In your professional context, comment on whether a serious game and input device would assist you as a therapist to identify children with midline crossing difficulties as well assist in your treatment process. Please include in your answer the following:
  - Advantages
  - Disadvantages
  - Suggested changes
  - Recommendations
- Comment on whether the use of the dashboard is beneficial and applicable to assist you as a therapist in identifying whether a child has pathology with midline crossing. Does the dashboard provide constructive tracking of the child's progress during treatment?
  - Advantages
  - Disadvantages
  - Suggested changes
  - Recommendations
- Comment whether the solution can be reasonably implemented and used effectively in areas where occupational therapy resources are limited.
  - Advantages
  - Disadvantages
  - Suggested changes
  - Recommendations

XSL•FO
RenderX

Results

Objective One: Designing a Remote Monitoring System for a Midline Crossing Serious Game

Adapted Serious Game

The initial game, as described in the previous section, was used as a base for the new telehealth system. The new system comprised an adapted serious game (with distinct levels), an input device, and a web-based backend system that enabled occupational therapists to access dashboards and behavioral information about the intervention. During the collaborative design sessions, it was decided that the serious game would be divided into different stages. Each stage has a specific aim, and different game variables can be set according to each child's individual needs. The game comprises 4 stages: an assessment stage, 2 intervention stages, and a maintenance stage [24]. In each stage, there are 3 variables: distance, speed, and time, which are used to create a specific environment for testing. The time variable, set at 2 minutes, was kept constant for all stages. The chosen duration was advised by the occupational therapist. The duration was purposely set to be short to allow the child to concentrate for a short amount of time without getting bored or losing focus. Each stage addresses a certain intervention, which will be explained below; therefore, the input to each stage is different. The occupational therapist can set these inputs however they see fit (or advice the caretaker at home on which values to use during remote treatment).

The game environment is chosen by selecting a stage and choosing the required values for the inputs for that stage. The game commences with the hat (controlled by the child using the input device) starting either on the left or right side of the screen depending on the handedness of the child. For the purpose of this explanation, a child who is right-handed will be used; therefore, the sprite starts on the left. If the child were left-handed, the sprite would start on the right so that the child's first movement would be crossing the midline. The balls that fall from the sky are strategically placed, depending on the stage, which will be discussed below. When the balls fall, the child must catch the ball in the hat and return the hat to the starting point (either the left or right side of the screen) to earn a point. Textbox 4 shows the variables used in the equations. Table 1 shows the inputs and equations for the variable set. The game flow is illustrated in Figure 5; depending on the stage, the equations shown in Table 1 will be inserted.

Textbox 4. Variables used in the equations to set the game environment.

Variables

- X_0 is the starting distance
- I is the incremental distance
- S is the score
- DV is the distance variance
- Random (a,b) is the random function that provides a random number between a and b
- Sp_0 is the initial speed
- Is_p is the incremental speed

Table 1. Equations used in each stage to set the game environment.

Stage	Input	Distance	Speed	Time
1	X_0[a], I[b]	$X_0+(IX_s$[c]$)$	Constant	Constant
2	X_0, DV[d]	X_0+Random (–DV, DV)	Constant	Constant
3	X_0, Sp_0[e], Is_p[f]	X_0	$Sp_0+(Is_pSX)$	Constant
4	N/A[g]	Random (–7.5,7.5)	Constant	Constant

[a]X_0: starting distance.

[b]I: incremental distance.

[c]IX_S: incremental distance as function of score.

[d]DV: distance variance.

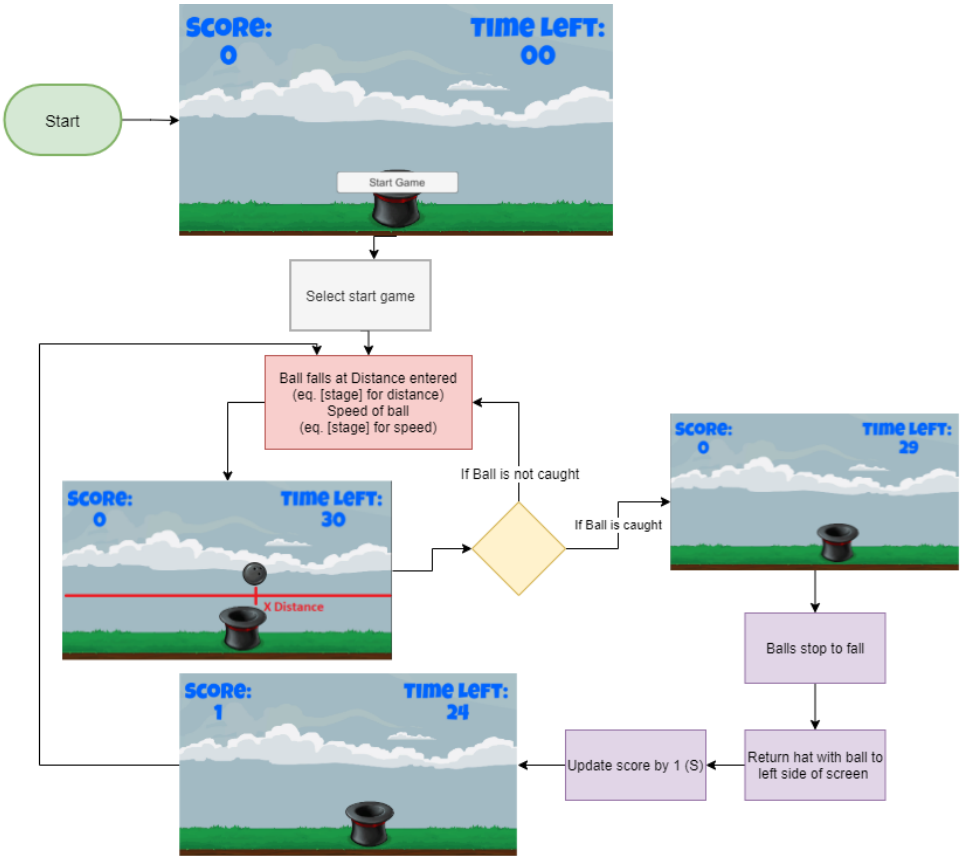
[e]Sp_0: initial speed.

[f]Is_p: incremental speed.

[g]N/A: not applicable.

XSL•FO
RenderX

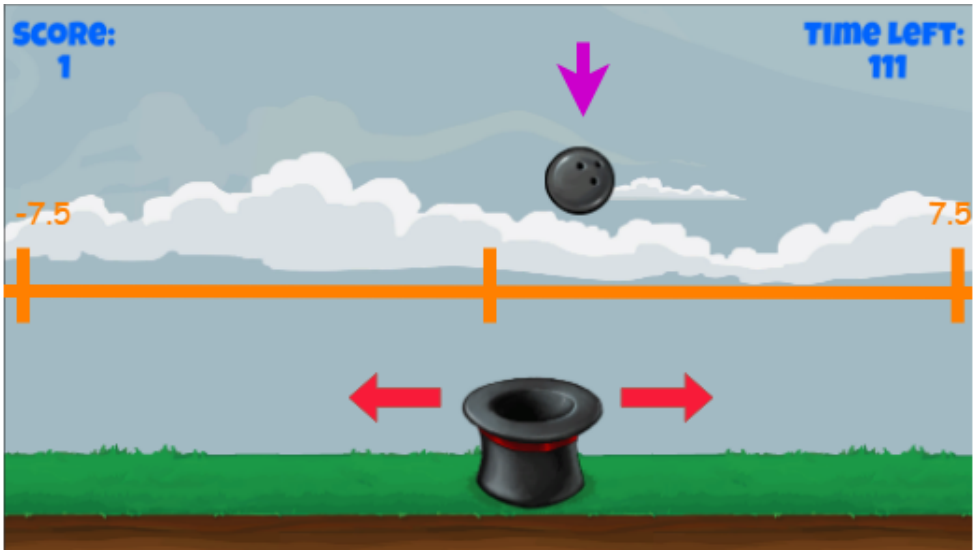
Figure 5. General game flow. eq: equation.



Stage 1

Stage 1 takes the *starting distance* (X_0) and *incremental distance* (I) as input. The starting distance is the point on the x-axis where the first ball will be spawned from the sky (top of the screen). The x-axis is depicted in Figure 6. The *incremental distance* is the distance that is incremented from the starting distance. The point on the x-axis where the balls are spawned is a function of the *score* (S); as the *score* increases, so does the distance of the spawned ball. The score represents the number of balls caught and returned. A ball is considered *returned* when the hat reaches the left side of the screen if the child is right-handed; if the child is left-handed, the hat would need to be *returned* to the right side of the screen. The score will update only after the hat has been returned. Consequently, the point at which the next ball falls will increase by the incremental distance (increase to the right if the child is right-handed). If a ball is missed, intuitively, the score will not increase, thus remaining the same; therefore, the ball will be spawned in the same place instead of being spawned at the incremented distance.

Figure 6. Position on x-axis in game.

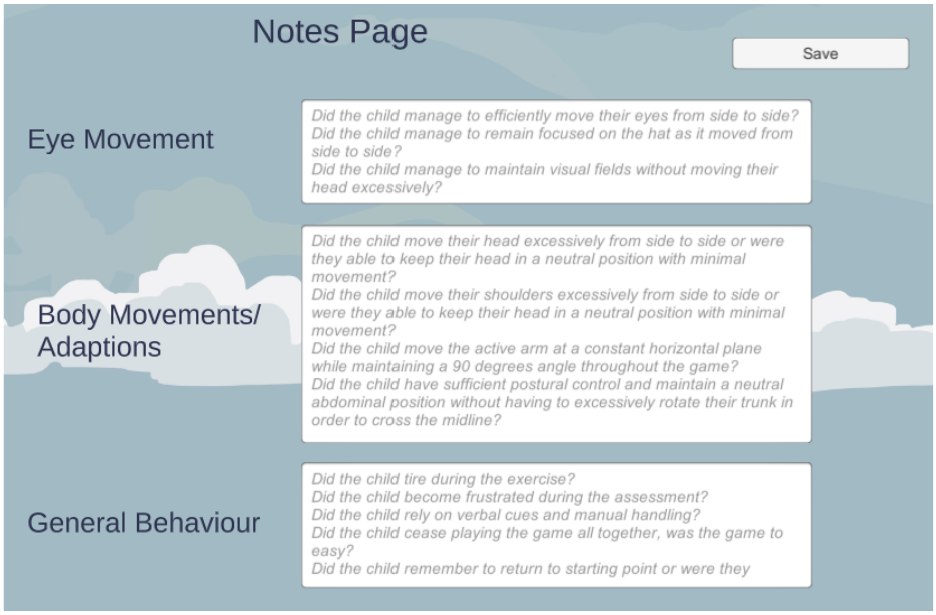


XSL•FO
RenderX

Stage 1, which is the assessment stage, aims to determine the child's gross ability to cross the midline. In addition to the stages incorporated into the game, an interface to input the specified variables determined by the therapist was introduced. These variables create a tailored game environment for the child. In stage 1, the speed and time variables remain constant, whereas the distance variable is incremented. The input device shown in Figure 4 controls the movement of the hat illustrated in Figure 6 from side to side, simulating the movement of crossing the midline while writing. After the child catches the ball, the hat needs to be moved all the way to the other side of the screen, ensuring that the midline is crossed.

A reference point for the child's baseline functioning, including their ability to cross the midline, was established in stage 1. The amount of external input required to achieve midline crossing would be noted by the caretaker on the system. Although the game is designed to determine the core impairments of crossing the midline, clinical reasoning, observations, and external guidance from the occupational therapist are vital for a holistic and accurate assessment. In the event that an occupational therapist is not performing the assessment (ie, caretaker, parent, teacher, or guardian), a notes page, shown in Figure 7, will need to be completed, which is presented to the caretaker at the end of the game. A guidance script with key points of reference will be provided to the caretaker before the child plays the game to assist and ensure accurate observations. Using the results, an individualized treatment plan tailored to the child's needs would be constructed and developed by the occupational therapist, which can be done remotely.

Figure 7. Notes page to fill in after each stage is completed.



Stage 2

Stage 2 takes the *starting distance* (X_0) and *distance variance* as inputs. Here, the *distance variance* is the distance from the *starting distance*, and the distance can be varied; for example, if the *starting distance* is 1.5 and the *distance variance* is 0.5, the ball can be spawned between points 1 and 2 on the x-axis.

Stage 2 is the first intervention stage, aimed at improving the child's gross ability to cross the midline. This stage allows for the distance variable to be adjusted, whereas the speed and time variables are constant. The caretaker sets the distance at the specific distance where the child began experiencing problems with midline crossing. This distance can then be adapted and graded during the intervention.

Stage 3

Stage 3 takes the *starting distance* (X_0), *initial speed* (Sp_0), and *incremental speed* (Is_p) as inputs. The *initial speed* is the speed at which the ball starts to fall. The *incremental speed* is the increase in speed from the starting point. The ball will continue to fall at the specified starting point; however, the speed at which the ball falls will increase as each ball is caught.

Stage 3, the second intervention stage, aims to enhance the child's gross ability to cross the midline. The distance and time variables will remain constant, whereas the speed variable will be incremented. This increases sufficient accuracy and skill.

Stage 4

Stage 4 does not require any input as the position at which each ball is spawned is random. The point at which each ball is spawned is calculated using the random function *Rand()*. As shown in Figure 6, the x-axis spans from –7.5 to 7.5; therefore, the ball can be spawned at a random point on the axis. The speed of the ball remains constant and does not increase as a function of the score. The gameplay time is constant for stage 4.

The purpose of stage 4, the unregulated round, is to maintain the skills the child has obtained in the previous stages. In this stage, the speed and time variables will remain constant, whereas the distance variable will be random.

Scoring

The occupational therapist advised on the scoring criteria for the stages during the collaborative design sessions. These scoring criteria were based on observations that occupational therapists would typically make during traditional treatments.

XSL•FO
RenderX

The first three criteria were generated from the gameplay data received from the game, and the following five criteria were recorded by observation. To confirm that the child did not display any form of pathology with midline crossing, a child would need to achieve a total of ≥4 points on their score card. It should be noted that all 8 criteria should be used in conjunction with one another and not assessed independently. Table 2 shows the 8 criteria, the observation being assessed, and the measurement to score the criterion for stage 1. If a child meets the scoring for a criterion, a point will be awarded. The criteria for the other stages are the same as those for stage 1, except for criterion 4. In the scoring criteria for stage 2, the effect of the random displacement of balls in a targeted area is determined rather than the increasing distance. In the scoring criteria for stage 3, the effect of increasing the speed of the balls is determined rather than increasing the distance. Finally, in the scoring criteria for stage 4, the effect of the balls randomly falling is determined.

Table 2. This table shows the scoring criteria for an assessment (stage 1).

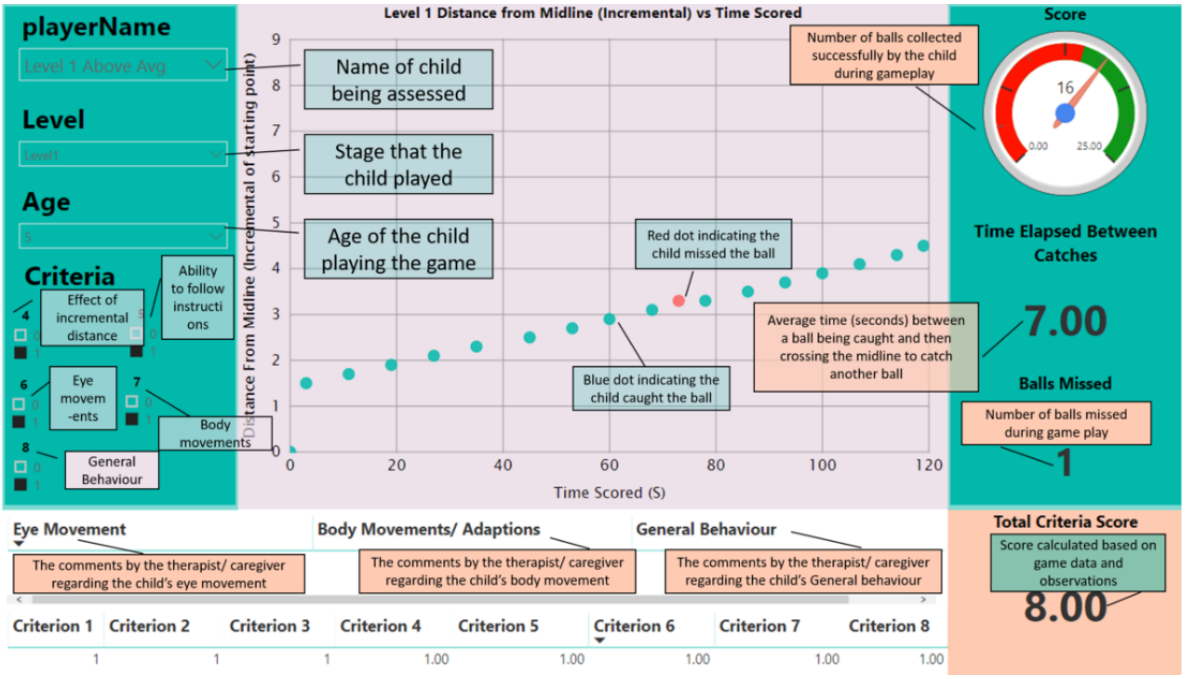
Criterion	Observation	Scoring
1	Balls caught	>14
2	Balls missed	<4
3	Average time between balls	<7.5
4	Distance increasing	No pattern
5	Follow instructions	Yes
6	Eye movements	Maintain visual focus
7	Body movement	No gross adjustments
8	General behavior	No verbal cues

Data Visualization

Occupational therapists can support children with difficulty crossing the midline by helping them develop the skills needed to perform activities of daily living. During therapy sessions, occupational therapists use a variety of techniques to support rehabilitation, such as functional electrical stimulation, constraint-induced motor training, facilitation, and virtual reality applications [27]. This combination of therapy, exercise, and context-specific retraining is critical for neuroplasticity, as mentioned in the *Introduction*. Although retrospective recall and exercise diaries can gather subjective data, the quality of these data is limited and relies on the notes and observations written by an occupational therapist; as such, occupational therapists lack objective data about the degree to which exercises have been performed. Consequently, a dashboard was designed to aid occupational therapists by providing a collection of objective data about children who have difficulty crossing the midline. The proposed monitoring tool provides a solution in which the gameplay of the child during an assessment or treatment is recorded, and the data collected in the game can be visualized by the therapist. The design of the system, and the monitoring functionality in particular, allows for the game and input device to be used by parents or guardians with the children in a home environment and to be used remotely. Therefore, the game would not be restricted to only being played during therapy sessions but rather can be played outside of these times as well. A dashboard was designed using Microsoft's PowerBI to display the data illustrated in Figure 8.

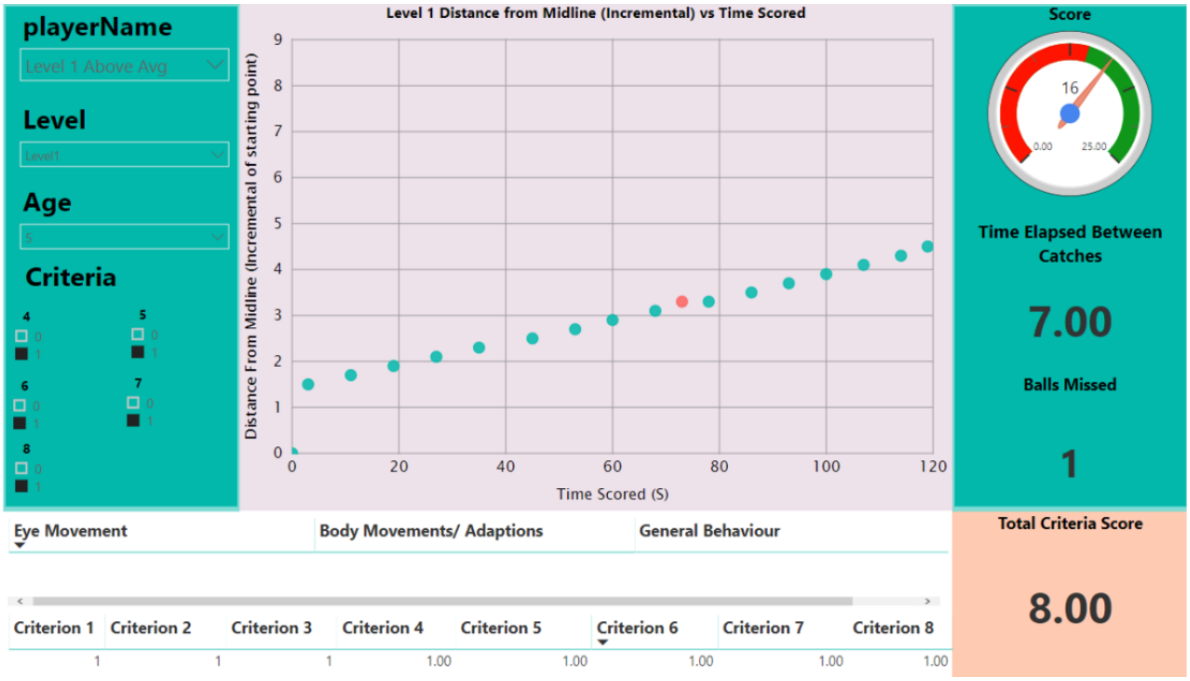
XSL•FO
RenderX

Figure 8. Representative dashboard showing how the tooltips will be displayed using Microsoft PowerBI.



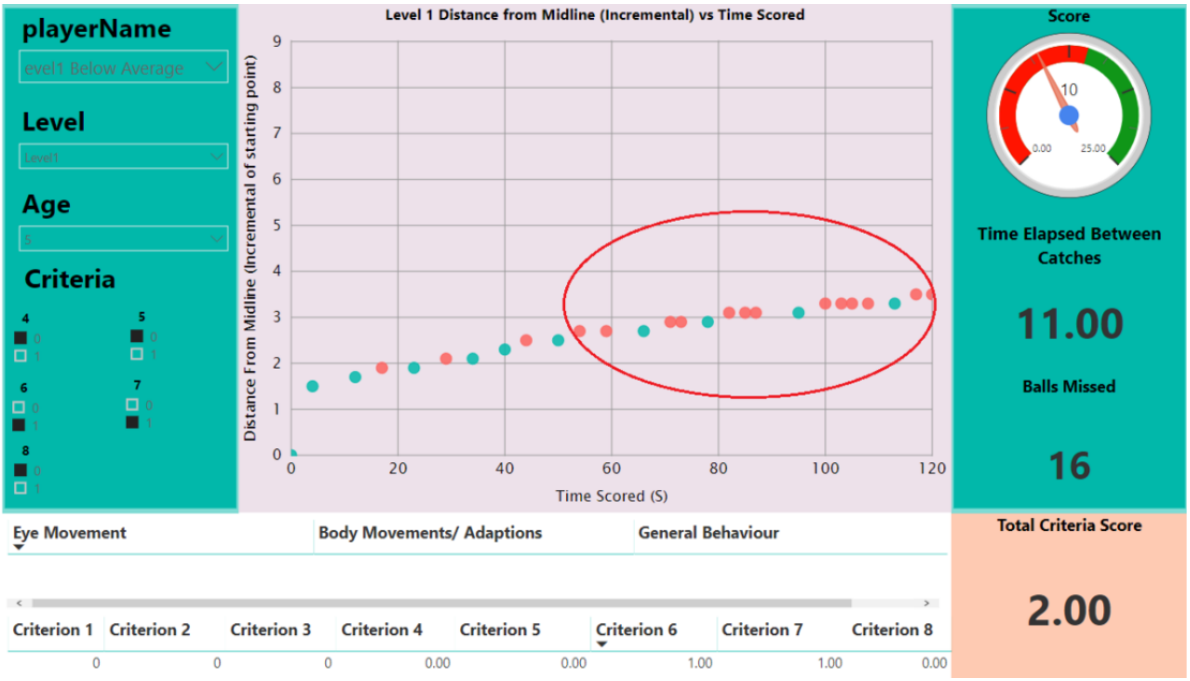
During the evaluation process, the occupational therapists were asked to play the game themselves and get familiar with how the system works. Figure 9 shows an illustrative example of how the dashboard would look for a child without midline crossing difficulties. The dashboard is accessible remotely. Figure 10 is an illustrative example of a dashboard for a child with midline crossing difficulties. The red circle indicates that as the distance increased, the *child* missed more balls. In this case, there are various factors that could impact the *child's* results. For example, the *child* may have felt bored and, therefore, did not play as they should have. A child with an average performance result needs to be further assessed to see whether their results are because of pathology with midline crossing or whether there are other factors yielding these results.

Figure 9. Representative dashboard for a child whose performance was above average for stage 1.



XSL•FO
RenderX

Figure 10. Representative dashboard for a child whose performance was below average for stage 1.



Objective Two: Evaluate Whether the Web-Based Game Is Fit for Purpose

The therapists commented on the convenience of integrating both assessment and treatment into the same application as it assists the therapists when grading a child. The stages developed in the game were also commended by the therapists, as each stage has a different focus point with a specific outcome. They collectively agreed that the quantitative aspect that the game creates by providing measurable and standardized data proves advantageous when compared with traditional methods of assessment and treatment. When using a traditional method, such as asking the child to build a puzzle or draw a line across a page, there are no quantitative measures that can be deduced; instead, the assessments are limited to mere observation. The proposed solution elicits not only quantitative data but also allows observational data to be recorded using the notes page.

Furthermore, they referred to the notes page as not only descriptive but also appreciated that it was designed to be understood by nonmedical professionals, thereby making the observational recording more effective and user-friendly. The notes page is viewed as an extremely beneficial aspect of the serious game as it ensures that the assessor covers all the observational components included in the assessment and treatment. As a result of the movements triggered by the game and the quantitative and observational data recorded by the application, the occupational therapists concluded that the serious game, accompanied by the input device, would assist them in identifying whether a child has pathology with midline crossing. Moreover, the solution would also assist in treating the child's pathology. Finally, the therapists thought that the solution could assist in exposing and identifying other gross and fine motor pathologies that a child might have, or it might even reveal underlying behavioral, cognitive, or physical impairments. It was suggested that the assessment stage should be standardized per age group.

One of the ways to identify whether a child has pathology with midline crossing is by observing whether the child has a midline jerk. This can be noticed when the child is performing actions where the midline is crossed, and their eyes move from side to side in a rapid manner. Consequently, the therapist proposed incorporating eye tracking to detect if the child's oculomotor functions are impaired. Additional monitoring through the use of a laptop camera can also be provided to observe the eye movement, head adjustments, and posture of the child.

As part of the game, when the child catches the object falling in the hat, the child is required to return the hat to the side of the screen to receive a point. Even though an alert is sounded when the object returns to the side, a suggestion was proposed to display a reminder on the screen to return the object to the side at the point where the child has caught the object. In addition, when the child is playing the game and controlling the input device, the child may release the device. Consequently, it was recommended that if the child releases the device during the game, an alert should be shown to remind the child to grasp the device.

Specific attention was given to the dashboard design as it is an important part of enabling remote monitoring. Everyone unanimously agreed that the dashboard's ability to track the progress of the child would benefit their reports immensely. The results represented in the dashboard would assist them in supporting and verifying their observational conclusions. Furthermore, when submitting reports to their clients' medical aids, the data from the dashboard can be used to solidify their conclusions and can be included in their reports.

The only disadvantage that was raised by the occupational therapists was that the dashboard might not be particularly easy to interpret for a layman; however, a suggestion was made to provide training to therapists on how to interpret and analyze the dashboard. The other disadvantage was that the dashboard was limited to only showing the results for a midline crossing

XSL•FO
RenderX

assessment and no other difficulties that could be identified by the game.

There were suggestions to provide a way of showing the standardized results for different categories by searching by age, for example, and, consequently, the results for that particular age group would be displayed. Introducing standardized norms would allow the child to be compared with other children within the specific area being examined. Another suggestion that was made was to allow therapists to add additional comments on the results of the child displayed for further reference. Finally, it was proposed that artificial intelligence could be incorporated so that when comments are made on the notes page by the therapist or guardian, the observations can be translated into quantitative data that can be used to calculate the scoring criteria instead of physically inputting the scores for the criteria that require observations.

The occupational therapists agreed that this solution could be used as a telehealth device. To further enhance these benefits, it was suggested that the game be deployed as a mobile app. This would allow for the game to be available on smart devices, which would, therefore, make it even more accessible than if it were only available on a laptop or computer. A second proposition was to record the gameplay so that it could be played back for reference. It was recommended that a video or chat capability be introduced so that the child or guardian could communicate with the therapist.

Discussion

Principal Findings

A telehealth system consisting of a serious game accompanied by an input device and a dashboard can be implemented to address children's midline crossing difficulties. The scoring system provides a quantitative aspect that proves advantageous when compared with traditional methods of assessment and treatment, where the assessment is limited to mere observation. In addition, the notes page that is completed at the end of the serious game ensures that the observational components that are vital in assessing a child are still included in the overall assessment. Therefore, the telehealth system elicits not only quantitative data but also allows observational data to be recorded using the notes page.

Although there are many input devices that can be used for therapeutic reasons, as described in the literature section, they each have shortcomings that would need to be addressed to make them more suitable when used by a child with a pathology with midline crossing. The common drawback of all the devices is their cost. It is not feasible to expect parents to buy expensive devices for treatment at home that will only be used for a limited time. A Kinect system will cost approximately US $399 and a Wii, US $164, whereas the proposed system costs approximately US $32 to manufacture. This price may also decrease if large quantities are manufactured.

The Wii provides a suitable solution for physical training and balance in particular; however, there are no games that explicitly aim to treat or assess children with midline crossing [18]. At first, Leap Motion seemed to be a suitable device for the development of a VT; however, with inaccurate readings stemming from the device affecting the game experience, the device would not be suitable for children [21]. The RealSense is a fitting tool; however, because of the high price of the device, it is not suitable [19].

Studies with a serious game and Kinect showed that individuals are more motivated, enjoy therapy, and even enable therapy to be more accessible [12]. On the basis of the feedback from the occupational therapists, the child will also experience these benefits from the system presented in this paper. However, this will only be proven in clinical studies.

Limitations

When determining whether serious games are feasible as a treatment option for midline crossing difficulties and are fit for purpose, there appears to be a lack of clinical evidence about the benefit to children from the application of serious games [28]. Owing to ethical constraints regarding testing the solution directly on children, professional opinions of occupational therapists were gathered to validate the solution. The next step would be to obtain the needed clearance to test directly on children without difficulties crossing the midline to attain a baseline. Thereafter, tests can be performed on children who have difficulty crossing the midline. Although the dashboard was configurable, it was found to be slightly difficult to interpret. It was suggested that therapists could be trained to interpret and analyze the dashboard. The telehealth system focuses on only one intervention. Therefore, the possibility of using the serious game and input device to assist in exposing and identifying other gross and fine motor difficulties that a child might have could be investigated.

Future Work

The therapists were confident that the telehealth system presented will assist them in identifying and treating children with midline crossing difficulties. Therefore, the possibility of using this solution to expose and identify other gross and fine motor difficulties that a child might have could be investigated. Furthermore, the solution could even reveal underlying behavioral, cognitive, or physical impairments. To clarify, the therapists would be able to identify other aspects, such as hand dominance, eye tracking, postural control, range of motion, attention and focus of the child, and the child's hand functions (grips and grasps) when playing the game. Furthermore, when identifying whether a child has midline crossing difficulties, a midline jerk can be observed. When a child is confronted with actions in which the midline is crossed, their eyes move from side to side rapidly. As a result, adding eye-tracking capabilities to the solution was put forward. To mitigate the concern raised by the therapists regarding the fact that an older child may achieve the game objectives quicker than a younger child, a standardized assessment stage per age group can be introduced.

One of the actions required in the game when a falling object is caught is to return the object to the side to receive a point. Currently, when the object is returned, an alert is sounded, indicating that a point is scored. A suggestion was proposed to display a reminder to return the object to the side. Moreover, when the child is playing the game using the input device, the

XSL•FO
RenderX

child may release the device. Consequently, it was suggested that in the event that the device is released during the game, an alert should be displayed to remind the child to remain holding the device. In addition, the results presented on the dashboard are limited to displaying the results for 1 player at a time. A recommendation was proposed to provide a way of showing standardized results for different categories, such as age. This would allow the child to be compared with other children within a specific area examined.

In future iterations, artificial intelligence can be incorporated so that when comments are made on the notes page by the therapist or guardian, the observations can be translated into quantitative data that can be used to calculate the scoring criteria instead of physically inputting the scores for the criteria that require observations. Finally, the solution proposed in this dissertation extracts gameplay data that can assist occupational therapists in identifying and treating children with difficulty crossing the midline. A greater number of occupational therapists adopting the solution in their practices means that more children will play the game and, ultimately, means more data are accumulated. These data can be analyzed using machine learning algorithms to find trends and enhance the assessment and treatment processes and visualizations produced.

Conclusions

A solution was designed to determine whether a telehealth system comprising a serious game can assist occupational therapists in identifying whether a child has pathology with midline crossing and can assist in treating the child remotely. Serious games are introduced as a need to meet objectives that go beyond entertainment and benefit the user in the area that needs to be mitigated. Through collaboration with occupational therapists, the telehealth system was designed to make use of different levels in the serious game, where each level addressed a different need of the therapy process. A novel, low-cost input device accompanies the serious game to track the movement of a child's hand from side to side and transform the position into the game. All results are saved on the web, and occupational therapists can access a dashboard that displays the results of each child. In addition, observed behavioral information will also be saved to assist occupational therapists in making decisions regarding changes to the intervention. During the interviews, occupational therapists indicated that the dashboard would support their treatment plan and that the end-to-end solution was indeed feasible.

Acknowledgments

The authors would like to thank all the occupational therapists who took part in this study.

Conflicts of Interest

None declared.

References

1. Dahl-Popolizio S, Carpenter H, Coronado M, Popolizio NJ, Swanson C. Telehealth for the provision of occupational therapy: reflections on experiences during the COVID-19 pandemic. Int J Telerehabil 2020 Dec 08;12(2):77-92 [FREE Full text] [doi: 10.5195/ijt.2020.6328] [Medline: 33520097]
2. Zylstra SE. Evidence for the use of telehealth in pediatric occupational therapy. J Occup Ther Sch Early Interv 2013 Dec 18;6(4):326-355. [doi: 10.1080/19411243.2013.860765]
3. Carson J. Telehealth and occupational therapy: integral to the triple aim of health care reform. Am J Occup Ther 2018;69(2):6902090010p1-6902090010p8. [doi: 10.5014/ajot.2015.692003] [Medline: 26122676]
4. Cermak SA, Quintero EJ, Cohen PM. Developmental age trends in crossing the body midline in normal children. Am J Occup Ther 1980 May;34(5):313-319. [doi: 10.5014/ajot.34.5.313] [Medline: 7405768]
5. Buchele Harris H, Cortina KS, Templin T, Colabianchi N, Chen W. Impact of coordinated-bilateral physical activities on attention and concentration in school-aged children. Biomed Res Int 2018 May 28;2018:2539748 [FREE Full text] [doi: 10.1155/2018/2539748] [Medline: 29998131]
6. Devi L. Child Development An Introduction. New Delhi: Institute for Sustainable Development Lucknow and Anmol Publications; 1998.
7. Payne S. Standardised tests: an appropriate way to measure the outcome of paediatric occupational therapy? Br J Occup Ther 2002 Mar 1;65(3):117-122. [doi: 10.1177/030802260206500303]
8. Crossing the Body's Midline. Kids Sense. URL: https://childdevelopment.com.au/areas-of-concern/fine-motor-skills/crossing-the-bodys-midline/ [accessed 2020-05-20]
9. Held Jr LI. Mysteries of the midline. In: Quirks of Human Anatomy: An Evo-Devo Look at the Human Body. Cambridge, United Kingdom: Cambridge University Press; 2009.
10. Masselink CE. Considering technology in the occupational therapy practice framework. Open J Occup Ther 2018 Jul 01;6(3). [doi: 10.15453/2168-6408.1497]
11. Jia A, Qi T, Ching SL, Ying LS. Serious games in nursing education: an integrative review. In: Proceedings of the 2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games). 2017 Presented at: 2017

XSL•FO
RenderX

9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games); Sep 6-8, 2017; Athens, Greece. [doi: 10.1109/vs-games.2017.8056599]
12. Muñoz AP, Guerra PI, Bykbaev YR. New approach of serious games in ludic complements created for rehabilitation therapies in children with disabilities using kinect. In: Proceedings of the 2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON). 2018 Presented at: Proceedings in the IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON); Aug 8-10, 2018; Lima , Peru. [doi: 10.1109/intercon.2018.8526464]
13. Tresser S. Case study: using a novel virtual reality computer game for occupational therapy intervention. Presence (Camb) 2012 Aug;21(3):359-371. [doi: 10.1162/pres_a_00118]
14. Matthew L, Theresa D. At the heart of it all: the concept of presence. J Comput Mediat Commun 1997 Sep 1;3(2):1-2. [doi: 10.1111/j.1083-6101.1997.tb00072.x]
15. Ruiz A, Cantos MM. Telematics application for therapeutic use using Kinect for Windows and SDK v. 1.6. 1library. 2012. URL: https://1library.co/document/zx9knodz-aplicacion-telematica-uso-terapeutico-utilizando-kinect-windows-sdk.html [accessed 2020-05-20]
16. Chang Y, Han W, Tsai Y. A Kinect-based upper limb rehabilitation system to assist people with cerebral palsy. Res Dev Disabil 2013 Nov;34(11):3654-3659. [doi: 10.1016/j.ridd.2013.08.021] [Medline: 24012594]
17. Chen PJ, Du YC, Shih CB, Yang LC, Lin HT, Fan SC. Development of an upper limb rehabilitation system using inertial movement units and kinect device. : IEEE; 2016 Presented at: 2016 International Conference on Advanced Materials for Science and Engineering (ICAMSE); November 12-13, 2016; Tainan, Taiwan. [doi: 10.1109/icamse.2016.7840309]
18. Addio GD, Iuppariello L, Gallo F, Bifulco P, Cesarelli M, Lanzillo B. Comparison between clinical and instrumental assessing using Wii Fit system on balance control. In: Proceedings of the 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA). 2014 Presented at: Proceedings in the IEEE International Symposium on Medical Measurements and Applications; Jun 11-12, 2014; Lisbon, Portugal. [doi: 10.1109/MeMeA.2014.6860124]
19. Hariadi RR, Herumurti D, Yuniarti A, Kuswardayan I, Suciati N, Mooy T. Virtual Sasando using leap motion controller. 2017 Presented at: 2017 International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA); October 12-14 , 2017; Surabaya, Indonesia. [doi: 10.1109/icamimia.2017.8387578]
20. Maggie S, Ahmed E, Dirk R. Development of a virtual coach scenario for hand therapy using LEAP motion. 2016 Dec 06 Presented at: Future Technologies Conference (FTC); 2016; San Francisco, CA, USA. [doi: 10.1109/FTC.2016.7821736]
21. Julio C, Yaroslava RB, John AD, Tatiana C, Vladimir RB. Serious game to improve fine motor skills using Leap Motion. 2018 Presented at: Proceedings in the IEEE Congreso Argentino de Ciencias de la Informatica y Desarrollos de Investigacion (CACIDI); 2018; Argentina p. 1-5. [doi: 10.1109/CACIDI.2018.8584370]
22. Chhor J, Yun Gong Y, Rau PLP. Design and Evaluation of a Serious Game for Health Employing Intel RealSense. 2017 Presented at: Proceeding in the Springer International Publishing Cross-Cultural Design conference; July 9-14 , 2017; Vancover, Canada. [doi: 10.1007/978-3-319-57931-3_42]
23. Annett M, Anderson F, Goertzen D, Halton J, Ranson Q, Bischof WF, et al. Using a multitouch table top for upper extremity motor rehabilitation. In: Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7. 2009 Presented at: OZCHI '09: Proceedings of the 21st conference of the computer-human interaction special interest group of Australia on Computer-human interaction: design; Nov 23-27, 2009; Melbourne Australia p. 261-264. [doi: 10.1145/1738826.1738869]
24. Jacobs J, Greeff JR. Interactive game for children with difficulty crossing the midline. In: Proceedings of the 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH). 2018 Presented at: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH 2018); May 16-18, 2018; Vienna, Austria. [doi: 10.1109/segah.2018.8401348]
25. Jacobs J, Greeff JR, Heymann R. Input device to assist occupational therapists treat children with difficulty crossing the midline. In: Proceedings of the 2019 IEEE AFRICON. 2019 Presented at: Proceedings of the IEEE Region 8 flagship conference in Africa (Africon)-2019; Sep 25-27, 2019; Accra, Ghana. [doi: 10.1109/africon46755.2019.9133800]
26. Kultima A, Kuittinen J, Niemala J, Paavilainen J. Casual games discussion. In: Proceedings of the 2007 conference on Future Play.: Nov 14 - 17, 2007; 2007 Presented at: Proceedings of the 2007 conference on Future Play; 2007; Toronto Canada. [doi: 10.1145/1328202.1328221]
27. Basteris A, Nijenhuis SM, Stienen AH, Buurke JH, Prange GB, Amirabdollahian F. Training modalities in robot-mediated upper limb rehabilitation in stroke: a framework for classification based on a systematic review. J Neuroeng Rehabil 2014 Jul 10;11:111 [FREE Full text] [doi: 10.1186/1743-0003-11-111] [Medline: 25012864]
28. Kokol P, Vošner HB, Završnik J, Vermeulen J, Shohieb S, Peinemann F. Serious game-based intervention for children with developmental disabilities. Curr Pediatr Rev 2020;16(1):26-32. [doi: 10.2174/1573396315666190808115238] [Medline: 31393252]

Abbreviations

IMU: inertial measurement unit

XSL•FO
RenderX

VT: virtual therapist

Edited by N Zary; submitted 05.02.21; peer-reviewed by I Manuaba, M Stein; comments to author 19.03.21; revised version received 11.05.21; accepted 19.06.21; published 01.11.21.

Please cite as:
Jacobs J, Heymann R, Greeff JJ
A Telehealth System Incorporating a Serious Game Intervention to Aid Occupational Therapists in Identifying and Treating Children With Difficulty Crossing the Body's Midline: Key Informant Interviews Among Occupational Therapists
JMIR Serious Games 2021;9(4):e27761
URL: https://games.jmir.org/2021/4/e27761
doi:10.2196/27761
PMID:34723821

©Jonathan Jacobs, Reolyn Heymann, Jacob Jacobus Greeff. Originally published in JMIR Serious Games (https://games.jmir.org), 01.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Feasibility of a Sensor-Controlled Digital Game for Heart Failure Self-management: Randomized Controlled Trial

Kavita Radhakrishnan[1], PhD; Christine Julien[2], PhD; Tom Baranowski[3], PhD; Matthew O'Hair[4], MA; Grace Lee[2], MS; Atami Sagna De Main[1], MSN; Catherine Allen[1], MPH; Bindu Viswanathan[5], PhD; Edison Thomaz[2], PhD; Miyong Kim[1], PhD

[1]School of Nursing, The University of Texas Austin, Austin, TX, United States

[2]Department of Electrical and Computer Engineering, Cockrell School of Engineering, The University of Texas Austin, Austin, TX, United States

[3]Baylor College of Medicine, Houston, TX, United States

[4]Good Life Games, Inc, Austin, TX, United States

[5]Department of Statistics and Data Sciences, The University of Texas Austin, Austin, TX, United States

Corresponding Author:
Kavita Radhakrishnan, PhD
School of Nursing
The University of Texas Austin
1710 Red River St
Austin, TX, 78701
United States
Phone: 1 512 471 7936
Email: Kradhakrishnan@mail.nur.utexas.edu

Abstract

Background: Poor self-management of heart failure (HF) contributes to devastating health consequences. Our innovative sensor-controlled digital game (SCDG) integrates data from sensors to trigger game rewards, progress, and feedback based on the real-time behaviors of individuals with HF.

Objective: The aim of this study is to compare daily weight monitoring and physical activity behavior adherence by older adults using an SCDG intervention versus a sensors-only intervention in a feasibility randomized controlled trial.

Methods: English-speaking adults with HF aged 55 years or older who owned a smartphone and could walk unassisted were recruited from Texas and Oklahoma from November 2019 to August 2020. Both groups were given activity trackers and smart weighing scales to track behaviors for 12 weeks. The feasibility outcomes of recruitment, retention, intervention engagement, and satisfaction were assessed. In addition to daily weight monitoring and physical activity adherence, the participants' knowledge, functional status, quality of life, self-reported HF behaviors, motivation to engage in behaviors, and HF-related hospitalization were also compared between the groups at baseline and at 6, 12, and 24 weeks.

Results: A total of 38 participants with HF—intervention group (IG; 19/38, 50%) and control group (CG; 19/38, 50%)—were enrolled in the study. Of the 38 participants, 18 (47%) were women, 18 (47%) were aged 65 years or older, 21 (55%) had been hospitalized with HF in the past 6 months, and 29 (76%) were White. Furthermore, of these 38 participants, 31 (82%)—IG (15/19, 79%) and CG (16/19, 84%)—had both weight monitoring and physical activity data at the end of 12 weeks, and 27 (71%)—IG (14/19, 74%) and CG (13/19, 68%)—participated in follow-up assessments at 24 weeks. For the IG participants who installed the SCDG app (15/19, 79%), the number of days each player opened the game app was strongly associated with the number of days the player engaged in weight monitoring (r=0.72; P=.04) and the number of days with physical activity step data (r=0.9; P<.001). The IG participants who completed the satisfaction survey (13/19, 68%) reported that the SCDG was easy to use. Trends of improvement in daily weight monitoring and physical activity in the IG, as well as within-group improvements in HF functional status, quality of life, knowledge, self-efficacy, and HF hospitalization in both groups, were observed in this feasibility trial.

Conclusions: Playing an SCDG on smartphones was feasible and acceptable for older adults with HF for motivating daily weight monitoring and physical activity. A larger efficacy trial of the SCDG intervention will be needed to validate trends of improvement in daily weight monitoring and physical activity behaviors.

Trial Registration: ClinicalTrials.gov NCT03947983; https://clinicaltrials.gov/ct2/show/NCT03947983

XSL•FO
RenderX

(JMIR Serious Games 2021;9(4):e29044) doi:10.2196/29044

KEYWORDS

heart failure; digital game; sensor; self-management; older adults; weight monitoring; physical activity; behaviors; mobile phone

Introduction

Need for Heart Failure Self-management Behaviors

Despite significant advances in treatment and management, heart failure (HF) continues to be the leading cause of hospitalization among older adults in the United States [1,2], with an estimated annual cost of US $32 billion [1]. Improved self-management behaviors, defined as "behaviors that maintain physiological stability" and enable "response to symptoms when they occur," can help reduce the adverse effects of an HF diagnosis [3,4]. Daily weight monitoring is such a behavior; weight gain is typically the first sign of volume overload in patients with HF, and if weight gain is treated promptly, clinically significant HF exacerbations can be avoided [5,6]. Similarly, physical activity improves myocardial function and functional capacity [7] and reduces depressive symptoms [8]. Engagement in these 2 critical HF self-management behaviors improves quality of life and reduces health care use [9,10]. Yet, weight monitoring and physical activity show significantly poorer adherence than other HF self-management behaviors [11,12] because of poor knowledge and lack of motivation [13].

Role of Innovative Digital Health Interventions for Motivating HF Self-management Behaviors

Recent advances in wearable technology, digital sensor devices, and mobile health (mHealth) apps have allowed application of these technologies to observe and motivate health behaviors in real-world conditions. Yet, despite the advantages of portability and scalability offered by digital devices and apps, long-term adoption rates remain low. There have been many instances of abandonment of the use of activity trackers in prior research, with more than 50% of the activity tracker users stopping use within a period of 2 weeks to 6 months [14,15]. Long-term maintenance rates of such devices by patients with HF are yet to be explored.

Digital games that serve as affordable, portable, scalable tools while being enjoyable and easy to use can help contextualize health behaviors [16] and motivate clinically significant behavior changes, thereby improving health outcomes [17]. Digital games can combine appealing stories [18] and interests and hobbies [19] with active learning [20], incentives [21], and social connections [22] to offer an accessible, engaging, and immersive habit-forming medium [23] while objectively measuring these behaviors [24,25]. Well-designed digital games have significantly improved physical (eg, balance and mobility) and cognitive (eg, processing speed) health outcomes [26] as well as behavioral outcomes for chronic diseases (eg, dietary changes for diabetes) [17,27]. Digital game-playing is increasing among older adults; the number of adults aged 50 years or older who enjoy video games in any format increased by 11 million from 2016 to 2019 [28], suggesting a receptiveness to playing games for health.

Digital games incorporating gamification principles such as competition, leaderboards, and incentives have improved learning outcomes and have been highly acceptable among participants with HF [29,30], but these games do not involve data on real-time HF self-management behaviors. In a sensor-controlled digital game (SCDG), data on behaviors from sensors, including those in wearable devices, are synchronized with a mobile gaming app to trigger game progress, rewards, personalized and contextually relevant feedback (eg, reduce fluid intake or call physician for weight gain), and incentives based on participants' real-time behaviors [24]. Combining digital games and sensors thus offers a powerful way to improve behavior adherence and potentially make health care participatory, personalized, predictive, and preventive as defined by the Precision Medicine Initiative [31].

Thus, the primary goals of this study are to obtain preliminary efficacy data on behavior adherence and undertake a comprehensive feasibility assessment of an SCDG intervention called Heart Health Mountain. This app synchronizes with a Bluetooth-enabled weighing scale and activity tracker to activate game rewards and feedback based on the real-time weight monitoring and exercise behaviors of older adult participants with HF.

Methods

Study Design and Scope

We conducted a prospective feasibility randomized controlled trial (1:1) with 2 parallel groups (sensors only or sensors plus SCDG app) from November 2019 to January 2021 (ClinicalTrials.gov: NCT03947983). The results presented here are from the 12-week SCDG intervention measuring objective behaviors of daily weight monitoring and physical activity and the 24-week follow-up measuring quality of life, functional status, HF self-management efficacy, and HF hospitalizations.

Study Population and Recruitment

Before the COVID-19 outbreak, we identified potential participants through chart review at a cardiac rehabilitation center as well as through referrals by clinical case managers at an inpatient cardiac floor in central Texas. Research staff members provided information on the study to adults who were aged 55 years or older, English-speaking, diagnosed with HF, and classified according to the New York Heart Association's HF classification as class II or III [32] during their inpatient stay or outpatient visit to the cardiac center. Other eligibility criteria included smartphone ownership as a proxy measure to indicate prior familiarity with smartphone use, ability to independently walk without a walker or human assistance, and a score of 4 or higher on the Mini-Cog [33] cognitive screen. The exclusion criteria included severe visual or tactile impairments (eg, legal blindness or severe arthritis), which would prevent the use of a smartphone, or end-stage renal failure

XSL•FO
RenderX

or terminal illness (eg, cancer), both of which adversely affect HF prognosis [34].

To continue the trial during the COVID-19 pandemic, all in-person interactions were converted to remote interactions. We used a secure, Health Insurance Portability and Accountability Act–compliant email system to continue receiving referrals from clinical case managers at the inpatient HF center. In addition, we contracted with the participant recruitment company Trialfacts [35] to recruit participants on the web from the states of Texas and Oklahoma. The eligibility criteria remained the same, but the formal screening process then required a narrative description of HF history or confirmation from the potential participant's health care provider to confirm the individual's HF diagnosis.

Intervention

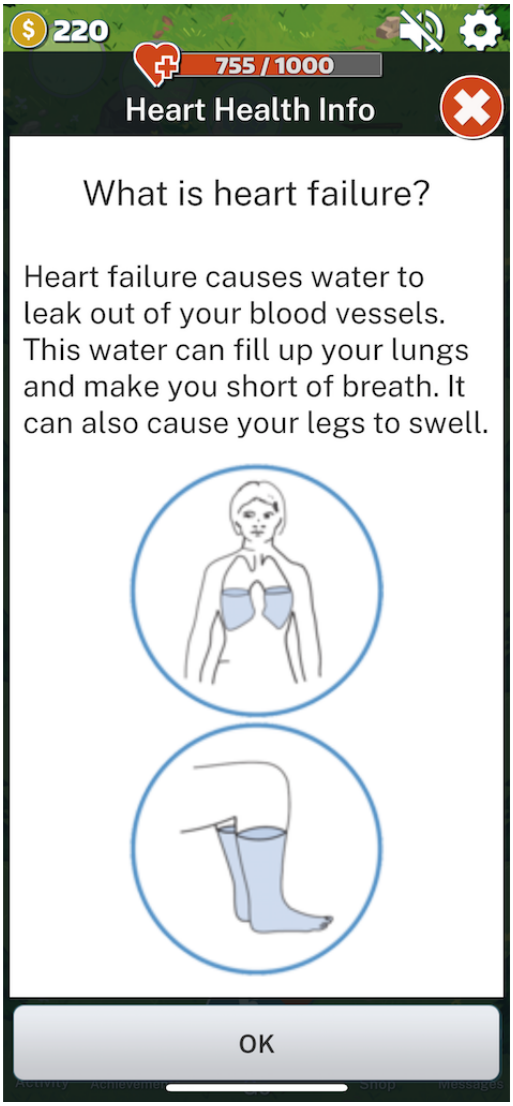
The description of the design, development, and usability assessment phase of the SCDG has been detailed elsewhere [25]. Guided by the Fogg behavioral model [36] and played on a smartphone, the *Heart Health Mountain* SCDG presented a narrative in which the older adult player helps an avatar to climb a mountain in a forested area, with the game's goal being to avoid hospitalization. The SCDG motivates the player's engagement in critical behaviors related to HF self-management, helping the avatar to climb the mountain. This is done by daily weight monitoring and attaining physical activity steps based on goal steps tailored to the individual player. The step goal in the SCDG ranged from 3000 to 15,000 steps.

According to the Fogg model [36], for a person to perform a target health behavior, they must (1) be sufficiently *motivated*, (2) have the *ability* to perform the behavior, and (3) be *triggered* to perform the behavior. For example, based on the *motivation* concept of the Fogg model [36], positive attitudes toward engaging in HF self-management and empathy with the game character will be aided by positive feedback, competition (eg, leaderboards), and game rewards (eg, coins to buy low-salt food items or accessories for the game character for engaging in real-time behaviors and in-game challenges). On the basis of the *ability* concept of the Fogg model [36], knowledge about HF self-management within the SCDG was provided using language especially chosen for those with low literacy levels (from the *Living Well With Heart Failure* booklet) [37,38] in bite-sized chunks (Figure 1). Built-in quizzes tested content mastery, and problem-solving strategies provided opportunities for higher game rewards (Figure 2).

Finally, based on the *trigger* concept of the Fogg model [36], game alerts, the avatar's health status, messages (Figure 3), and incentives were tailored to the participant's real-time HF self-management behaviors based on data from the behavior-tracking sensors (eg, climbing up the mountain by 2 steps if the physical activity goal was attained that day). HF game players are expected to make meaningful connections between game events and real-time HF self-management behaviors, increasing the likelihood that they retain and apply their newly acquired knowledge, skills, and habits for HF self-management behaviors outside the game's world [39] and improve their health outcomes.

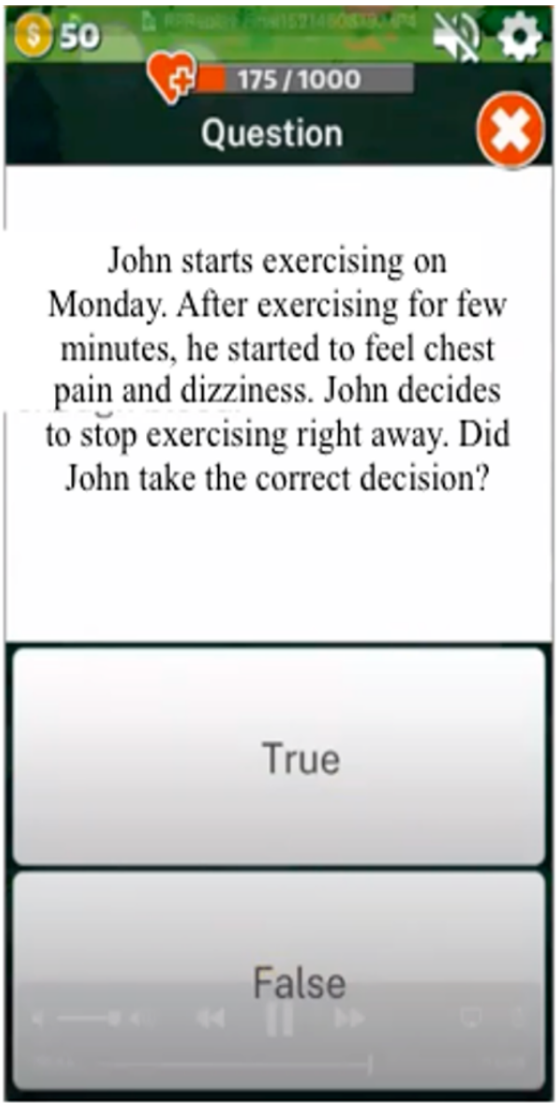
XSL•FO
RenderX

Figure 1. Knowledge embedded in the sensor-controlled digital game.



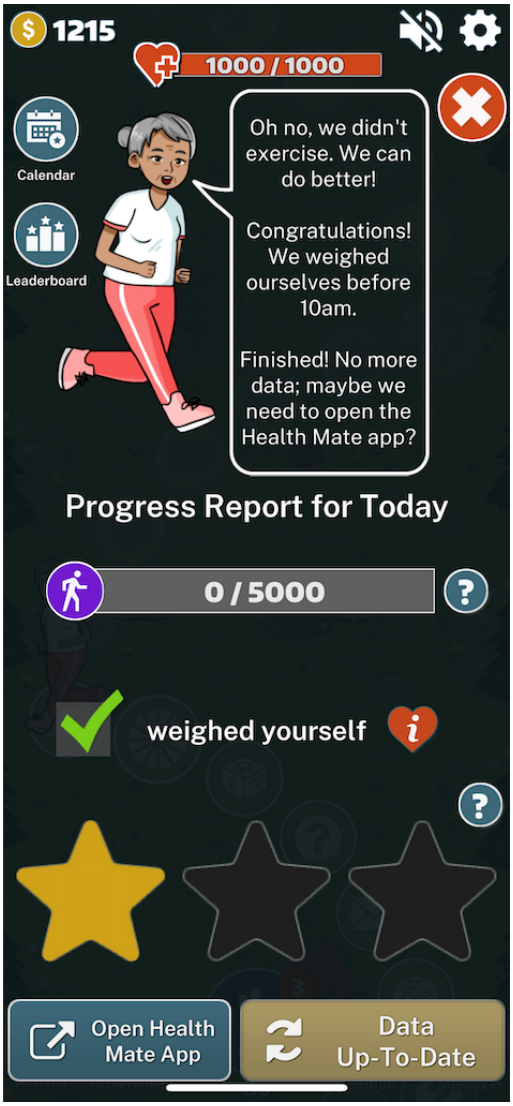
XSL•FO
RenderX

Figure 2. Built-in quiz in the sensor-controlled digital game.



XSL•FO
RenderX

Figure 3. Example of game alert messages based on real-time behavior.



Intervention Procedures

CONSORT (Consolidated Standards of Reporting Trials) guidelines [40] informed this parallel-group randomized controlled trial and the reporting of outcomes (Multimedia Appendix 1). The SCDG intervention group (IG) received sensors tracking weight monitoring and physical activity and played the SCDG app on a smartphone; the sensors-only control group (CG) received sensors tracking weight monitoring and physical activity only. Both IG and CG participants were given the Withings Go activity tracker (Withings) [41], Withings Body smart weighing scale [42], and Withings Health Mate app [43] to record, store, and transmit daily weight and physical activity data, but the CG did not receive the SCDG app. The Withings Body smart weighing scales have been found to accurately measure body mass [44]. We selected the Withings Go tracker because it has long battery life, is waterproof, is reasonably accurate in measuring activity levels of people with HF [45,46], and uses a clock schema, which is a familiar interface for most adults, to represent the user's daily physical activity. The step goals for all participants were set based on their preferences as well as the nursing research assistant's assessment of their physical health status.

Whereas the participants in the IG received standardized HF education [6,38] embedded within the SCDG, the participants in the CG received the same information in written format. Thus, the difference between the IG and the CG was the receipt of the SCDG gaming elements. The SCDG app transmitted the IG participants' game-playing data, and the Health Mate app transmitted all participants' sensor data to the research team through cellular data service or home Wi-Fi. An intervention nursing research assistant remotely tracked user engagement for participants in both groups during the first week and contacted participants by phone in the absence of any data to offer help with troubleshooting.

Before the COVID-19 outbreak, an intervention nursing research assistant installed the apps on all participants' smartphones and trained the participants to use the sensor devices and apps at either their homes or at a location of their convenience. During the COVID-19 pandemic, in-person interactions were converted to remote interactions. Contactless delivery of study equipment, videoconferencing on smartphones, and printed pictorial and video instructions allowed remote support for installation, training, and troubleshooting of the devices and apps. The study team members' experience with the devices and apps during in-person interactions with the older adults at their homes helped

XSL•FO
RenderX

inform the training materials that were provided to the participants during the pandemic. Installation and training times by the research team member varied widely during the remote phase, from 0 minutes (participant self-installation) to 180 minutes compared with 45 to 120 minutes during the in-person phase.

Randomization

The participants were assigned to the IG or CG with a 1:1 randomization ratio such that the 2 groups were equivalent in terms of biological sex. The randomization was carried out after informed consent was obtained and the baseline survey completed. The allocation sequence list for block randomization was generated by an independent researcher at the Sealed Envelope Ltd website [47] with random block sizes of 2 and 4 and concealed until the trial group was assigned. To ensure blinding during assessment of the outcomes, the research assistant who delivered the intervention to the 2 groups was different from the research assistant who collected the baseline and follow-up surveys from the 2 groups. The researcher who performed the data analysis was blinded to the participant groups. Indication of participation in a digital health intervention study in the informed consent allowed us to blind all participants to their specific group assignment.

Outcome Measures

Feasibility Outcomes

To inform the development of a future large-scale randomized controlled clinical trial, we determined the feasibility, acceptability, and effectiveness of our SCDG in engaging older adults with HF by assessing recruitment (how many accepted the invitation to participate in the study), retention, engagement, and satisfaction. We considered the trial feasible if the enrollment rate was at least 20% of the patients with HF who were approached to participate in the study, if at least 50% of the IG and CG participants completed both the 12-week objective behavior assessment and the 24-week follow-up assessment (retention), and if at least 50% of the IG participants used the SCDG intervention for at least 50% of the days (engagement) and were satisfied with the SCDG intervention (satisfaction).

Retention

Retention was recorded as the number (proportion) of participants in both groups who used the sensor devices to measure physical activity and weight monitoring for 12 weeks and completed follow-up assessments at 24 weeks to assess maintenance of the SCDG intervention.

Engagement and Adherence

A Google Cloud console was created to store all participants' deidentified physical activity and weight monitoring behavioral data obtained from the Withings Health Mate app as well as data used to assess the IG participants' engagement with the SCDG and its features. The number of days of adherence to physical activity and weight monitoring as well as engagement with the SCDG were obtained from the Google Cloud console.

Satisfaction

We assessed the IG participants' satisfaction with the SCDG using a questionnaire at the end of 12 weeks based on the 4-item interest and enjoyment subscale of the Intrinsic Motivation Inventory (α=.8) [48], a multidimensional instrument intended to assess participants' subjective experience related to a target activity. The interest and enjoyment subscale assesses an individual's interest and inherent pleasure in performing a specific activity on a 4-point Likert scale (1=not satisfied, 4=very satisfied). In addition, the participants were asked open-ended questions about what they liked and disliked most about using the SCDG for HF self-management. A 4-point scale (1=strongly disagree, 4=strongly agree) was used to assess the key elements that might have helped the participants maintain their motivation to continue using the SCDG. Finally, the participants were asked whether they would recommend the SCDG to others with HF.

Health and Behavior Outcomes

Overview

All participants were asked to complete surveys at 6, 12, and 24 weeks after the baseline survey to assess the immediate effect of the intervention and maintenance of behavioral changes. Before the COVID-19 outbreak, the participants were asked to complete the surveys on the Qualtrics platform (Qualtrics XM) on a study iPad (Apple Inc). During the pandemic, the participants were emailed or texted links to the Qualtrics surveys. All participants completed a sociodemographic survey questionnaire at baseline. They also completed the 2-item Patient Health Questionnaire [49] to report on depressive symptoms at baseline and at 6, 12, and 24 weeks.

Weight Monitoring Behavior

The primary outcome of the days of weight monitoring was measured by the number of days with weight monitoring data. This measure was collected from sensor logs within the Health Mate app [43], with each day of weighing measured dichotomously (yes or no).

Physical Activity Behavior

Physical activity data were derived from the Withings Go sensor logs within the Health Mate app [43] by obtaining the cumulative steps for each day and averaging the steps for each participant over 6 and 12 weeks.

HF-Related Functional Status

For this measure, we used items 1-12 from the Kansas City Cardiomyopathy Questionnaire (KCCQ) [50]. The item scores are transformed to a range of 0 to 100 by subtracting the lowest possible scale score, dividing by the range of the scale, and multiplying by 100. The summary scores for the functional status thus range from 0 to 100, with higher scores indicating better functional status. The Cronbach α value for the KCCQ functional status is excellent at .93 [49]. The scale has demonstrated criterion validity, with high correlation with the New York Heart Association HF classification and the 6-minute Walk Test [50].

XSL•FO
RenderX

Quality of Life

Quality of life was measured with items 13-15 from the KCCQ [50]. The item scores are transformed to a range of 0 to 100 by subtracting the lowest possible scale score, dividing by the range of the scale, and multiplying by 100. The summary values for the quality-of-life domain thus range from 0 to 100, with higher scores indicating better quality of life. The Cronbach α value was .78 in prior studies [50]. In comparison with similar quality-of-life instruments, the KCCQ is sensitive to clinical changes in HF and has shown a significant and high correlation with the New York Heart Association HF classification [50].

HF Self-management Knowledge

The instrument used was the 30-item Atlanta Heart Failure Knowledge Test [51]. Each correct answer is scored as 1, with no additional weighting of items; the correct responses are then summed. Incorrect or skipped questions are scored as 0. The total scores range from 0 to 30, and higher scores indicate better HF knowledge. The content validity ratings for relevance and clarity ranged from 0.55 to 1.0, with 81% of the items rated 0.88 to 1.0. The Cronbach α value for reliability was .84 [51]. Construct validity has been demonstrated by directly correlating knowledge with clinical and self-care outcomes, including dietary sodium consumption, medication adherence, and health care use [51].

Self-reported HF Self-care Behaviors

The instrument used was the 9-item European Heart Failure Self-care Behavior Scale [52]. The items were scored on a Likert scale of 1 to 5. For calculating the standardized score, each item was reverse coded and then computed using the following formula: (sum of all even-coded items minus 9) times 2.7777. The standardized scores range from 0 to 100; every item is given equal weight, with a higher score indicating better self-care. The coefficient α value was .80 for this instrument [52].

HF Self-efficacy

The instrument used was the 6-item Section C (self-efficacy section) of the Self-Care of Heart Failure Index [53]. The item scores were standardized using the following formula: (sum of Section C items minus 6) times 5.56. The scores on the self-efficacy scale range from 0 to 100, with higher scores reflecting better self-efficacy. The Cronbach α value was .88, with good evidence for construct validity and contrasting group validity [53].

Motivation for HF Self-management Behaviors

The instrument used was the 19-item Treatment Self-Regulation Questionnaire [54], adapted for HF. Each item was scored on a scale of 1 to 7. The scale has 2 subscales: autonomous regulation and controlled regulation. The score for each subscale was the average of the items in that subscale. The average for controlled regulation was subtracted from the average for autonomous regulation to calculate the Relative Autonomy Index. The Cronbach α value for chronic disease–related behaviors ranged from .83 to .87 [54]. Construct and convergent validity have been confirmed using factor analysis and correlations with autonomy-related subscales of other behavior scales and health outcomes [55].

HF Hospitalization

This measure was obtained through the participants' self-report (yes or no) on the 6-, 12-, and 24-week surveys and was confirmed with hospitalization discharge summaries or communication with health care providers.

Statistical Analysis

Power

On the basis of 1:1 randomization with 80% power (α=5%), a sample size of 38 patients per group would have been required to detect a difference of 80% versus 50% in both groups of daily monitoring of weight monitoring and physical activity. We chose 80% as the cutoff for adequate weight monitoring at 12 weeks because patients with HF who completed *at least* 80% of the weight diaries (5.6/7 of the days per week) were found to have significantly reduced odds for HF-related hospitalizations in comparison with patients who completed less than 80% of the weight diaries [6]. In addition, only 50% of the participants with HF in a remote monitoring sensor group recorded their weights more than 50% of the time [56]. Allowing for 10% attrition, 49 patients per group (N=98) would have had to be recruited for a fully powered clinical trial of the SCDG intervention for HF behavior of weight monitoring. The power analysis was performed using G*Power (Heinrich Heine University) [57]. With a sample of 38, our study was underpowered. However, this sample size was sufficient to conduct a feasibility study [58], which can inform implementation of a fully powered study with fewer problems to test the SCDG's effectiveness.

Data Analysis

Descriptive statistics for feasibility included (1) the percentage of participants recruited from the total who were approached, (2) the percentage of participants (among those recruited) who were retained in the study at the end of 24 weeks in each arm and overall, (3) the number of days the IG participants played the SCDG, and (4) average satisfaction with the SCDG. All open-ended questions were coded by 2 team members (CA and AS) and then analyzed using a general inductive thematic approach [59].

The observed effect sizes (Cohen *d*) for the primary outcome of weight monitoring days at the end of 12 weeks was calculated using the following formula: mean of days in IG minus mean of days in CG divided by average SD of days in IG and days in CG. The potential clinical meaningfulness of the results (in addition to statistical significance) was based on the magnitude of the effects: small (Cohen d=0.20), medium (Cohen d=0.50), and large (Cohen d=0.80) [60].

All statistical analyses were conducted using SPSS software (version 26.0; IBM Corp). Baseline characteristics of the participants in the IG and CG were compared using independent 2-tailed t tests for continuous variables and chi-square tests for categorical variables. Although this study was a feasibility trial, we assessed within-group trends in the IG and CG using paired-sample t tests (2-tailed) at baseline, 6 weeks, 12 weeks, and 24 weeks. Within-group changes were presented as absolute changes from baseline. All data were presented as means with

XSL•FO
RenderX

SDs. Missing values were addressed using intent-to-treat principles with maximum likelihood estimations.

Ethical Considerations

The institutional review board at the University of Texas at Austin approved this study on July 30, 2018 (number 2017-12-0042). Interested participants provided their phone numbers to be contacted by a nursing graduate research assistant for formal screening through phone. If potential participants were found eligible, the research assistant scheduled visits at their homes to complete the written informed consent.

During the COVID-19 pandemic, the eligible potential participants completed the verbal informed consent with the research assistant on the phone. The participants were also emailed a copy of the informed consent form. To protect the participants' privacy, 6-digit unique ID numbers and dummy email addresses were generated for the profile information that was required by the Health Mate app [43] and Google Cloud console so that no personal patient information was shared or processed through the app or the console. Participants who scored higher than 4 on the Patient Health Questionnaire (2 items) for depression [49] were provided resources for local mental health services, including available mental health services at their HF clinical site. The eligible participants received US $75 as incentives spread over 4 time points for participating in the study. The participants who continued in the study during the COVID-19 pandemic were also allowed to keep the weighing scale and tracker to avoid reuse of the devices.

Results

Baseline Characteristics

The characteristics of the participants are presented in Table 1. There were no marked differences between the IG and the CG except for marital status, with significantly more IG participants being married or having a partner.

XSL•FO
RenderX

Table 1. Baseline characteristics of the cohort (N=38).

Baseline characteristics	SCDG[a] (IG[b]; n=19), n (%)	Sensors only (CG[c]; n=19), n (%)	Total, n (%)	*P* value
Age (years)				.20
55-64	11 (58)	9 (47)	20 (53)	
65-74	7 (37)	5 (26)	12 (32)	
≥75	1 (5)	5 (26)	6 (15)	
Sex				.99
Male	10 (53)	10 (53)	20 (53)	
Female	9 (47)	9 (47)	18 (47)	
Ethnicity				.07
Hispanic or Latino	3 (16)	0 (0)	3 (8)	
Non-Hispanic or Latino	16 (84)	19 (100)	35 (92)	
Race				.25
White	16 (84)	13 (69)	29 (76)	
African American	2 (11)	4 (21)	6 (16)	
Native American	0 (0)	1 (5)	1 (3)	
Other	1 (5)	1 (5)	2 (5)	
Highest level of education				.90
High school	4 (21)	3(16)	7 (18)	
University education	8 (42)	8 (42)	16 (42)	
Technical diploma	7 (37)	8 (42)	15 (40)	
Marital status				.05
Married or has a partner	13 (68)	7 (37)	20 (53)	
Divorced or widowed	6 (32)	12 (63)	18 (47)	
Living arrangement				.35
Living alone	5 (26)	6 (32)	11 (29)	
Living with others	12 (63)	12 (68)	24 (63)	
Other	2 (11)	0 (0)	2 (5)	
Duration with HF[d] diagnosis				.63
<6 months	6 (32)	5 (26)	11 (29)	
7 months to 1 year	2 (11)	3 (16)	5 (13)	
1-5 years	7 (37)	5 (26)	12 (32)	
>5 years	4 (21)	6 (32)	10 (26)	
Last HF hospitalization				.51
<1 month	5 (26)	6 (32)	11 (29)	
1-6 months	7 (37)	3 (16)	10 (26)	
7 months to 1 year	1 (5)	2 (11)	3 (8)	
>1 year	6 (32)	8 (42)	14 (37)	
Prior digital game-playing				.23
Yes	17 (90)	13 (68)	30 (79)	
No	2 (10)	6 (32)	8 (21)	
Depression (>2 on PHQ[e] -2)				.99
No	14 (74)	14 (74)	28 (74)	

XSL•FO
RenderX

Baseline characteristics	SCDG[a] (IG[b]; n=19), n (%)	Sensors only (CG[c]; n=19), n (%)	Total, n (%)	*P* value
Yes	5 (26)	5 (26)	10 (26)	
Number of comorbid conditions				.62
0-2	12 (63)	13 (68)	25 (66)	
>2	8 (42)	6 (32)	14 (37)	
Top 2 comorbid conditions				
Hypertension	10 (53)	10 (53)	20 (53)	.99
Diabetes	8 (42)	5 (26)	13 (34)	.55

[a]SCDG: sensor-controlled digital game.

[b]IG: intervention group.

[c]CG: control group.

[d]HF: heart failure.

[e]PHQ: Patient Health Questionnaire.

Feasibility Outcomes

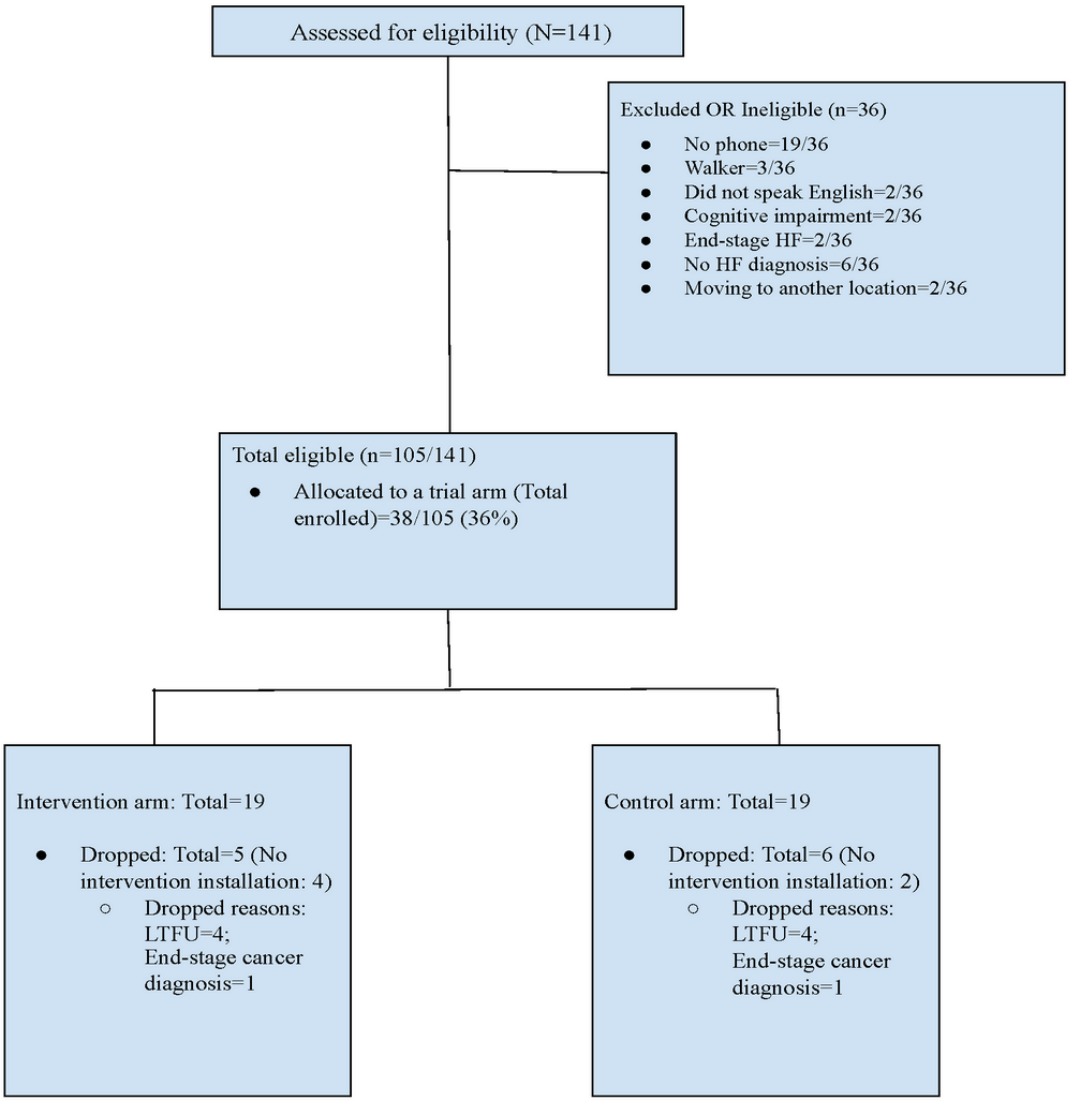
Recruitment and Enrollment

Before the COVID-19 outbreak, 91 older adults with HF were approached at clinical centers (cardiac rehabilitation center, cardiac hospital unit, and senior centers), of whom 25 (27%) were found ineligible for the study. The reasons for exclusion were as follows: not owning smartphones (19/25, 76%), moderate cognitive impairment (2/25, 8%), moving to another state (pre–COVID-19; 2/25, 8%), or not fluent in English (2/25, 8%). Of the 66 eligible participants, 14 (21%) enrolled in the study.

During the COVID-19 pandemic, 50 patients with HF were approached, of whom 11 (22%) were ineligible. The reasons for exclusion were as follows: moving to a hospice (2/11, 18%), not having a confirmed diagnosis of HF (6/11, 55%), or using a wheelchair or walker to ambulate (3/11, 27%). Of the 39 participants found eligible—11 (28%) from clinical sites and 28 (72%) through web-based recruitment—24 (61%) agreed to enroll in the study: 6 (25%) from clinical sites and 18 (75%) through web-based recruitment (Figure 4). The average distance of the participants' homes from the study team increased from 14 miles (range 2-52) during the pre–COVID-19 phase to 143 miles (range 11-465) during the COVID-19 phase.

XSL•FO
RenderX

Figure 4. Diagram of participant flow. HF: heart failure; LTFU: lost to follow-up.



Retention

Of the 38 participants who enrolled in the study, 6 (16%) completed the baseline survey but dropped out before installation of the devices or apps. Of the 6 participants who dropped out, 4 (67%) were lost to follow-up, and 2 dropped out because of end-stage disease unrelated to HF. We obtained the 12-week objective behavior data assessments from 84% (32/38) and 79% (30/38) of the total enrolled participants for weight monitoring and physical activity, respectively. Although 100% (32/32) of the participants who installed the apps were able to transmit data on daily weight monitoring, 6% (2/32; 1 in the CG and 1 in the IG) experienced issues with synchronizing the tracker and were unable to transmit data on physical activity. As the tracker became an outdated version during the course of the study, the device company was unable to support troubleshooting for these 2 trackers. Table 2 provides retention statistics on the survey data available to the study team at each data collection time point.

XSL•FO
RenderX

Table 2. Retention of participants at each time point (N=38).

Data collection	Participants, n (%)
12-week weight monitoring	32 (84)
12-week physical activity	30 (79)
6-week surveys	30 (79)
12-week surveys	30 (79)
24-week surveys	27 (71)

Engagement

Installation and training for the devices and apps was completed with 100% (32/32) of the participants who progressed in the study. Of the 32 participants, 14 (44%) owned a smartphone with the Android platform, whereas the remaining 18 (56%) owned a smartphone with the iOS platform. Of the 15 remaining IG participants, 11 (71%) played the SCDG more than 50% of the days (range 7%-96%). Using the daily data of the 15 IG participants, the number of days each player opened the game app was strongly associated with the number of days each player engaged in weighing (r=0.72; P=.04) and the number of days with physical activity step data (r=0.9; P<.001).

Satisfaction With the SCDG

At the end of 12 weeks of game-playing, of the 15 IG participants, 13 (87%) completed the survey regarding their satisfaction with the SCDG (Table 3).

Table 3. Intervention group participants' satisfaction with the sensor-controlled digital game (N=13).

Satisfaction parameters	Participants, n (%)
Interesting	11 (85)
Easy	13 (100)
Enjoyable	11 (85)
Satisfying to play	11 (85)
Satisfaction with the avatar's look	11 (85)
Satisfaction with sound in the game	12 (92)
Satisfaction with graphics	12 (92)
Satisfaction with using the sensor devices to progress in the game	8 (62)
Satisfaction with the content or information	13 (100)
Game motivated me to weigh myself daily	12 (92)
Game motivated me to exercise more	11 (85)
Will recommend this game to others with HF[a]	11 (85)
Prefer playing digital games over other ways to learn about managing HF	10 (77)

[a]HF: heart failure.

The themes related to playing the SCDG included its competitive nature (2/13, 15%), motivation to attain health behavior goals (5/13, 38%), ease of the interface between the sensor devices and the game (2/13, 15%), opportunity to learn about managing HF (2/13, 15%), and access to behavior data (1/13, 8%). A participant said that the duration of the game play for 12 weeks was not long enough to develop healthy habits:

> *I was disappointed that it ended so soon. It was motivational for me to continue to pay attention to my heart failure regimen until it became more of a routine. I'm at a stage in my life where I can continue to strengthen my heart by adding more to my routine.* [Participant]

We faced challenges in syncing data from some of the trackers within the SCDG in real time because the device company discontinued support by the time the study was implemented. Therefore, a common theme among the barriers to playing the game included problems with syncing steps from the Withings Go activity tracker with the game (5/13, 38%). Other barriers included inaccuracies in tracking bicycle activity as steps and the lowest attainable step goal in the game being too high for a participant with HF:

> *Steps and physical activity at unrealistic start levels. I started out at 280 steps, got up to 550 steps a day.* [Participant]

Of the 13 participants who completed the survey, 2 (15%) found the game to be simplistic and said they would prefer more features. A participant stated as follows:

> *Needs more bells and whistles. More color and animations.* [Participant]

XSL•FO
RenderX

Finally, of the 13 participants who completed the survey, 3 (23%) stated that the use of digital games to motivate HF behaviors was not appealing. Of these 3 participants, 2 (67%) preferred reading to playing games, whereas 1 (33%) stated as follows:

> *If it would help people do daily weight/vital sign checks it would be worthwhile. I didn't need motivation other than my cardiologist telling me to.* [Participant]

Health and Behavior Outcomes

Weight Monitoring

At the end of 12 weeks, in comparison with baseline self-reports of weighing behaviors, trends of increase by 40% in weighing 5 days or more in a week were observed in the IG. Trends of decrease by 6% were observed in the CG for weighing 5 days or more in a week (Table 4). An effect size of 0.53 for the SCDG intervention was obtained for the primary outcome of mean days with weight monitoring.

Table 4. Behavior outcomes at baseline and at 6 and 12 weeks.

Outcome and group	Baseline (self-report)	1-6 weeks (sensor data)	7-12 weeks (sensor data)	1-12 weeks (sensor data)
Weighing 5 days or more in a week, n (%)				
IG[a] (n=15)	6 (40)	13 (87)	11 (73)	12 (80)
CG[b] (n=17)	10 (53)	12 (71)	8 (47)	8 (47)
All (N=32)	16 (50)	25 (78)	19 (5)	20 (63)
Days with weighing (maximum: 42 in 6 weeks; 84 in 12 weeks), mean (SD); median				
IG (n=15)	N/A[c]	35.8 (6.0); 37	32.9 (10.5); 35	69.7 (16.1); 76
CG (n=17)	N/A	32.2 (10.2); 34	28.0 (10.9); 30.4	60.1 (18.7); 64
All (N=32)	N/A	34 (8.4); 37	30 (10.9); 31	64 (17.8); 64
Physical activity steps, mean (SD); median				
IG (n=15)	N/A	2742 (2499); 1986	3365 (2821); 2281	2887 (2821); 2133
CG (n=16)	N/A	2638 (1573); 2522	2444 (1757); 2244	2541 (1604); 2601
All (N=31)	N/A	2690 (2060); 2480	2737 (2596); 2262	2713 (2270); 2482

[a]IG: intervention group.

[b]CG: control group.

[c]N/A: not applicable.

Overall, across the 2 groups, the female participants had lower trends in weight monitoring than the male participants (60 vs 70 days, respectively, out of 84); however, both biological sex groups had higher trends in the IG (men, 80; women, 61) compared with the CG (men, 62; women, 57) for days with weight monitoring.

Physical Activity

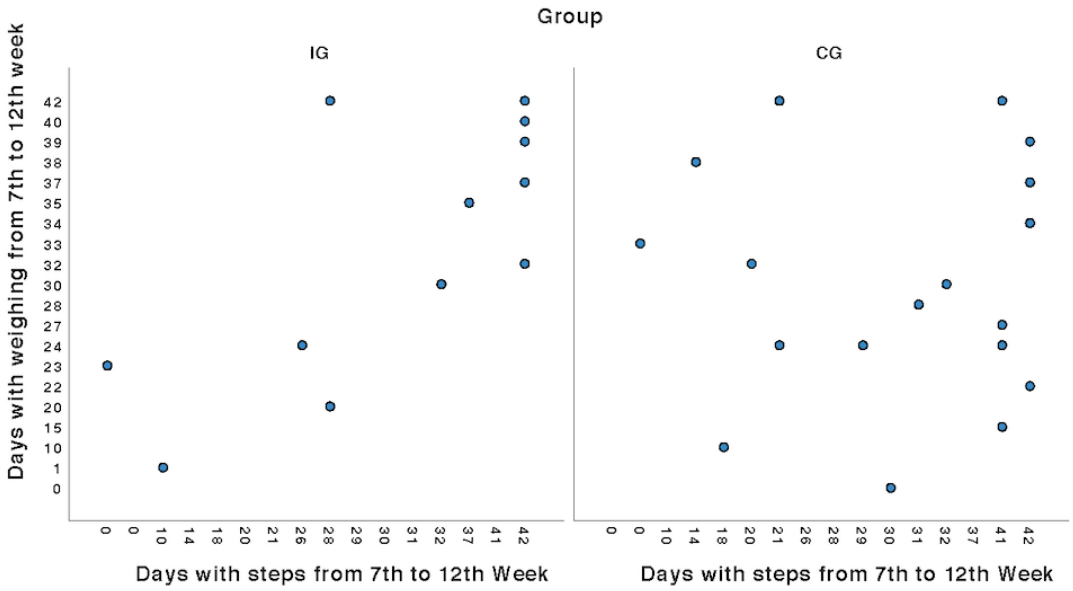
Trends in average physical activity steps from the 6th *week* to the 12th week were seen with modest increases in the IG and modest decreases in the CG (Table 4).

Multiple Behavior Engagement

HF self-management requires simultaneous engagement in multiple behaviors over an extended duration. The IG participants showed higher correlation (r=0.78; P<.001) between engaging in weight monitoring behaviors and engagement in physical activity behaviors from the 6th week to the 12th week than the CG participants (r=−0.01; P=.79; Figure 5).

XSL•FO
RenderX

Figure 5. Engagement in dual behaviors of weight monitoring and physical activity. CG: control group; IG: intervention group.



HF Functional Status

At the end of 6 weeks, the IG demonstrated a clinically significant increase of 7 points on the KCCQ [50,61]. At the end of 24 weeks, both groups retained a clinically and statistically significant increase in functional status over baseline levels (Table 5).

Quality of Life

Both groups demonstrated a clinically [62] and statistically significant within-group increase in quality of life at 6, 12, and 24 weeks compared with baseline on the KCCQ (Table 5).

HF Self-management Knowledge

Both groups demonstrated modest but statistically significant within-group improvement in knowledge at 6, 12, and 24 weeks compared with baseline (Table 5).

Self-reported HF Self-management Behaviors

The IG participants consistently demonstrated improvement in self-reported self-management behaviors compared with baseline and retained a statistically significant improvement at the end of 24 weeks. Self-reported behaviors in the CG participants changed minimally in comparison with baseline (Table 5).

HF Self-efficacy

Both groups demonstrated within-group improvement in self-efficacy across the different time points, although the CG participants demonstrated higher trends of improvement (Table 5).

Motivation for HF Self-management Behaviors

The changes in autonomous regulation of behaviors were nonsignificant in both the groups at 6, 12, and 24 weeks (Table 5).

HF Hospitalization

Decreases in 1-month and 6-month hospitalization rates were seen in both the IG and CG, albeit tempered by the worldwide COVID-19 effect of decrease in HF hospitalizations [63,64] (Table 5).

XSL•FO
RenderX

Table 5. Survey outcomes at baseline and at 6, 12, and 24 weeks (N=38).

Paired outcome and group	Baseline	From baseline to 6 weeks		From baseline to 12 weeks		From baseline to 24 weeks	
	Value	Value	*P* value	Value	*P* value	Value	*P* value
HF[a]-related functional status							
Population, n	N/A[b]	30	N/A	30	N/A	26	N/A
IG[c], mean (SD)	66.6 (28.4)	6.9 (14.1)	.048[c]	3.7 (14.1)	.27	7.4 (10.7)	.07
CG[d], mean (SD)	69.2 (20.8) 72.9	1.7 (11.8)	.54	4.1 (12.6)	.18	6.5 (11.7)	.03[e]
Quality of life							
Population, n	N/A	30	N/A	30	N/A	26	N/A
IG, mean (SD)	62.03 (24.1)	5.8 (13.1)	.07	8.2 (13.5)	.02[e]	10.8 (16.9)	.01[e]
CG, mean (SD)	60.2 (20.7)	8.7 (15.3)	.02[e]	13.0 (19.3)	.01[e]	15.4 (19.6)	.003[e]
HF self-management knowledge							
Population, n	N/A	30	N/A	30	N/A	26	N/A
IG, mean (SD)	24.5 (2.6)	1.9 (2.5)	.003[e]	1.4 (1.7)	.002[e]	1.1 (2.1)	.03[e]
CG, mean (SD)	25.0 (2.4)	2.1 (2.5)	.001[e]	1.1 (1.5)	.005[e]	1.1 (1.7)	.01[e]
Self-reported HF self-management behaviors							
Population, n	N/A	30	N/A	31	N/A	25	N/A
IG, mean (SD)	74.1 (21.1)	4.9 (20.4)	.31	9.3 (16.5)	.02[e]	10.2 (13.4)	.004[e]
CG, mean (SD)	81.7 (15.5)	−2.7 (9.9)	.26	2.6 (12.4)	.4	0.5 (9.8)	.8
HF self-efficacy							
Population, n	N/A	30	N/A	31	N/A	25	N/A
IG, mean (SD)	69.9 (20.7)	2.2 (25.7)	.7	1.0 (20.5)	.8	3.0 (18.0)	.5
CG, mean (SD)	62.6 (18.4)	6.9 (18.1)	.11	10.2 (12.3)	.002[b]	7.0 (16.5)	.08
Motivation for HF self-management behaviors							
Population, n	N/A	30	N/A	30	N/A	25	N/A
IG, mean (SD)	1.63 (2.2)	−0.1 (1.2)	.75	0.1 (2.6)	.9	−0.6 (2.5)	.4
CG, mean (SD)	1.64 (1.6)	−0.2 (2.2)	.8	–0.3 (3.0)	.7	−0.2 (2.7)	.8
HF hospitalization							
In the last month							
Population, n	N/A	31	N/A	31	N/A	26	N/A
IG, mean (SD)	5 (26.3)	1 (7)	N/A	0 (0)	N/A	1 (7)	N/A
CG, mean (SD)	6 (31.6)	1 (6)	N/A	1 (6)	N/A	0 (0)	N/A
In the past 6 months							
Population, n	N/A	N/A	N/A	N/A	N/A	N/A	N/A
IG, mean (SD)	12 (63)	N/A	N/A	N/A	N/A	2 (15)	N/A
CG, mean (SD)	9 (47)	N/A	N/A	N/A	N/A	2 (15)	N/A

[a]HF: heart failure.
[b]N/A: not applicable.
[c]IG: intervention group.
[d]CG: control group.
[e]Significant at $P<.05$.

XSL•FO
RenderX

Discussion

Principal Findings

We have demonstrated the feasibility of conducting a randomized controlled trial to compare an SCDG intervention and a sensors-only intervention for improving older adults' HF self-management behaviors of daily weight monitoring and physical activity and other health outcomes. Despite the constraints imposed by the COVID-19 pandemic, we successfully attained benchmark parameters for the feasibility outcomes of recruitment, retention, intervention engagement, and satisfaction with the intervention among older adults diagnosed with HF. The advantages of recruiting participants and conducting the trial remotely enabled us to increase access to our study from 2 counties to more than 18 counties, which included rural areas. The older adult participants in our study were able to use the digital game app and sensors regardless of their education level, familiarity with digital games, or smartphone platform (ie, iOS or Android).

Regarding the behavior outcomes, the IG demonstrated better trends in the primary outcome of weight monitoring behavior, with a higher average of weight monitoring days than the CG: 46% more IG participants than CG participants attained the clinically significant level of weighing 5 days or more in a week. Similarly, the IG participants demonstrated higher engagement in the dual behaviors of weight monitoring and physical activity than the CG participants. Moreover, the IG participants demonstrated significant within-group improvement in self-reported HF self-management behaviors (weight monitoring, exercise, salt restriction, medication adherence, crisis recognition, and related follow-up), which persisted at 24 weeks. This result is promising because improving overall HF outcomes depends on engagement in multiple HF self-management behaviors over long time intervals.

Ours is one of the few studies to capture daily physical activity steps of older adults with HF over 12 weeks. Although both IG and CG participants demonstrated an increase in average daily physical steps from the 6th week to the 12th week, the increase was still lower overall than recommended for physical activity steps [65]. Given the technical issues with an outdated tracker, the SCDG may motivate higher increases in step levels with better trackers. Although typical recommended levels for cardiovascular health benefits have ranged from 10,000 [65] to 15,000 [66] steps a day, a study with 16,000 older women found that even an average of 4400 steps a day resulted in significantly lower mortality rates [67]. However, older adults with HF often suffer from frailty and fatigue; therefore, it may be harder for them to reach the optimal step levels. In a recent study that examined the walking activity of adults with HF using a step counter over a year, younger age, higher ejection fraction, and lower HF classification were found to be significantly correlated with the number of daily steps [68]. Whether modest increases in steps result in health benefits among older adults with HF remains to be seen and may signify a need for additional supportive or palliative interventions that relieve physical or mental symptom burden to ensure optimal quality of life among patients with HF [69]. Nevertheless, studies like ours demonstrate actual physical activity engagement in real-world situations and examine the longitudinal relationship between clinically significant physical activity steps and optimal HF outcomes. A recent position paper on measuring physical activity by varied activity monitors for HF in real-world conditions has provided key criteria for the selection of activity monitors based on the aims of the research and observation metrics [70]. On the basis of our participants' experience with the SCDG, future iterations could include the ability to set realistic step goal levels for participants with HF that are lower than 3000 steps and game play duration that is longer than 12 weeks.

The within-group trends of improvement in functional status and quality of life in both the IG and CG that persisted until the end of the study demonstrate the potential of digital health interventions to improve outcomes among older adults with HF. Similarly, trends of improvement in HF self-efficacy were observed in both the IG and CG. Both groups were allowed to retain the sensor devices after 12 weeks of behavior data collection, which might have contributed to persistent improvement in perceived functional status, quality of life, and self-efficacy in comparison with baseline levels. The lack of perceptible changes in autonomous regulation of behaviors in either group could be explained by the participants' perceiving the SCDG and the sensors-only interventions as expecting accountability of their behaviors, which may be analogous to controlled regulation of behavior. Future research can help explore the need and preferences for either kind of behavior regulation for sustaining HF self-management behaviors.

Improved HF management at the clinical level may have resulted in the baseline knowledge of HF self-management being already high in both groups. Still, both groups demonstrated trends of modest improvements in knowledge at 6 and 12 weeks that were maintained at 24 weeks.

Reduction in HF hospitalization is the desired distal outcome of improved self-management behaviors. Although both groups demonstrated reduction in hospitalization, it is unclear whether the reduction resulted from the interventions or from the worldwide reduction in cardiac-related hospitalizations due to the COVID-19 pandemic [62,64]. The reduction in hospitalizations by more than 30% in both the groups persisted until the end of the 24-week study period.

Comparison With Prior Work

Recent studies have explored the feasibility of presenting data from fitness tracker sensors or Bluetooth-enabled weighing scales to track behaviors and provide feedback to patients with HF through an mHealth app. In a single-group study with 20 individuals with HF who were followed over 6 months to track weight monitoring, physical activity, and medication adherence, 60% of the participants wore trackers and 45% used the weighing scale for more than 70% of the days [71]. In our study, among the 32 participants who installed the SCDG app, 27 (80%) wore the trackers for more than 70% of the days and used the weighing scale for more than 80% of the days at the end of 12 weeks. In addition, the IG participants in our study experienced an increase in HF functional status and similar increases in knowledge in comparison with the IG participants

XSL•FO
RenderX

in another recent study with an mHealth app intervention for HF that also employed a Bluetooth weighing scale [72]. Although our study's results with the SCDG app compare favorably with those of similar studies that have examined mHealth apps and sensor devices with patients with HF, given the small sample size in our feasibility trial, the results and findings from our trends analysis should be treated with caution.

Limitations

Digital literacy is often a concern when digital health interventions are implemented with older adults [73]. Our study team's prior experience in conducting usability assessments with older adults with HF as well as in installation of, and training for, devices and apps at the participants' homes [25] helped us to design the information and content to support the participants' self-installation of the apps and use of the devices. Nevertheless, the study might have suffered from selection bias because participation was voluntary and participants already comfortable with using technology might have been especially motivated to participate. In addition, despite block randomization and stratification by biological sex, a higher proportion of participants in the IG were married or had a partner. Having a partner to support self-management efforts might have resulted in better trends of behavior outcomes in the IG.

Another limitation was the research team's inability to intervene immediately in the event of missing behavioral data or unsafe weight changes throughout the 12-week game play duration. Our approach of using the Twilio app (Twilio Inc) [74] for alerting the nursing research assistant to the absence of physical activity or weight monitoring data or to sudden weight changes became unmanageable because of the high volume of text messages. In future studies, developing a dashboard that will allow easier visualization of behavioral data from all participants will provide us with the opportunity to intervene immediately in the event of absent or unsafe levels of behavioral data.

Conclusions

This pilot randomized controlled trial indicates that playing an SCDG app on smartphones is feasible and acceptable for older adults with HF for motivating daily weight monitoring and physical activity. The participants who played the SCDG demonstrated positive trends in daily weight monitoring and physical activity behaviors and reduction in HF hospitalization compared with the sensors-only group. This study justifies fully powered efficacy trials with digital gaming solutions to motivate adherence to HF self-management behaviors and improve the health outcomes of older adults with HF.

Acknowledgments

This research was supported by the National Institute of Nursing Research of the National Institutes of Health under award number R21NR018229. The content is the sole responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Editorial support with manuscript development was provided by Dr John Bellquist at the Cain Center for Nursing Research at the University of Texas at Austin School of Nursing. The authors would like to thank Cathy B Spranger, DrPH, manager at Seton Medical Center Cardiac Rehab Laboratory; Dr Clay Cauthen, cardiologist at Seton Medical Center; nurse educator Rebecca Cox; and the Continuum of Care clinicians at Seton Medical Center for helping us recruit participants to this study. The authors would also like to thank Hira Asim and Katelyn Leggio, student research volunteers and assistants at the University of Texas at Austin.

Authors' Contributions

KR conceptualized the study together with TB, CJ, MO, MK, and ET. KR, CA, and AS recruited the participants. CA collected survey outcomes data, which was supervised by KR. AS helped with technical installation and support, which was supervised by KR. MO and GL developed the technology, which was supervised by CJ, TB, and KR. Data were curated by GL, which was supervised by KR. BV conducted the data analysis. KR wrote the manuscript, which was critically reviewed by TB, CJ, BV, and MK. All authors approved the final draft of the submitted manuscript.

Conflicts of Interest

MO is the owner of Good Life Games, Inc, a company that develops health games for hire, which provided the prototype game images for this study.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2146 KB - games_v9i4e29044_app1.pdf]

References

1. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart disease and stroke statistics-2020 update: a report from the American Heart Association. Circulation 2020 Mar 03;141(9):e139-e596. [doi: 10.1161/CIR.0000000000000757] [Medline: 31992061]

XSL•FO
RenderX

2. Murad K, Goff DC, Morgan TM, Burke GL, Bartz TM, Kizer JR, et al. Burden of comorbidities and functional and cognitive impairments in elderly patients at the initial diagnosis of heart failure and their impact on total mortality: the cardiovascular health study. JACC Heart Fail 2015 Jul;3(7):542-550 [FREE Full text] [doi: 10.1016/j.jchf.2015.03.004] [Medline: 26160370]
3. Lainscak M, Blue L, Clark AL, Dahlström U, Dickstein K, Ekman I, et al. Self-care management of heart failure: practical recommendations from the Patient Care Committee of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2011 Feb;13(2):115-126 [FREE Full text] [doi: 10.1093/eurjhf/hfq219] [Medline: 21148593]
4. Heart Failure Society of America, Lindenfeld J, Albert NM, Boehmer JP, Collins SP, Ezekowitz JA, et al. HFSA 2010 comprehensive heart failure practice guideline. J Card Fail 2010 Jun;16(6):e1-194. [doi: 10.1016/j.cardfail.2010.04.004] [Medline: 20610207]
5. Jones CD, Holmes GM, Dewalt DA, Erman B, Broucksou K, Hawk V, et al. Is adherence to weight monitoring or weight-based diuretic self-adjustment associated with fewer heart failure-related emergency department visits or hospitalizations? J Card Fail 2012 Jul;18(7):576-584 [FREE Full text] [doi: 10.1016/j.cardfail.2012.05.004] [Medline: 22748492]
6. Jones CD, Holmes GM, DeWalt DA, Erman B, Wu J, Cene CW, et al. Self-reported recall and daily diary-recorded measures of weight monitoring adherence: associations with heart failure-related hospitalization. BMC Cardiovasc Disord 2014 Jan 31;14:12 [FREE Full text] [doi: 10.1186/1471-2261-14-12] [Medline: 24483776]
7. Ismail H, McFarlane JR, Nojoumian AH, Dieberg G, Smart NA. Clinical outcomes and cardiovascular responses to different exercise training intensities in patients with heart failure: a systematic review and meta-analysis. JACC Heart Fail 2013 Dec;1(6):514-522 [FREE Full text] [doi: 10.1016/j.jchf.2013.08.006] [Medline: 24622004]
8. Tu R, Zeng Z, Zhong G, Wu W, Lu Y, Bo Z, et al. Effects of exercise training on depression in patients with heart failure: a systematic review and meta-analysis of randomized controlled trials. Eur J Heart Fail 2014 Jul;16(7):749-757 [FREE Full text] [doi: 10.1002/ejhf.101] [Medline: 24797230]
9. Piepoli MF, Conraads V, Corrà U, Dickstein K, Francis DP, Jaarsma T, et al. Exercise training in heart failure: from theory to practice. A consensus document of the Heart Failure Association and the European Association for Cardiovascular Prevention and Rehabilitation. Eur J Heart Fail 2011 Apr;13(4):347-357 [FREE Full text] [doi: 10.1093/eurjhf/hfr017] [Medline: 21436360]
10. Jaarsma T, Klompstra L, Ben Gal T, Boyne J, Vellone E, Bäck M, et al. Increasing exercise capacity and quality of life of patients with heart failure through Wii gaming: the rationale, design and methodology of the HF-Wii study; a multicentre randomized controlled trial. Eur J Heart Fail 2015 Jul;17(7):743-748 [FREE Full text] [doi: 10.1002/ejhf.305] [Medline: 26139585]
11. Nieuwenhuis MM, Jaarsma T, van Veldhuisen DJ, Postmus D, van der Wal MH. Long-term compliance with nonpharmacologic treatment of patients with heart failure. Am J Cardiol 2012 Aug 01;110(3):392-397. [doi: 10.1016/j.amjcard.2012.03.039] [Medline: 22516525]
12. van der Wal MH, van Veldhuisen DJ, Veeger NJ, Rutten FH, Jaarsma T. Compliance with non-pharmacological recommendations and outcome in heart failure patients. Eur Heart J 2010 Jun;31(12):1486-1493. [doi: 10.1093/eurheartj/ehq091] [Medline: 20436049]
13. Bos-Touwen I, Jonkman N, Westland H, Schuurmans M, Rutten F, de Wit N, et al. Tailoring of self-management interventions in patients with heart failure. Curr Heart Fail Rep 2015 Jun;12(3):223-235 [FREE Full text] [doi: 10.1007/s11897-015-0259-3] [Medline: 25929690]
14. Herrmann LK, Kim J. The fitness of apps: a theory-based examination of mobile fitness app usage over 5 months. Mhealth 2017 Jan 30;3:2 [FREE Full text] [doi: 10.21037/mhealth.2017.01.03] [Medline: 28293619]
15. Attig C, Franke T. Abandonment of personal quantification: a review and empirical study investigating reasons for wearable activity tracking attrition. Comput Human Behav 2020 Jan;102:223-237. [doi: 10.1016/j.chb.2019.08.025]
16. Vieira A, Gabriel J, Melo C, Machado J. Kinect system in home-based cardiovascular rehabilitation. Proc Inst Mech Eng H 2017 Jan;231(1):40-47. [doi: 10.1177/0954411916679201] [Medline: 27913628]
17. Thompson D, Baranowski T, Buday R, Baranowski J, Thompson V, Jago R, et al. Serious video games for health how behavioral science guided the development of a serious video game. Simul Gaming 2010 Aug 01;41(4):587-606 [FREE Full text] [doi: 10.1177/1046878108328087] [Medline: 20711522]
18. Baranowski MT, Lu AS, Buday R, Lyons EJ, Schell J, Russoniello C. Stories in games for health: more pros or cons? Games Health J 2013 Oct;2(5):256-263 [FREE Full text] [doi: 10.1089/g4h.2013.0052] [Medline: 26196925]
19. Marston H. Older adults as 21st century game designers. Comput Game J 2012 May 15;1(1):90-102 [FREE Full text] [doi: 10.1007/bf03392330]
20. Beale IL, Kato PM, Marin-Bowling VM, Guthrie N, Cole SW. Improvement in cancer-related knowledge following use of a psychoeducational video game for adolescents and young adults with cancer. J Adolesc Health 2007 Sep;41(3):263-270. [doi: 10.1016/j.jadohealth.2007.04.006] [Medline: 17707296]
21. Serious Games: Mechanisms and Effects. New York: Routledge; 2009.
22. McGonigal J. Reality is Broken: Why Games Make Us Better and How They Can Change the World. Delhi, India: RHUK; 2012.

XSL•FO
RenderX

23. Gardner B. A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. Health Psychol Rev 2015;9(3):277-295 [FREE Full text] [doi: 10.1080/17437199.2013.876238] [Medline: 25207647]
24. Brox E, Fernandez-Luque L, Tøllefsen T. Healthy gaming - video game design to promote health. Appl Clin Inform 2011 Apr 27;2(2):128-142 [FREE Full text] [doi: 10.4338/ACI-2010-10-R-0060] [Medline: 23616865]
25. Radhakrishnan K, Julien C, O'Hair M, Baranowski T, Lee G, Allen C, et al. Usability testing of a sensor-controlled digital game to engage older adults with heart failure in physical activity and weight monitoring. Appl Clin Inform 2020 Oct;11(5):873-881. [doi: 10.1055/s-0040-1721399] [Medline: 33378780]
26. 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]
27. Theng Y, Lee JW, Patinadan PV, Foo SS. The use of videogames, gamification, and virtual environments in the self-management of diabetes: a systematic review of evidence. Games Health J 2015 Oct;4(5):352-361. [doi: 10.1089/g4h.2014.0114] [Medline: 26287926]
28. Gaming attitudes and habits of adults ages 50-plus. AARP Research. 2019. URL: https://www.aarp.org/content/dam/aarp/research/surveys_statistics/technology/2019/2020-gaming-trends-older-adults.doi.10.26419-2Fres.00328.001.pdf [accessed 2021-10-02]
29. Dithmer M, Rasmussen JO, Grönvall E, Spindler H, Hansen J, Nielsen G, et al. "The Heart Game": using gamification as part of a telerehabilitation program for heart patients. Games Health J 2016 Feb;5(1):27-33 [FREE Full text] [doi: 10.1089/g4h.2015.0001] [Medline: 26579590]
30. Radhakrishnan K, Toprac P, O'Hair M, Bias R, Kim MT, Bradley P, et al. Interactive digital e-health game for heart failure self-management: a feasibility study. Games Health J 2016 Dec;5(6):366-374 [FREE Full text] [doi: 10.1089/g4h.2016.0038] [Medline: 27976955]
31. Flores M, Glusman G, Brogaard K, Price ND, Hood L. P4 medicine: how systems medicine will transform the healthcare sector and society. Per Med 2013;10(6):565-576 [FREE Full text] [doi: 10.2217/pme.13.57] [Medline: 25342952]
32. Dolgin M. Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels. 9th edition. Philadelphia: Lippincott Williams and Wilkins; 1994.
33. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini-cog: a cognitive 'vital signs' measure for dementia screening in multi-lingual elderly. Int J Geriatr Psychiatry 2000 Nov;15(11):1021-1027. [doi: 10.1002/1099-1166(200011)15:11<1021::aid-gps234>3.0.co;2-6] [Medline: 11113982]
34. McAlister FA, Ezekowitz J, Tonelli M, Armstrong PW. Renal insufficiency and heart failure: prognostic and therapeutic implications from a prospective cohort study. Circulation 2004 Mar 02;109(8):1004-1009. [doi: 10.1161/01.CIR.0000116764.53225.A9] [Medline: 14769700]
35. Trialfacts homepage. Trialfacts. URL: https://trialfacts.com [accessed 2020-12-31]
36. Fogg B. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Presented at: Persuasive '09: Proceedings of the 4th International Conference on Persuasive Technology; Apr 26-29, 2009; Claremont California USA. [doi: 10.1145/1541948.1541999]
37. DeWalt DA, Schillinger D, Ruo B, Bibbins-Domingo K, Baker DW, Holmes GM, et al. Multisite randomized trial of a single-session versus multisession literacy-sensitive self-care intervention for patients with heart failure. Circulation 2012 Jun 12;125(23):2854-2862 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.081745] [Medline: 22572916]
38. Heart Failure Society of America homepage. Heart Failure Society of America. URL: https://hfsa.org/heart-failure-educational-modules [accessed 2020-12-23]
39. Thompson D, Baranowski T, Buday R. Conceptual model for the design of a serious video game promoting self-management among youth with type 1 diabetes. J Diabetes Sci Technol 2010 May 01;4(3):744-749 [FREE Full text] [doi: 10.1177/193229681000400331] [Medline: 20513343]
40. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010 Jun 01;152(11):726-732 [FREE Full text] [doi: 10.7326/0003-4819-152-11-201006010-00232] [Medline: 20335313]
41. Withings homepage. Withings. URL: https://support.withings.com/hc/en-us/categories/202313048-Go [accessed 2020-12-31]
42. Body weight and BMI Wi-Fi scale. Withings. URL: https://www.withings.com/us/en/body [accessed 2020-12-31]
43. Withings Health Mate: a fitness, activity and health tracker app. Withings. URL: https://www.withings.com/us/en/health-mate [accessed 2020-12-31]
44. Collier SR, McCraw C, Campany M, Lubkeman A, StClair P, Ji H, et al. Withings body cardio versus gold standards of pulse-wave velocity and body composition. J Pers Med 2020 Mar 11;10(1):17 [FREE Full text] [doi: 10.3390/jpm10010017] [Medline: 32168728]
45. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]

XSL•FO
RenderX

46. Vetrovsky T, Siranec M, Marencakova J, Tufano JJ, Capek V, Bunc V, et al. Validity of six consumer-level activity monitors for measuring steps in patients with chronic heart failure. PLoS One 2019 Sep 13;14(9):e0222569 [FREE Full text] [doi: 10.1371/journal.pone.0222569] [Medline: 31518367]
47. Simple randomization service. Sealed Envelope. URL: https://www.sealedenvelope.com/simple-randomiser/v1 [accessed 2021-03-09]
48. McAuley E, Duncan T, Tammen VV. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis. Res Q Exerc Sport 1989 Mar;60(1):48-58. [doi: 10.1080/02701367.1989.10607413] [Medline: 2489825]
49. Piepenburg SM, Faller H, Gelbrich G, Störk S, Warrings B, Ertl G, et al. Comparative potential of the 2-item versus the 9-item patient health questionnaire to predict death or rehospitalization in heart failure. Circ Heart Fail 2015 May;8(3):464-472. [doi: 10.1161/CIRCHEARTFAILURE.114.001488] [Medline: 25878325]
50. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. J Am Coll Cardiol 2000 Apr;35(5):1245-1255 [FREE Full text] [doi: 10.1016/s0735-1097(00)00531-3] [Medline: 10758967]
51. Reilly CM, Higgins M, Smith A, Gary RA, Robinson J, Clark PC, et al. Development, psychometric testing, and revision of the Atlanta Heart Failure Knowledge Test. J Cardiovasc Nurs 2009;24(6):500-509 [FREE Full text] [doi: 10.1097/JCN.0b013e3181aff0b0] [Medline: 19858959]
52. Lee CS, Lyons KS, Gelow JM, Mudd JO, Hiatt SO, Nguyen T, et al. Validity and reliability of the European Heart Failure Self-care Behavior Scale among adults from the United States with symptomatic heart failure. Eur J Cardiovasc Nurs 2013 Apr;12(2):214-218 [FREE Full text] [doi: 10.1177/1474515112469316] [Medline: 23263273]
53. Vellone E, Riegel B, Cocchieri A, Barbaranelli C, D'Agostino F, Antonetti G, et al. Psychometric testing of the Self-Care of Heart Failure Index Version 6.2. Res Nurs Health 2013 Oct;36(5):500-511. [doi: 10.1002/nur.21554] [Medline: 23832431]
54. Levesque CS, Williams GC, Elliot D, Pickering MA, Bodenhamer B, Finley PJ. Validating the theoretical structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors. Health Educ Res 2007 Oct;22(5):691-702. [doi: 10.1093/her/cyl148] [Medline: 17138613]
55. Williams GC, McGregor HA, Zeldman A, Freedman ZR, Deci EL. Testing a self-determination theory process model for promoting glycemic control through diabetes self-management. Health Psychol 2004 Jan;23(1):58-66. [doi: 10.1037/0278-6133.23.1.58] [Medline: 14756604]
56. Ong MK, Romano PS, Edgington S, Aronow HU, Auerbach AD, Black JT, Better Effectiveness After Transition–Heart Failure (BEAT-HF) Research Group. Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure: the better effectiveness after transition -- heart failure (BEAT-HF) randomized clinical trial. JAMA Intern Med 2016 Mar;176(3):310-318 [FREE Full text] [doi: 10.1001/jamainternmed.2015.7712] [Medline: 26857383]
57. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
58. Sim J, Lewis M. The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. J Clin Epidemiol 2012 Mar;65(3):301-308. [doi: 10.1016/j.jclinepi.2011.07.011] [Medline: 22169081]
59. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
60. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Routledge; 1988.
61. Spertus J, Peterson E, Conard MW, Heidenreich PA, Krumholz HM, Jones P, Cardiovascular Outcomes Research Consortium. Monitoring clinical changes in patients with heart failure: a comparison of methods. Am Heart J 2005 Oct;150(4):707-715. [doi: 10.1016/j.ahj.2004.12.010] [Medline: 16209970]
62. Butler J, Khan MS, Mori C, Filippatos GS, Ponikowski P, Comin-Colet J, et al. Minimal clinically important difference in quality of life scores for patients with heart failure and reduced ejection fraction. Eur J Heart Fail 2020 Jun;22(6):999-1005 [FREE Full text] [doi: 10.1002/ejhf.1810] [Medline: 32239794]
63. Vest AR, Upshaw JN, Dean K, Donovan E, Diakos N, Hong C, et al. The decrease in hospitalizations for heart failure during the COVID-19 pandemic: a community and academic hospital comparison study. J Card Fail 2020 Oct;26(10):S71. [doi: 10.1016/j.cardfail.2020.09.207]
64. Hall ME, Vaduganathan M, Khan MS, Papadimitriou L, Long RC, Hernandez GA, et al. Reductions in heart failure hospitalizations during the COVID-19 pandemic. J Card Fail 2020 Jun;26(6):462-463 [FREE Full text] [doi: 10.1016/j.cardfail.2020.05.005] [Medline: 32405232]
65. Tudor-Locke C, Craig CL, Brown WJ, Clemes SA, De Cocker K, Giles-Corti B, et al. How many steps/day are enough? For adults. Int J Behav Nutr Phys Act 2011 Jul 28;8:79 [FREE Full text] [doi: 10.1186/1479-5868-8-79] [Medline: 21798015]
66. Tigbe WW, Granat MH, Sattar N, Lean ME. Time spent in sedentary posture is associated with waist circumference and cardiovascular risk. Int J Obes (Lond) 2017 May;41(5):689-696. [doi: 10.1038/ijo.2017.30] [Medline: 28138134]
67. Lee I, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of step volume and intensity with all-cause mortality in older women. JAMA Intern Med 2019 Aug 01;179(8):1105-1112 [FREE Full text] [doi: 10.1001/jamainternmed.2019.0899] [Medline: 31141585]

XSL•FO
RenderX

68. Gade JD, Spindler H, Hollingdal M, Refsgaard J, Dittmann L, Frost L, et al. Predictors of walking activity in patients with systolic heart failure equipped with a step counter: randomized controlled trial. JMIR Biomed Eng 2020 Nov 30;5(1):e20776. [doi: 10.2196/20776]
69. Sobanski PZ, Alt-Epping B, Currow DC, Goodlin SJ, Grodzicki T, Hogg K, et al. Palliative care for people living with heart failure: European Association for Palliative Care Task Force expert position statement. Cardiovasc Res 2020 Jan 01;116(1):12-27. [doi: 10.1093/cvr/cvz200] [Medline: 31386104]
70. Klompstra L, Kyriakou M, Lambrinou E, Piepoli MF, Coats AJ, Cohen-Solal A, et al. Measuring physical activity with activity monitors in patients with heart failure: from literature to practice. A position paper from the Committee on Exercise Physiology and Training of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2021 Jan;23(1):83-91 [FREE Full text] [doi: 10.1002/ejhf.2035] [Medline: 33111464]
71. Sohn A, Speier W, Lan E, Aoki K, Fonarow GC, Ong MK, et al. Integrating remote monitoring into heart failure patients' care regimen: a pilot study. PLoS One 2020 Nov 19;15(11):e0242210 [FREE Full text] [doi: 10.1371/journal.pone.0242210] [Medline: 33211733]
72. Wei KS, Ibrahim NE, Kumar AA, Jena S, Chew V, Depa M, et al. Habits heart app for patient engagement in heart failure management: pilot feasibility randomized trial. JMIR Mhealth Uhealth 2021 Jan 20;9(1):e19465 [FREE Full text] [doi: 10.2196/19465] [Medline: 33470941]
73. Smith B, Magnani JW. New technologies, new disparities: the intersection of electronic health and digital health literacy. Int J Cardiol 2019 Oct 01;292:280-282 [FREE Full text] [doi: 10.1016/j.ijcard.2019.05.066] [Medline: 31171391]
74. APIs for reliable SMS delivery. twilio. URL: https://www.twilio.com/sms [accessed 2021-09-24]

Abbreviations

CG: control groupCONSORT: Consolidated Standards of Reporting Trials
HF: heart failure
IG: intervention group
KCCQ: Kansas City Cardiomyopathy Questionnaire
mHealth: mobile health
SCDG: sensor-controlled digital game

Edited by N Zary; submitted 30.03.21; peer-reviewed by B Dinesen, R Krukowski, M Kamada; comments to author 29.05.21; revised version received 03.06.21; accepted 28.06.21; published 08.11.21.

Please cite as:
Radhakrishnan K, Julien C, Baranowski T, O'Hair M, Lee G, Sagna De Main A, Allen C, Viswanathan B, Thomaz E, Kim M
Feasibility of a Sensor-Controlled Digital Game for Heart Failure Self-management: Randomized Controlled Trial
JMIR Serious Games 2021;9(4):e29044
URL: https://games.jmir.org/2021/4/e29044
doi:10.2196/29044
PMID:34747701

©Kavita Radhakrishnan, Christine Julien, Tom Baranowski, Matthew O'Hair, Grace Lee, Atami Sagna De Main, Catherine Allen, Bindu Viswanathan, Edison Thomaz, Miyong Kim. Originally published in JMIR Serious Games (https://games.jmir.org), 08.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development and Validation of a Mobile Game for Culturally Sensitive Child Sexual Abuse Prevention Education in Tanzania: Mixed Methods Study

Maria Proches Malamsha[1], BSc; Elingarami Sauli[2], PhD, MD; Edith Talina Luhanga[3], PhD

[1]School of Computation and Communication Science and Engineering, Nelson Mandela African Institution of Science and Technology, Arusha, United Republic of Tanzania

[2]School of Life Science and Bioengineering, Nelson Mandela African Institution of Science and Technology, Arusha, United Republic of Tanzania

[3]Carnegie Mellon University Africa, Kigali, Rwanda

Corresponding Author:
Maria Proches Malamsha, BSc
School of Computation and Communication Science and Engineering
Nelson Mandela African Institution of Science and Technology
PO Box 447
Nelson Mandela Rd
Arusha
United Republic of Tanzania
Phone: 255 715261702
Email: selestinapro@gmail.com

Abstract

Background: Globally, 3 out of 20 children experience sexual abuse before the age of 18 years. Educating children about sexual abuse and prevention is an evidence-based strategy that is recommended for ending child sexual abuse. Digital games are increasingly being used to influence healthy behaviors in children and could be an efficient and friendly approach to educating children about sexual abuse prevention. However, little is known on the best way to develop a culturally sensitive game that targets children in Africa—where sexual education is still taboo—that would be engaging, effective, and acceptable to parents and caretakers.

Objective: This study aimed to develop a socioculturally appropriate, mobile-based game for educating young children (<5 years) and parents and caretakers in Tanzania on sexual abuse prevention.

Methods: HappyToto children's game was co-designed with 111 parents and caretakers (females: n=58, 52.3%; male: n=53, 47.7%) of children below 18 years of age and 24 child experts in Tanzania through surveys and focus group discussions conducted from March 2020 to April 2020. From these, we derived an overview of topics, sociocultural practices, social environment, and game interface designs that should be considered when designing child sexual abuse prevention (CSAP) education interventions. We also conducted paper prototyping and storyboarding sessions for the game's interface, storylines, and options. To validate the application's prototype, 32 parents (females: n=18, 56%; males: n=14, 44%) of children aged 3-5 years and 5 children (females: n=2, 40%; males: n=3, 60%) of the same age group played the game for half an hour on average. The parents undertook a pre-post intervention assessment on confidence and ability to engage in CSAP education conversations, as well as exit surveys on the usability and sociocultural acceptability of the game, while children were quizzed on the topics covered and their enjoyment of the game.

Results: Parents and caregivers showed interest in the developed game during the conducted surveys, and each parent on average navigated through all the parts of the game. The confidence level of parents in talking about CSAP increased from an average of 3.56 (neutral) before using the game to 4.9 (confident) after using the game. The ability scores, calculated based on a range of topics included in CSAP education talks with children, also increased from 5.67 (out of 10) to 8.8 (out of 10) after the game was played. Both confidence level and ability scores were statistically significant ($P<.001$). All 5 children were interested in the game and enjoyed the game-provided activities.

Conclusions: The HappyToto game can thus be an effective technology-based intervention for improving the knowledge and skills of parents and children in CSAP education.

XSL•FO
RenderX

(JMIR Serious Games 2021;9(4):e30350) doi:10.2196/30350

KEYWORDS

child sexual abuse; social cultural belief; ecological setting; prevention; parents; caretakers; child experts; mobile game

Introduction

Background

Child sexual abuse (CSA) is increasingly becoming a serious public health issue affecting both boys and girls. It has been associated with physical, mental, and behavioral health problems that may be long lasting throughout the course of a child's life [1]. According to global evaluations and meta-analyses, the rates of CSA are estimated to be between 7.6% and 7.9% for boys and 18.0% to 19.7% for girls [2,3]. According to studies conducted in Uganda, India, Israel, and the United States, the COVID-19 pandemic has resulted in an upsurge in CSA [4-6]. In Tanzania, almost 3 out of every 10 females and 1 out of every 7 males experience some form of sexual harassment before the age of 18 [7]. Records from Tanzania's Ministry of Health, Community Development, Gender, Elderly and Children (MoHCDGEC) show that the numbers have fluctuated regularly but have consistently increased: from 4423 cases in 2016 to 4307 in 2017 and from 9408 cases in 2018 to 10,750 cases in 2019. Such a trend is worrying and demands a whole-society approach in solving CSA. In most cases globally, the perpetrator has a close relationship with the abused child (eg, they are a parent, friend, or dating partner) [8]. Likewise, in Tanzania, the perpetrator is usually someone familiar to the child, such as a neighbor, teacher, dating partner, or relative (uncle or cousin) [7,9].

When children are sexually abused, it affects their physical, mental, and psychological health. There have been reports of both short-term effects, such as pain, bleeding, and harm to private parts, including tearing and even discharge; and long-term effects, such as psychological abuse, emotional torture, low academic performance, becoming sexual offenders, depression, anxiety, infertility in girls, and the acquisition of chronic diseases such as HIV/AIDS [1]. In recognition of these serious effects of sexual violence and exploitation of children, the United Nations sustainable development goals include a target (Target 16.2) to "end abuse, exploitation, trafficking and all forms of violence against and torture of children." The World Health Organization, in partnership with various partners, also developed INSPIRE, a set of 7 evidence-based strategies that can be used by communities and countries to eliminate various forms of violence against children [10]. Among the strategies recommended is educating children about sexual abuse and its prevention.

CSAP education (CSAPE) is not a new concept. School-based programs have been adapted into the education systems of many countries, and they have proven to be very effective in transmitting information to children [11,12]. In Ireland for example, children aged between 6 and 12 years old are educated on topics such as safe and unsafe touches, not keeping secrets from parents, and the danger of strangers [13]. In North America, the United Kingdom, Australia, and New Zealand, school children are taught sexual abuse recognition, appropriate and inappropriate touches, the difference between good and bad secrets, how to say no and avoid unwanted advances, how to report abuse to an adult, and how abuse is never the fault of a child [12,14]. In Tanzania, no specific programs exist in the school curriculum to teach children about sexual abuse prevention [1]. However, schools may implement their own programs if needed. Unfortunately, organizing professional visits for students specifically for CSAPE is costly in terms of time and money and may therefore be out of reach for the children enrolled in primary schools, where scarcity of resources sometimes means even essential items, such as classrooms and desks, are unavailable [11,12,15].

Apart from school-based programs, CSAPE can also be conducted through places of worship and media campaigns. In Tanzania, a the 2019-2020 report by the MoHCDGEC highlighted that churches, mosques, books, and campaigns from various ministries and government agencies, among others, had offered CSAPE programs in the 2019-2020 period to 72,832 children. Similarly, in the United States, a community preventative technique that includes home visits and technological enhancement tools (eg, mobile phones, websites, phoning, SMS text messaging) to educate parents on positive parenting has shown encouraging results [16-18].

A key limitation of school and other public-based programs is that younger children are not yet in school. Younger children who are not yet in school (usually under 6 years old) are not considered in the development of the CSAPE programs. Yet, sexual abuse among this age group is present and can be perpetuated by myths, such as the belief that HIV infection can be cured through sexual contact with a child [9]. The lesser-developed communication skills of this age group may also mean that they are unable to report incidents of abuse. Home-based CSAPE is thus important to ensure their safety. From the age of 1 to 6 years, children grow sexually, recognize their biological differences, exhibit sexual gestures, and become explorative [19]. Children between 3 and 6 years look to their parents and caretakers for answers to questions such as the difference between girls and boys. Studies have proposed that parents should create home environments where they are engaged and ready to respond to and openly discuss these kinds of topics [20-23]. Child psychologists recommend that talks on sexual abuse with this age group should focus on the same topics offered to older children, including those regarding private parts, safe and unsafe touches, the importance of not keeping secrets from parents and caregivers, requesting permission before going anywhere, and reporting situations where they did not feel safe, and that these talks be offered in an easier language [24]. For many parents, this can be a daunting task. In Tanzania, there are additional sociocultural barriers, such as the discussion of sex being generally taboo, especially that occurring between parents and caretakers and their children [25].

Digital games have increasingly been used to offer health interventions to children and their parents. Mobile learning

XSL•FO
RenderX

games are games created with the defined goal of deploying on a mobile phone to motivate and engage society, especially children, on a predetermined topic [26]. Mobile learning games provide learning anywhere and anytime, and allow children to take more risks [27] and present the opportunity to retry following failure [26]. Mobile learning games geared toward health include eBug and modified Mario Brothers games. Mobile learning games specifically targeting sexual abuse prevention include Cool and Safe, Orbit, and SAP MobAPP. With the rising mobile and internet use penetration in Tanzania (29 million citizens in June 2020), mobile learning games can be a useful resource for providing CSAPE to parents and their younger children. However, none of these games have been developed with the African sociocultural context in mind. For instance, body parts like the buttocks are referred to using their proper names, which is widely unacceptable in Africa, and extended families are not featured. Their general acceptability and effectiveness may therefore be hindered.

Objectives

This study aimed to develop a CSAPE mobile learning game to target the 3 to 5-year-old group, tailored for Tanzania as a case study. The acceptance of CSAPE game is associated with cultural relevance [28] and co-designed with users of digital apps. Therefore, the game was designed with a mixed methods approach to identify parents' and caretakers' requirements, as well as with child experts' recommendations for the game content and design [29]. The following research questions were thus examined: What are the parents', caretakers', and child experts' topics and content related to CSAPE that should be contained in the game? What are the challenges related to parents', caretakers', and child experts' sociocultural and environmental condition that should be considered in the CSAPE in Tanzania? What are the parents', caretakers, and child experts' requirements for the game interface, features, and interaction designs? What would the acceptability and usability of a CSAPE mobile app be to society?

Related Work

Examples of computer-aided sexual abuse prevention games for children include Cool and Safe [30], which is an effective web-based prevention training program containing film clips, stories, tasks, and games developed for French and German elementary school children. Safety for Kids 3 is another mobile animation app to teach children about safety in their environment, including preventing CSA [31]. Other educational games include Orbit, which is Australian web-based software for children aged 8 to 10 years that offers information on togetherness, listening, understanding, belief, and courage [27]. There is also SAP MobAPP, a mobile app for primary school children in Korea [28]; and the Child Sexual Abuse Prevention Education Using Hybrid Application, which includes a combination of a website and a mobile app as a behavioral self-protection tool for fifth graders in South Korea [32]. Stewards of Children is also another combination of website and mobile app used for raising awareness on the prevention, recognition, and reaction to CSA using real people's stories [33].

In designing game-based learning, cognitive, affective, motivational, and sociocultural theoretical foundations play a significant role [26]. The cognitive foundation suggests that all the tasks and activities in the game should reflect the lessons to be learned. The motivation foundation emphasizes the need for the game to have activities that will interest and engage a player in the tasks. Moreover, the affective foundation explains how features of a game make a player feel, thus affecting their learning (eg, warm colors that may attract a child). Lastly, social and cultural foundations have a role to play in game design by connecting it to the player for easy learning. Understanding a community's sociocultural practices can significantly contribute toward a reliable and appropriate prevention program. Certain traits exist in Tanzanian communities that should be considered when developing a prevention program, such as family structures, sex and gender taboos (modesty, gender roles), sexual norms, separation (male and female), the culture of silence, foreign influence, religious teachings, unquestionable obedience by children, the value of children, poverty, and lack of explicitness in teachings [34]. As a result, any preventive measures must consider the child's social environment (working together with the entire environment to which a child is exposed), and the appropriateness of any such approach must include parents [35]. The design must also consider children's engagement in the game (flow theory: balancing between boredom and anxiety) [36]. This flow includes the magic circle of game-based learning, which involves response, feedback, and challenge [26].

Methods

A preliminary survey was conducted in 3 regions of Tanzania: Dar es Salaam in Ubungo district, Arusha in Meru district, and Morogoro in Mahenge district, along with other regions with a very small percentage of participation. The regions were selected to provide the sociocultural and social setting requirements for game development. Dar es Salaam was selected because of the presence of multicultural interactions in an urban setting, Morogoro was chosen because of multicultural interactions in a rural setting. and validation was done in Arusha where there is also multicultural interactions in an urban setting. The first phase of the survey, completed in Dar es Salaam, Morogoro, and other areas, involved questionnaire and focus group discussion from March 2020 to April 2020. The second phase of the survey was completed for validation and was conducted in Arusha from March 2020 to April 2020.

Sample Size

A total of 172 people participated in the study. In the first phase, 111 parents from various districts (Ubungo, Ulanga, Meru, and others) were involved in the study survey, including 58 females (52.3%) and 53 males (47.7%). Twenty four child experts forming four focus groups from two districts (Ubungo and Ulanga) took part in discussion from March 2020 to April 2020 to design the game content, features, interface, and interactions. In the second phase, 32 parents and 5 children from Meru district participated in the validation. Participants were given information about the facts and benefits of the research before they agreed to participate.

XSL•FO
RenderX

Sampling and Recruitment

To obtain a representative sample of the population in the first phase survey, a snowball sampling technique was used [37]. Parents were selected from strategic locations: outside of their children's schools, hospitals, market places, home visits. and online Google forms via email contact lists or WhatsApp groups for answering questionnaires. Participants were asked if they would be willing to recommend another person who fulfilled the criteria, was willing to take part, and lived nearby. Similarly to the online questionnaire, participants were requested to forward it to others who met the criteria. We recognized that participation in the study may have unintended implications, such as the likelihood of emotional distress for persons who have experienced sexual abuse. To minimize this, both the introduction and the form of consent stated that the nature of questions was focused solely on app design and not on experiences. The survey questions were reviewed and approved by Kibong'oto Infectious Diseases Hospital-Nelson Mandela African Institution of Science and Technology-Centre for Educational Development in Health, Arusha Health Research Ethics Committee. The ability to exit the survey at any time was emphasized. Child experts were selected from the local government gender desks to district social service offices. One or two hosts were identified and informed about the study, and they recommended which experts should be included in the focus group discussion. During validation in the second-phase survey, parents were sampled using the snowballing technique during home and office visits.

Inclusion Criteria

To be included in the study, participants had to be 18 years of age or older, own a smartphone, and have at least 1 child aged 1 to 18 years. Both male and female parents and caretakers were eligible. To be regarded as a child expert, the individual had to possess the necessary abilities and knowledge to protect children's safety and well-being. These participants completed a first-stage survey for the requirement selection.

To be included in the second-phase survey, participants had to be 18 years of age or older, own a smartphone, and have at least 1 child aged 3 to 5 years. Both male and female parents and caretakers were eligible. Before participating in the survey, participants were required to sign a consent form after they were informed about the study.

Parents and Caretakers Requirements

For data collection in the first survey, which was conducted face to face and online, a paper questionnaire with 33 questions separated into 4 sections was prepared (Multimedia Appendix 1). There were different types of questions, including dichotomous-response (yes or no), multiple-choice, and open-ended questions. Section 1 collected demographics, such as age, gender, marital status, number of children, and educational background. No personal information was collected, but parents and caretakers who agreed to take part in additional research phases were asked to provide their names and contact information separately after completing the survey. Parents and caretakers were free to provide or not provide such information. Section 2 included questions aimed to determine current CSAPE practices in Tanzania, such as parents' skills, knowledge, the child's age at which they are comfortable for their children to begin learning CSA prevention education, and challenges and enablers of parents talking to their children about CSAPE. Game specifications, mobile games, general games that children enjoy playing, and whether or not parents allow their children to use their phones were all discussed for Tanzanian children in Section 3. This section also featured elements that encourage a child to play a specific game, as well as play mode preferences for parents or caretakers and their children. The draw-write-tell method was used in this section to allow parents and caregivers to contribute in the design of the gamed user interface (UI) [38]. Section 4 inquired about the social environment factors that contribute to CSA in Tanzanian families, such as family structures and how often parents examine their children for symptoms of abuse.

Child Expert Requirements

The purpose of the focus groups discussion was to try to understand the challenges for parents and caretakers in talking to their children about CSAP, topics to include in CSAPE, and how the game should be designed (Multimedia Appendix 2). In each session, the facilitator introduced the purpose, and then participants could discuss them while the facilitator made sure everyone was heard. Interviews with child care experts were then conducted. The interviews were semistructured with open-ended questions and split into 4 sections. Section 1 consisted of demographic information. Section 2 comprised questions on the existing policies and methods used to protect children, topics and sociocultural requirements for prevention education against child sexual exploitation, and the obstacles that come with it in Tanzania. Section 3 asked about the social environment that contributes to child sexual abuse. Section 4 was about game UI design and using the draw-write-tell method to explain their design ideas [38]. To promote co-design with the mobile users, we also conducted paper prototyping and storyboarding sessions for the game's interface, storylines, and options in focus groups discussions with child experts.

Game Development Approach

Transforming the user's requirements into a software platform is referred to as development process. In this study, the game development life cycle (GDLC) model [39] was used as the development process. The GDLC involves preproduction, production, and postproduction phases. The benefit of GDLC is that it can handle the multidisciplinary nature of game development, which includes a mix of art, sound, control systems, human factors, and artificial intelligence to form a creative concept in order to entertain and teach, as opposed to conventional software development which aims at solving problems [40,41]. Due to time and budget constraints, the scope of the app was limited to include content for 3 to 5-year-old children and parents acting as a guiding presence.

The preproduction phase involved the requirements data collection. After data analysis, game design and game prototype were used to put together a comprehensive design document detailing the game goals, storyline, fun factors, level designs, gameplay mechanics, and overall blueprint of the game [40]. The production phase consisted of the development of different

XSL•FO
RenderX

game components including storyboard production, assets creation, development platform, game engine, programming, implementation of HappyToto, and gameplay. The evolutionary prototyping approach was used, where the first version is created after obtaining input from the users, and then subsequent versions are made with additional functionality. The tools and technologies used in development of the game included a flutter platform that provided the design for the game by combining artwork, graphics, sound, code, and engine. It is a single code-based platform app that can be used to build software for both Android and iOS devices [42]. Dart, a Google-developed object programming language, was also used to develop the game. Dart focuses on front-end creation for mobile and web applications. Additionally, Adobe Illustrator software was used to create and draw 2D vector graphic objects, while Adobe After Effect was used to develop the motion graphics and visual effects of the game. Sound was recorded by sound recorder software. The postproduction phase consisted of game beta testing and validation. The researcher and 2 parents tested HappyToto, after which development and errors were rectified.

Game Validation

Validation for this HappyToto game took place between February 2021 and March 2021. The application prototype was validated using a survey with 32 parents of children aged between 3 and 5 years, including 18 females (56%) and 14 males (44%); and with 5 children, including 2 girls (40%) and 3 boys (60%) of the same age group. A pre- and postintervention assessment was used to determine the comfort of parents in allowing their children to use HappyToto as an education toolkit for CSAPE, parents' usability of the game, and children's' perception of the game (Multimedia Appendix 3). Questionnaires, interviews, and observations were used for both intervention assessments. The questionnaire was used to ask parents about their comfort in talking to their children about CSAP before using the game. The parents were then allowed to engage with the game while being led by the researcher. Noninterruptive data were collected based on observation as the user interacted with the system, and all the comments made by users while using the game were noted. Parents were asked to score their confidence in talking to their children about CSA after playing the game, and usability questions were asked. Open-ended questions were also asked, such as their favorite feature, the least-liked feature, and what could be added or improved to make the game more appealing and appropriate. The data collected from children included how well they could follow instructions and navigate through the game and their interest in levels while playing.

Data Analysis

Questionnaires, interviews, and observations were used to collect both qualitative and quantitative data. For qualitative data, the inductive method for qualitative analysis was used. For quantitative data, descriptive statistics (frequency) were used to investigate the data while *t* tests were conducted using R software (The Foundation for Statistical Computing) [43] to detect any significance difference between the ranks of CSAPE description from parents before and after using the game.

Ethical Considerations

Before taking part in the study, the qualified participants were informed of their rights, free will to participate, the type of research, what data were needed, how their data would be published in aggregate, and that there would be no adverse impact whether they chose to participate or not. Participation was limited to the parents and child experts whose informed consent was signed, while for the children, this involved a trusted adult. The HappyToto game development study received ethical clearance for each stage of its development and validation from Kibong'oto Infectious Diseases Hospital-Nelson Mandela African Institution of Science and Technology-Centre for Educational Development in Health, Arusha Health Research Ethics Committee.

Results

Parent, Caretaker, and Child Expert Requirements

A total of 111 parents and caretakers (females: 58, 52.3%; male: 53, 47.7%) were distributed across the regions of Arusha (n=11, 9.9%), Dar es Salaam (n=54, 48.7%), Morogoro (n=20, 18%), and others (Dodoma, Iringa, Kagera, Kigoma, Kilimanjaro, Manyara, Mbeya, Mtwara, Mwanza, Pwani, Shinyanga, Songwe, and Zanzibar: 26, 23.4%) during phase 1 of the survey. Interviews involved 24 participants from Dar es Salaam (n=19, 79%) and Morogoro (n=5, 21%) and included 18 (75%) females and 6 males (25%) as indicated in Table 1.

XSL•FO
RenderX

Table 1. Socioeconomic and demographic characteristics of participant parents, caretakers, and experts.

Characteristic	Phase 1 (requirement gathering)		Phase 2 interview (validation), n (%) (N=32)
	Survey, n (%) (N=111)	Interview, n (%) (N=24)	
Regions reached			
Dar-es Salaam	54 (48.7)	19 (79)	0 (0)
Morogoro	20 (18)	5 (21)	0 (0)
Arusha	11 (9.9)	0 (0)	32 (100)
Others[a]	26 (23.4)	0 (0)	0 (0)
Gender identity			
Women	58 (52.3)	18 (75)	18 (56)
Men	53 (47.7)	6 (25)	14 (44)
Age (years)			
20 -30	43 (38.7)	4 (17)	14 (44)
31- 40	40 (36)	15 (63)	14 (44)
41 and above	28 (25.2)	6 (21)	4 (12)
Education			
Primary school	9 (8.1)	3 (13)	10 (31)
Secondary school	12 (10.8)	5 (21)	10 (31)
Tertiary level	81.1 (90)	16 (67)	12 (38)
Employment			
Employed	70 (63.1)	N/A[b]	21 (66)
Social officer	N/A	19 (79)	N/A
Teacher	N/A	3 (13)	N/A
Police officer	N/A	1 (4)	N/A
House maid	N/A	1 (4)	N/A
Self-employed	35 (31.5)	N/A	10 (31)
Farmer	18 (5.4)	N/A	1 (3)
Number of children			
1	43 (38.7)	N/A	12 (38)
2	27 (24.3)	N/A	11 (34)
3	17 (15.3)	N/A	5 (16)
4	9 (8.1)	N/A	3 (9)
5 and above	15 (13.5)	N/A	1 (3)

[a]Including Dodoma, Iringa, Kagera, Kigoma, Kilimanjaro, Manyara, Mbeya, Mtwara, Mwanza, Pwani, Shinyanga, Songwe, Zanzibar.
[b]N/A: not applicable.

Topic Suggestion

The main topics which parents, caretakers, and child experts were interested in were dangers of receiving presents (83/111, 74.8%), modesty and body changes during different stages of growth and development (71/111/, 64.4%), what is an abuse and who can be an abuser (70/111, 63.0%), dangerous people and environments and what to do if encountered (63/111, 57.0%), safe and healthy versus unsafe and unhealthy touches (60/111, 54.1%), abandoning bad traditions (43/111, 38.5%), private parts names (32/111, 28.9%), fear of God (16/111, 14.8%) and finally confidence (12/111, 11.1%). The locations that parents, caretakers, and child experts felt children spent the most time (social environment criteria) and could be vulnerable to abuse were the “home,” “school,” “on the way to/from school,” “playing outside with friends,” and “in religious institutions.” As a result, they felt the game should discuss the recommended topics in these settings (eg, receiving gifts at school, on the way home).

XSL•FO
RenderX

Sociocultural Requirement

Around 70 of the 111 parents, caretakers, and child experts (63.1%) feared that exposure to CSAPE would affect children's innocence and thus potentially make them more sexually active. Other concerns highlighted by parents, caretakers, and child experts were that they did not have enough skills to teach their children (46/111, 41.5%) and that Tanzanian society did not favor openness to children (40/111, 39.3%) and therefore felt it prudent to not expose children to taboo subjects. The other fear included busy schedules (21/111, 19.3%), globalization (16/111, 14.1%), the CSA threat not being that great (13/111, 11.9%), poverty (11/111, 9.6%), and increase of existing temptations from society (10/111 8.9%).

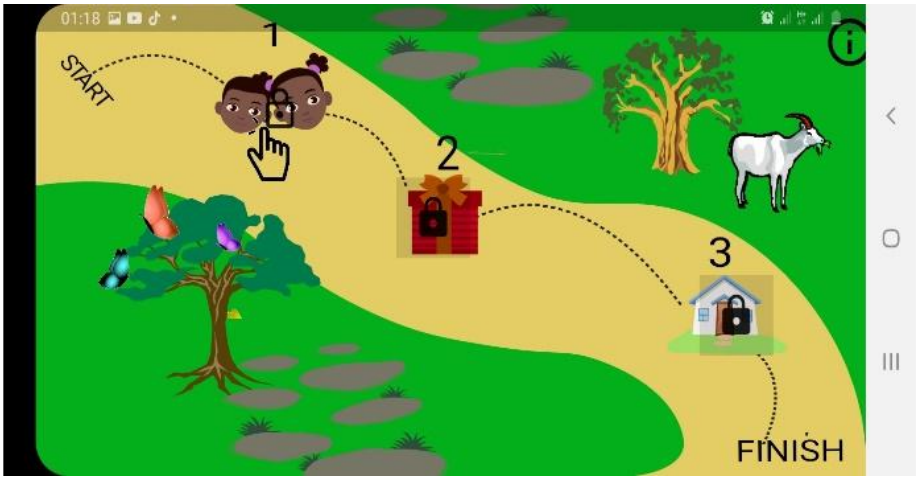
UI Design From Parents, Caretakers, and Child Experts

Parents, caretakers, and child experts designed their ideas for the game in the first phase of the survey by drawing, writing, and telling [38] what they wanted to see as the game's UI. The UI sketches and explanations are used as guidance when designing the game's UI and include layout and environment design. Parents, caretakers, and child experts also preferred the Tanzanian setting, characters, and objects surrounding the children; use of Swahili language; and the and less explicit content. Furthermore, a preference survey for operating systems was conducted, with Android (Google; 93/111, 83.8%) and iOS (Apple; 9/111, 14.4%) being the most preferred.

HappyToto Game Design and Development

HappyToto is the name given to the mobile game developed for 3 to 5-year-old children. The story script was composed in Swahili and English languages with consideration to the parents', caretakers' and child experts' opinions from the surveys and the stipulated child prevention education from previous studies. The game consists of 3 levels: private parts, presents or gifts, and a safe environment (see Figure 1).

Figure 1. Game landing page (levels page).



Storyline and Storyboard Production

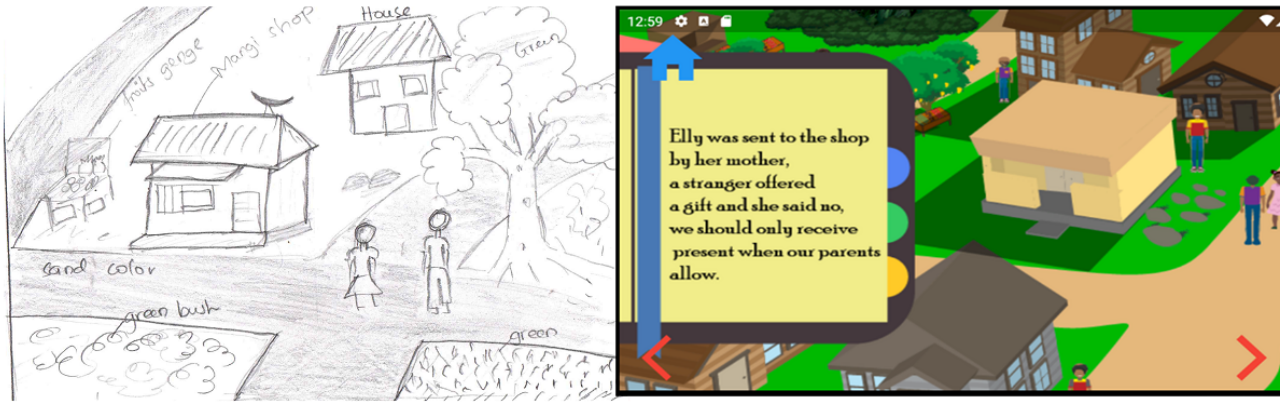
The storyline is one developed based on today's Tanzania where Ibra and Elly are twins from a middle-class household. As they are playing outside, their mother teaches them different lessons related to private parts. On the second level, their mother teaches them about proper presents to receive in the flipping game and about not receiving presents from unfamiliar people. After remembering the lessons their mother taught them, the story ends with the mother teaching Ibra and Elly about safe places and safe people inside the house and how they behave outside their home. Following each lesson, there is an exercise to check understanding. If the exercise is passed, the level is complete.

HappyToto contains 2 main cartoon-type characters Ibra (a boy) and Elly (a girl) with body shape, facial features, hair, brown skin, and clothes that reflect Tanzanian children that are 5 years old. Tanzanian children like to play and to listen to their parents. Other characters include a mother who is not physically present in the game but whose orders are narrated, a guest, and a stranger.

Game scenarios were sketched with paper and pencil for each setup in the levels, character, environment, and story flow for different scenes to depict Tanzanian settings. During the phase 1 survey, parents, caretakers, and child experts participated in the paper and pencil design of the game's UI (see Figure 2).

XSL•FO
RenderX

Figure 2. Sketch of the outdoor environment and 2D of the outdoor environment.



Assets Creation

The game assets included audio, interface design, and icons that are stored within the app. Background audio was obtained from open-source libraries, and narratives, instructions, and applause were recorded using a voice recorder application. The framework provided the user interface widgets, including icons for the next, back, and home buttons. The other user interface resources included images, video, buttons, and pop-up notifications. The environment and characters were created as 2D vector graphics in Adobe Illustrator using child-friendly colors like red, blue, and yellow. Adobe After Effect was used to create the game's video segment, including merging 2D objects and storylines and animating them.

App Screens

HappyToto consists of 2 welcome screens, where the first contains information about the game and the next contains the levels of the game. There is also an information button at the top right for adults in case they want more resources on CSA prevention, and access requires users to solve a mathematical problem. Level 1 consists of 4 screens: private parts, safe touch, dressing, and self-care. Level 2 consists of 4 screens: presents, flipping, presents outdoors, and dragging. Level 3 consists of 7 screens: safe places outdoors, safe places in the market, safe places indoors, safe places in the house, who to ask for help, safe people, and reactions when someone makes you sad.

Gameplay

When the level 1 button is pressed, a video on private parts with a play button instructs the player to play. The video tells the tale of what private parts are, and the next screen contains a safe touch exercise. The next screen contains a dressing game in which a child must drag clothes to the 2 characters. The final screen of level 1 includes 2 clickable characters that, when pressed, provide a narrative of what activities Ibra and Elly may do on their own and whom to ask for assistance. When the next button is pressed it gives instructions to activate the second round. When a child lands on level 2, narratives about receiving presents begin, and the following screen features a flipping card game that shows examples of gifts that a child might receive. The subsequent screen narrative focuses on the dangerous places to receive presents and is followed by an image-dragging exercise in which after a picture is completed, after which a lesson on what a child should be mindful of is given. The level ends with instructions on the landing page on level 3, which is subsequently unlocked. Level 3 starts with a narrative on a safe environment outdoors, followed by hide and seek as an exercise to locate safe and unsafe places. The fourth page in this level starts with a narrative on a safe environment indoors, followed by a hide-and-seek exercise to locate safe and unsafe places and people in the house. The end screen of the game describes what Ibra and his friends should do upon encountering dangerous people. The game has a flow mode storyline which follows the same children across different days, as this was the favored mode from the survey. There is also looping music played at each level to continue engaging a child.

Game Validation Results

Thirty-two parents in the Arusha Region participated in the phase 2 survey for validation. Of the participants, 44% (14/32) were men, 56% (18/32) were women, their ages ranged from 20 to 50 years, 31% (10/32) had received primary education, 31% (10/32) had received secondary education, and 38% (n=12/32) had received tertiary education (Table 2). Children who participated in the evaluation were between 3 and 5 years old, with 60% (3/5) being male and 40% (2/5) being female.

XSL•FO
RenderX

Table 2. Comparison scores for confidence level and ranked results before and after using the game.

Education level	Confidence			Ranked results[a]		
	Topic 1	Topic 2	Topic 3	Topic 1	Topic 2	Topic 3
Primary level						
Pretest mean	3.20	4.20	4.3	5	5.6	5.8
Posttest mean	4.90	4.9	5.00	9	8.9	8.8
Secondary level						
Pretest mean	3.7	4.2	4.2	6.3	6.6	6.6
Posttest mean	5	5	5	9.2	9.2	9.2
Tertiary level						
Pretest mean	2.58	3.08	2.83	4.75	5.16	5.25
Posttest mean	4.83	4.83	4.92	8.38	8.29	8.38

[a]Two-tailed ($T \leq t$) P value <.001.

Game Acceptability

Parents were generally happy with the game; out of 32 interviewed, 81% (n=26) scored as a "satisfied parent" (awarding it a 5 out of 5), while a score of 4 out of 5 ("somewhat satisfied") was given by 19% (n=6), as shown in Table 2. Two of the participants (T3 and T2) requested permission to download the game to their phones immediately. One hundred percent of the participants (n=32) would allow their children to play the game immediately if they were around and said they would recommend this game to other parents. On the features generally, 90% (n=29) of the parents liked game level 1, 81% (n=26) liked game level 2, and 78% (n=25) liked game level 3. The reasons for parents' preference for all the levels were the way the lessons were well integrated into the game activities (15/32, 50%). The safe touch activity, dragging, story narration, and visuals used were all liked by parents. The storyline for teaching sensitive topics in a creative and appropriate way and the in-game activities were praised by all participants (n=32). For example, participant S8 said, "I love how the lessons incorporated in hiding and seek feature because children can relate with it," and T11 commented, "It is very educative to children especially the private part lesson and the dragging features."

The private parts topic was completely acceptable because "colloquial words used" and the names of private parts were not called explicitly, respectful and motherly language was used, and the game conformed to Tanzanian culture, with 100% (n=32) of the participants rating it 5 out of 5 ("appropriate"). Also, 94% (n=30) of the participants ranked the "receiving presents" subject 5 out of 5 ("appropriate") and the "dangerous environment and people" subject was ranked 5 out of 5 ("appropriate") by 84% (n=27) of all participants. The language (Swahili), not explicitly naming private parts, characters and climate, indoor and outdoor settings, and family structures were the basic features listed as conforming to the culture.

In general, the participants discussed some features including respectful, friendly, and nonexplicit language; age-appropriate lessons; inclusion of local games such as "kombolela" (hide and seek); and graphics of Tanzanian characters and environments as very useful for displaying the UI.

Although they were satisfied, when asked "What is your least favorite feature?" and "What features should be added/modified," 28% (n=9) of participants indicated technical corrections such as the need for more instructions, icons, animation, and game analytics. For instance, participant T4 said, "I wish there were more animation in the game and more restrictions in the level where a child will play for some time before going to another level"; T6 said, "You can add game analytics to allow parents so see the progress of their children and what features they enjoy and play a lot." Twenty-eight percent (n=9) of the participants recommended modification to the story by adding more scenes and adding detail on how to talk to children. Participants T10 and P3 suggested to "add more dangerous environment and dangerous people scenes" and "add scenes where the emphasis is put on a child to report to a parent how their day went," respectively. Nineteen percent (n=6) recommended variation of a narrator sound incorporating a man's and child's voice with strictness in the unacceptable behavior. Participant P4 suggested, "You need to add more lessons related songs to make the game more fun."

When asked about the possible effects of the game, 84% (n=27) of the participants rated it a 1 out of 5 ("not affected"). This was because the game was engaging but not to the point of addiction, and because it is offline, children cannot reach other sites. The remaining 16% (n=5) of participants were concerned about the effects of light from the mobile phone as well as addiction to the game.

Pre- and Postconfidence and Ability Scores

Before and after using the app, the trust level in educating children about private parts, obtaining gifts, and unsafe environments of people was measured using a 5-point Likert scale ranging from 1 (uncomfortable) to 5 (comfortable), as shown in Table 2.

The ranked results were obtained when parents were asked objectively how they intended to educate a child, and the results were ranked by taking into account the key points in each topic

XSL•FO
RenderX

and guidelines from MoHCDGECs National Parenting Education Manual for Families. The parents' opinions of the game before and after using it are summarized in Table 3. Prior to interacting with the game, parents had different views on the topics. Some parents assumed that their children were still too young (3 years) and that they would not understand until they were approximately 5 years old or teenagers. Some parents spoke with their children in order to protect them because occurrences on the streets had given the impression that their children were not resistant to violence. Of the 32 parents who were asked whether they were familiar with the cartoons that their children enjoyed watching and playing, 30 parents indicated having seen popular Swahili language children's cartoons (eg, *Akili and Me* and *Ubongo Kids*).

Table 3. Selected parents opinion before and after using the game.

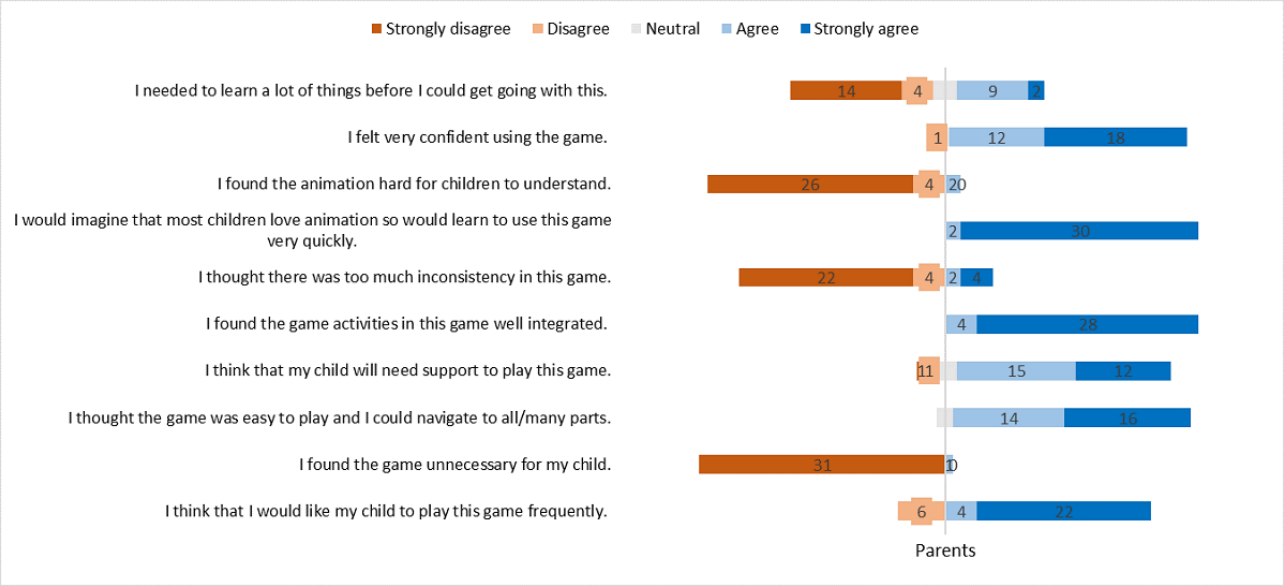
Participant characteristic	Exemplar quote
Attitude before playing the game	
31 to 40-year-old woman	"I know there are effects, but I do not want to be very open to the child unless they ask; I am a doctor, so I know what to do and the effects. I do normal disciplining and emphasize a safe environment and people."
41 to 50-year-old man	"I provide child self-protection like no touch of the pupu [anus] and susu [vagina] parts; only receive presents from parents and no one else."
31-40-year man	"I do not talk to them about these issues; their mother does, and they are still young to be taught about sexual abuse."
41 to 50-year-old man	"They were not exposed to a dangerous environment, so there was no reason to tell them, but as their brother is approaching adolescence, he has started asking questions about his body changes."
41 to 50-year-old female	"I feel like my child is too young to understand all dangerous environments. He understands some, but not all the scenarios are easy to explain to him."
21 to 30-year-old female	"The neighboring child was abused, and there was a case I do not know how it ended; I have to scare my child so that she is not abused."
31 to 40-year-old female	"These abuses happen in town areas. My children live in the interior village, so they are very safe".
Overall acceptability	
31 to 40-year-old female	"The language is very well understood, attractive environment, lessons are well understood, and it matches with the age of the child."
31 to 40-year-old female	"The language used is not very explicit in naming the private parts; polite way."
41 to 50-year-old female	"This is very creative way of teaching."
Addiction potential	
20 to 30-year-old female	"For what I see this game, a child can't not be addicted because there are no features to continuously engage a child. It is just for lesson and may be repeat a few times."

Usability

Most parents felt confident using the game after reading the instructions. It was suggested that children will also need instructions during the first-time interacting with the game (see Figure 3).

XSL•FO
RenderX

Figure 3. Usability test results.



Children Usability Results

All 5 children enjoyed flipping, dragging, and the ability to move around the game. When the level page loaded and they could hear the background sound, all the children became curious. After playing the game, the 4 to 5 year olds were able to correctly answer questions such as “What do you do if a neighbor tells you to take off your clothes?” A researcher met a 5-year-old girl after a day, and she showed an interest in playing the game again. They all expressed the need for more animations and activities or actions in the game, similar to what they had seen in movies to hold their attention while playing.

Discussion

We have attempted to develop and validate a user-friendly children’s game app for prevention of sexual abuse. The distinguishing features for this study compared to other CSA prevention studies are the following: it is the first study in Tanzania to consider development and validation of the acceptability and usability of a game as a prevention education tool for children aged 3 to 5 years, Tanzania's culture, society, and social environment were considered in the design of the game, the content of the game touched on topics that parents and child experts indicated would be suitable, and a mobile game was used to improve child's learning experience without time or space limitations. Although other research has found significant results when using games [30], these studies considered other cultural contexts. CSAPE game assessment needs to be performed in Tanzania due to cultural differences and the increasing numbers of parents now using smartphones and allowing their children to play games on their phones. CSAPE games can provide another tool for creatively educating CSAP preschool children and keeping them engaged. We discovered that parents were afraid of taking away their children's innocence but also wanted them to learn about presents, body changes as they grow older, who an abuser is, reporting, and safe touches in a fun and culturally sensitive way. After the game was developed, parents and children were very interested in how the game was structured and showed gained skills. Herein, we discuss the usefulness and efficacy of the HappyToto game in CSA prevention.

All 32 parents were eager to introduce CSAPE to their 3 to 5 year old using HappyToto as a CSA prevention method because the lessons were age appropriate [19], but parents also believed that good parenting was the best way to protect children [20]. Proper parenting can be defined in 2 ways: (1) teaching children how to defend themselves (ie, teaching them to report to their parents when someone makes them feel uncomfortable) [20] and (2) limiting parenting for child safety, such as prohibiting children from going to specific locations where there is a high risk of seeing a perpetrator and imposing a curfew time to be back at home [20]. During the presurvey, each parent expressed an opinion about parenting in their personal situation. Without having seen a definition of parenting, each parent described parenting as one or both of the 2 definitions above. It is important to remember that even the notion of “proper parenting” is a transitory statement, as the definition of parenting has evolved as a result of the changing social environment and globalization. Most parents confessed to being aware of abuse but having difficulty discussing it openly with their young children [44] (ie, on the private parts of their bodies). Parents also expressed struggling to describe concepts such as “your uncle could abuse you.”

Previous research has shown that parents find it interesting how games can be used to teach prevention education and encourage their children to use them and agree that this idea should be used in the concept and design of programs and applications [45]; we therefore used this concept in this study and received positive feedback. It was surprising to see the parent, caretaker, and child expert participants’ contributions to CSA prevention tailored to a game, perhaps because they had not used such a game before. The role of a parent as a guiding presence while children play allows a parent to participate in CSA prevention [9] by providing additional clarification to children on concepts. HappyToto also fills the skills gap, where parents previously depended on skills learned as children [25], which were

XSL•FO
RenderX

insufficient. HappyToto provides the skillset parents need to be more transparent and ready to overcome the taboo of not talking to children about issues of child sexual abuse. The confidence level of parents on talking about CSAPE with children significantly increased after using HappyToto ($P<.001$), as compared to before interacting with HappyToto, with the confidence level increasing from an average of 3.56 "neutral" before using the game to a 4.9 "confident" after using the game. Furthermore, when asked to describe how they would talk to their child before using the HappyToto game they scored an average of 5.67 out of 10 as compared to 8.8 out of 10 after using the game.

HappyToto offered several suggestions for talking to children about challenging topics in a nondirective manner that parents were willing to try. Due to its cultural appropriateness, the opportunity to connect to the game played an essential role in ensuring that it was acceptable to parents. Beginning with language (a woman in the game softly narrated in Swahili, Tanzania's official and national language), a safe environment was created for children to feel comfortable with the content [46]. Although CSAPE recommends that children be explicitly taught about healthy and improper interaction by giving proper names to private parts, Tanzanian culture's conservative nature makes it much more suitable to use colloquial words in the game [25,34]. The deliberate use of cartoons with brown skin color who are traditionally dressed provided commonly demanded values for society's children. The inclusion of fathers in the CSA prevention criteria debunked the idea that mothers were solely responsible for their children's education and gave fathers a chance to participate [47]. Parents who liked the game the least indicated that it was due to a lack of details in the lessons they would like to see and expressed the need for additional functionality.

In the first phase survey, parents with a secondary level education had an average of 3.9 ("somewhat confident") while parents with a primary level education had an average of 4 ("somewhat confident"), indicating that they were more interested in educating their children about CSAPE than were parents with a tertiary level education, who had an average of 2.83 ("neutral"). This may possibly be due to the community lifestyle where the children are more likely to play in streets [45] and where people with a higher education are more likely to live in an area with fenced perimeters, where children generally do not move outside of their parent's view. It was okay, for example, for 3 to 5-year-old children who did not live inside the fences to receive a lesson in the game on how to behave on the walk to the shop because this is where sexual harassment occurs [9]. The fact that the game depicted real locations where abuse occurs was appealing to both parents and children, as it made the game relatable while still making the lessons easy to remember [28].

The storylike game with a narrative flow and basic activities were sufficiently fun for the children to play and help understand lessons while not being too addicting [46]. This feature increased parents' eagerness to encourage their children to play, knowing that they are still going to balance the game's use with other activities. Although the repetition function is necessary to ensure that children remember the material, this needs to be a balanced with the risk of addiction.

Parents were likely comfortable with the cartoon presentation and narration of HappyToto because the game was designed with the interests of preschoolers being considered [48]. The HappyToto design is also similar to culturally appropriate and relatable children educational cartoons in Swahili, such as *Akili and Me* and *Ubongo Children*. Suggested improvements from participants were to include variations in the voices of women, men, and children but also strictness in the voice when anything that should not be done was being discussed so that children can be properly informed [47].

Several limitations became evident in this study. First, only parents from 3 regions in Tanzania were included, and cultural differences between ethnic groups and communities may introduce diversity in the development and validation of this app. Tanzania has approximately 125 ethnic groups, and future studies could consider the inclusion of participants from other regions. Second, the study was carried out with no control group. Future studies should consider using a case-control model to eliminate potential confounders not evident in this study.

Additionally, although parents responded to the questionnaire immediately after using the game, there may be a disparity between the questionnaire response and parents' behavior in real life when talking to their children. Moreover, measures of courtesy and recall bias were not performed in this study to determine if children could remember the materials. Further study would benefit from longitudinal evaluation of children's behavior following use of the game. Finally, the COVID-19 pandemic has changed the world's social ecosystem, including social restrictions and more indoor activities resulting in fewer interactions. The COVID-19 global pandemic also limited the total number and demographic of participants that could be included the study.

Parents are aware that they need to teach their children about sexual education, yet they are unaware of the most appropriate way to deliver this information. Hence, there is a need for game developers to design and validate different approaches to CSAPE based on the sociocultural differences of the target population. This study has contributed to the empirical discourse in designing an appropriate platform and media application against CSA using preventive games. HappyToto was designed in collaboration with game users, and was consequently able to incorporate the social cultural perspectives of the target community. Thus, the design and activities in the HappyToto game complement previous efforts in CSAPE in a culturally relevant way. This is important in ensuring that parents are comfortable with and adopt the use of CSAPE materials. There is a need for government and other appropriate authorities to support needed research on more attractive interactive features in teaching prevention and encouraging children to repeat the activities just enough to be able to remember the skills without becoming addicted to the game. Parents need to be informed that children as young as 3 years can be taught prevention education and can understand this education. This developed game will be further piloted in a wider population of children to assess its effectiveness. After the large-scale pilot and analysis

XSL•FO
RenderX

of comments on the game design and content included in the game are completed, the game will be made available for free on the App Store and Google Play. Following this, we will develop a monitoring, assessment, and continuous improvement strategy, and increase awareness through schools and the media. Furthermore, we will perform technological demonstrations to private and governmental organizations that deal with children's welfare on a local and national level to promote awareness of CSAP and educate on how to use mobile technology to prevent CSA.

Acknowledgments

We would like to express our gratitude to the African Development Bank for partially funding this research study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire for parents and caretakers on child sexual abuse prevention education.
[PDF File (Adobe PDF File), 268 KB - games_v9i4e30350_app1.pdf]

Multimedia Appendix 2
Questionnaire for child experts on child sexual abuse prevention education.
[PDF File (Adobe PDF File), 265 KB - games_v9i4e30350_app2.pdf]

Multimedia Appendix 3
Questionnaire to parents and caretakers on the validation of the game as an educational tool for child sexual abuse prevention.
[DOCX File , 36 KB - games_v9i4e30350_app3.docx]

References

1. Kisanga F, Nystrom L, Hogan N, Emmelin M. Child sexual abuse: community concerns in urban Tanzania. J Child Sex Abus 2011 Mar;20(2):196-217. [doi: 10.1080/10538712.2011.555356] [Medline: 21442533]
2. Pereda N, Guilera G, Forns M, Gómez-Benito J. The prevalence of child sexual abuse in community and student samples: a meta-analysis. Clin Psychol Rev 2009 Jun;29(4):328-338. [doi: 10.1016/j.cpr.2009.02.007] [Medline: 19371992]
3. Stoltenborgh M, Bakermans-Kranenburg MJ, Alink LRA, van IJzendoorn M. The prevalence of child maltreatment across the globe: review of a series of meta-analyses. Child Abuse Rev 2014 Oct 20;24(1):37-50. [doi: 10.1002/car.2353]
4. Tener D, Marmor A, Katz C, Newman A, Silovsky JF, Shields J, et al. How does COVID-19 impact intrafamilial child sexual abuse? Comparison analysis of reports by practitioners in Israel and the US. Child Abuse Negl 2021 Jun;116(Pt 2):104779 [FREE Full text] [doi: 10.1016/j.chiabu.2020.104779] [Medline: 33143870]
5. Sserwanja Q, Kawuki J, Kim JH. Increased child abuse in Uganda amidst COVID-19 pandemic. J Paediatr Child Health 2021 Feb;57(2):188-191. [doi: 10.1111/jpc.15289] [Medline: 33296120]
6. Poddar S, Mukherjee U. Ascending child sexual abuse statistics in India during COVID-19 lockdown: a darker reality and alarming mental health concerns. Indian J Psychol Med 2020 Sep;42(5):493-494 [FREE Full text] [doi: 10.1177/0253717620951391] [Medline: 33414605]
7. UNICEF Tanzania, Centers for Disease Control and prevention, Muhimbili University of Health and Allied Sciences. Violence against Children in Tanzania: Findings from a National Survey, 2009. UNICEF, Dar es Salaam. Dar es Salaam: UNICEF; 2011. URL: https://resourcecentre.savethechildren.net/document/violence-against-children-tanzania-findings-national-survey-2009/ [accessed 2021-05-06]
8. A familiar face: Violence in the lives of children and adolescents. United Nations Children's Fund. New York, USA: UNICEF; 2017. URL: https://data.unicef.org/wp-content/uploads/2017/10/EVAC-Booklet-FINAL-10_31_17-high-res.pdf [accessed 2021-06-25]
9. Ezekiel MJ, Kisanga F, Mosha IH, Anaeli A, Kamazima SR, Mpembeni R, et al. Factors associated with child sexual abuse in Tanzania: a qualitative study. Tanzania J Hlth Res 2017 Mar 27;19(2):1-10. [doi: 10.4314/thrb.v19i2.9]
10. Violence against children. World Health Organization. 2020 Jun 08. URL: https://www.who.int/news-room/fact-sheets/detail/violence-against-children [accessed 2021-09-05]
11. Bustamante G, Andrade MS, Mikesell C, Cullen C, Endara P, Burneo V, et al. "I have the right to feel safe": evaluation of a school-based child sexual abuse prevention program in Ecuador. Child Abuse Negl 2019 May;91:31-40. [doi: 10.1016/j.chiabu.2019.02.009] [Medline: 30822629]
12. Zhang H, Shi R, Li Y, Wang Y. Effectiveness of School-Based Child Sexual Abuse Prevention Programs in China: A Meta-Analysis. Research on Social Work Practice 2021 Jul 08;31(7):693-705. [doi: 10.1177/10497315211022827]

XSL•FO
RenderX

13. MacIntyre D, Carr A, Lawlor M, Flattery M. Development of the Stay Safe programme. Child Abuse Review: Journal of the British Association for the Study and Prevention of Child Abuse and Neglect 2000 May;9(3):200-216. [doi: 10.1002/1099-0852(200005/06)9:3<200::aid-car596>3.0.co;2-t]
14. Buller AM, Pichon M, McAlpine A, Cislaghi B, Heise L, Meiksin R. Systematic review of social norms, attitudes, and factual beliefs linked to the sexual exploitation of children and adolescents. Child Abuse & Neglect 2020 Jun;104:104471. [doi: 10.1016/j.chiabu.2020.104471]
15. Lyanga AA, Chen M. The Impacts of Fee - Free Education Policy in Junior Secondary Schools in Tanzania. AJESS 2020 Dec 15;13(3):36-47. [doi: 10.9734/ajess/2020/v13i330333]
16. Self-Brown S, Osborne MC, Boyd JC, DeVeausse Brown N, Rostad W, Patterson A, et al. The impact of SafeCare® Dads to Kids program on father maltreatment risk and involvement: Outcomes and lessons learned from an efficacy trial. Child Abuse Negl 2018 Sep;83:31-41. [doi: 10.1016/j.chiabu.2018.06.014] [Medline: 30016743]
17. Lefever JEB, Bigelow KM, Carta JJ, Borkowski JG, Grandfield E, McCune L, et al. Long-term impact of a cell phone-enhanced parenting intervention. Child Maltreat 2017 Nov;22(4):305-314. [doi: 10.1177/1077559517723125] [Medline: 28845676]
18. Cowart-Osborne M, Jackson M, Chege E, Baker E, Whitaker D, Self-Brown S. Technology-based innovations in child maltreatment prevention programs: examples from SafeCare®. Soc Sci (Basel) 2014 Aug 15;3(3):427-440 [FREE Full text] [doi: 10.3390/socsci3030427] [Medline: 25606347]
19. Cacciatore RSM, Ingman-Friberg SML, Lainiala LP, Apter DL. Verbal and behavioral expressions of child sexuality among 1-6-year-olds as observed by daycare professionals in Finland. Arch Sex Behav 2020 Oct;49(7):2725-2734. [doi: 10.1007/s10508-020-01694-y] [Medline: 32300911]
20. Rudolph J, Zimmer-Gembeck MJ. Parents as protectors: A qualitative study of parents' views on child sexual abuse prevention. Child Abuse Negl 2018 Nov;85:28-38. [doi: 10.1016/j.chiabu.2018.08.016] [Medline: 30180986]
21. Rudolph J, Zimmer-Gembeck M, Shanley DC, Hawkins R. Child sexual abuse prevention opportunities: parenting, programs, and the reduction of risk. Child Maltreat 2018 Feb;23(1):96-106. [doi: 10.1177/1077559517729479] [Medline: 28920456]
22. Ezekiel MJ, Kisanga F, Mosha IH, Anaeli A, Kamazima SR, Mpembeni R, et al. Factors associated with child sexual abuse in Tanzania: a qualitative study. Tanzania J Hlth Res 2017 Mar 27;19(2):1-10. [doi: 10.4314/thrb.v19i2.9]
23. Leclerc B, Feakes J, Cale J. Child sexual abuse in youth-oriented organisations: tapping into situational crime prevention from the offender's perspective. Crime Sci 2015 Oct 16;4(1):1-12. [doi: 10.1186/s40163-015-0044-3]
24. Preventing child sex abuse. Center of Disease Control and Prevention.: CDC; 2021 Apr 20. URL: https://www.cdc.gov/violenceprevention/childsexualabuse/fastfact.html [accessed 2021-08-21]
25. Bastien S, Kajula LJ, Muhwezi W. A review of studies of parent-child communication about sexuality and HIV/AIDS in sub-Saharan Africa. Reprod Health 2011 Sep 24;8:25 [FREE Full text] [doi: 10.1186/1742-4755-8-25] [Medline: 21943095]
26. Plass JL, Homer BD, Kinzer CK. Foundations of game-based learning. Educational Psychologist 2016 Feb 06;50(4):258-283. [doi: 10.1080/00461520.2015.1122533]
27. Scholes L, Jones C, Stieler-Hunt C, Rolfe B. Serious games for learning: games-based child sexual abuse prevention in schools. International Journal of Inclusive Education 2014 Jan 21;18(9):934-956. [doi: 10.1080/13603116.2013.860195]
28. Moon KJ, Park KM, Sung Y. Sexual abuse prevention mobile application (SAP_MobAPP) for primary school children in Korea. J Child Sex Abus 2017 Jul;26(5):573-589. [doi: 10.1080/10538712.2017.1313350] [Medline: 28661824]
29. Principles for digital development. Digital Principles. URL: https://digitalprinciples.org/ [accessed 2021-08-05]
30. Müller AR, Röder M, Fingerle M. Child sexual abuse prevention goes online: Introducing "Cool and Safe" and its effects. Computers & Education 2014 Sep;78:60-65. [doi: 10.1016/j.compedu.2014.04.023]
31. Child abuse prevention-Kid Safety 3 android gameplay. YouTube. URL: https://www.youtube.com/watch?v=gPlS4yAiub0 [accessed 2021-05-08]
32. Kang SR, Kim S, Kang K. Effects of child sexual abuse prevention education program using hybrid application (CSAPE-H) on fifth-grade students in South Korea. J Sch Nurs 2020 Jul 21:1059840520940377. [doi: 10.1177/1059840520940377] [Medline: 32691681]
33. Stewards of children. Darkness to light: End child sexual abuse. 2021. URL: https://www.d2l.org/education/stewards-of-children/ [accessed 2021-05-09]
34. Plummer CA, Njuguna W. Cultural protective and risk factors: professional perspectives about child sexual abuse in Kenya. Child Abuse Negl 2009 Aug;33(8):524-532. [doi: 10.1016/j.chiabu.2009.02.005] [Medline: 19758700]
35. Garbarino J. The ecology of human development. In: Children and Families in the Social Environment. New Brunswick, USA: A Division of Transaction Publishers; 1999:10-33.
36. Kiili K, De Freitas S, Arnab S, Lainema T. The Design Principles for Flow Experience in Educational Games. Procedia Computer Science 2012;15:78-91. [doi: 10.1016/j.procs.2012.10.060]
37. Snowball sampling: definition, method, advantages and disadvantages. QuestionPro.: QuestionPro; 2021. URL: https://www.questiopro.com/ [accessed 2021-10-05]
38. Angell C, Alexander J, Hunt JA. 'Draw, write and tell': A literature review and methodological development on the 'draw and write' research method. Journal of Early Childhood Research 2014 Aug 05;13(1):17-28. [doi: 10.1177/1476718x14538592]

XSL•FO
RenderX

39. Aleem S, Capretz LF, Ahmed F. Game development software engineering process life cycle: a systematic review. J Softw Eng Res Dev 2016 Nov 9;4(1):1-30. [doi: 10.1186/s40411-016-0032-7]
40. Ramadan R, Widyani Y. Game development life cycle guidelines. : IEEE; 2013 Sep Presented at: 2013 International Conference on Advanced Computer Science and Information Systems; Sep 28, 2013; Sanur Bali, Indonesia p. 95-100 URL: https://ieeexplore.ieee.org/document/6761558 [doi: 10.1109/icacsis.2013.6761558]
41. Jain S. Game development life cycle. LinkedIn.: LinkedIn; 2017 May 01. URL: https://www.linkedin.com/pulse/game-development-life-cycle-sumit-jain/ [accessed 2021-04-18]
42. Kuzmin N, Ignatiev K, Grafov D. Experience of developing a mobile application using Flutter. In: Information Science and Applications. Singapore: Springer; 2020:571-575.
43. Integrated Development Environment for R. RStudio. Boston, MA; 2021. URL: http://www.rstudio.com/ [accessed 2021-11-01]
44. Mlekwa FM, Nyamhanga T, Chalya PL, Urassa D. Knowledge, attitudes and practices of parents on child sexual abuse and its prevention in Shinyanga district, Tanzania. Tanzania J Hlth Res 2016 Nov 23;18(4):1-9. [doi: 10.4314/thrb.v18i4.6]
45. Madu SN. Child sexual abuse and implications for the ecology of human development. Journal of Human Ecology 2017 Oct 24;14(4):253-262. [doi: 10.1080/09709274.2003.11905621]
46. Zhou C, Occa A, Kim S, Morgan S. A Meta-analysis of Narrative Game-based Interventions for Promoting Healthy Behaviors. Journal of health communication 2020;25(1):54-65. [doi: 10.1080/10810730.2019.1701586] [Medline: 31829829]
47. Panter-Brick C, Burgess A, Eggerman M, McAllister F, Pruett K, Leckman JF. Practitioner review: Engaging fathers--recommendations for a game change in parenting interventions based on a systematic review of the global evidence. Journal of child psychology and psychiatry, and allied disciplines 2014 Nov;55(11):1187-1212 [FREE Full text] [doi: 10.1111/jcpp.12280] [Medline: 24980187]
48. Arbianingsih, Rustina Y, Krianto T, Ayubi D. Developing a health education game for preschoolers: What should we consider? Enfermería Clínica 2018 Feb;28(1):1-4. [doi: 10.1016/s1130-8621(18)30025-1]

Abbreviations

CSA: child sexual abuse
CSAP: child sexual abuse prevention
CSAPE: child sexual abuse prevention education
GDLC: game development life cycle
MoHCDGEC: Tanzania's Ministry of Health, Community Development, Gender, Elderly and Children
UI: user interface.

Edited by N Zary; submitted 12.05.21; peer-reviewed by M Antoniou, JM Suelves; comments to author 16.07.21; revised version received 05.09.21; accepted 25.10.21; published 08.11.21.

Please cite as:
Malamsha MP, Sauli E, Luhanga ET
Development and Validation of a Mobile Game for Culturally Sensitive Child Sexual Abuse Prevention Education in Tanzania: Mixed Methods Study
JMIR Serious Games 2021;9(4):e30350
URL: https://games.jmir.org/2021/4/e30350
doi:10.2196/30350
PMID:34747703

©Maria Proches Malamsha, Elingarami Sauli, Edith Talina Luhanga. Originally published in JMIR Serious Games (https://games.jmir.org), 08.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effects of an Immersive Virtual Reality Exergame on University Students' Anxiety, Depression, and Perceived Stress: Pilot Feasibility and Usability Study

Wenge Xu[1], BSc, MSc, PhD; Hai-Ning Liang[2], BA, PhD; Nilufar Baghaei[3,4], BSc, PhD; Xiaoyue Ma[2], BSc; Kangyou Yu[2], BSc; Xuanru Meng[2], BSc; Shaoyue Wen[2], BSc

[1]Digital Media Technology Lab, School of Computing and Digital Technology, Birmingham City University, Birmingham, United Kingdom
[2]Department of Computing, School of Advanced Technology, Xi'an Jiaotong-Liverpool University, Suzhou, China
[3]School of Natural and Computational Sciences, Massey University, Auckland, New Zealand
[4]School of Information Technology and Electrical Engineering, University of Queensland, Brisbane, Australia

Corresponding Author:
Hai-Ning Liang, BA, PhD
Department of Computing
School of Advanced Technology
Xi'an Jiaotong-Liverpool University
111 Renai Road
Suzhou Industrial Park
Suzhou, 215123
China
Phone: 86 51288161516
Email: haining.liang@xjtlu.edu.cn

Abstract

Background: In recent years, there has been an increase in the number of students with depression, anxiety, and perceived stress. A solution that has been increasingly used for improving health and well-being is exergaming. The effects and acceptability of exergames have been studied widely but mostly with older adults. The feasibility and usability of exergames among university students, especially those of immersive virtual reality (iVR) exergames, remain unexplored.

Objective: This study aimed to explore the feasibility of a 6-week iVR exergame–based intervention in reducing anxiety, depression, and perceived stress among university students and to examine the usability and acceptability of such games.

Methods: A total of 31 university students were recruited to participate in a 6-week study in which they needed to play a boxing-style iVR exergame called *FitXR* (FitXR Limited) twice per week (30 minutes per session). Their anxiety (Beck Anxiety Inventory), depression (Beck Depression Inventory-II), and perceived stress (Perceived Stress Scale) levels were measured before and after intervention.

Results: A total of 15 participants completed the 6-week study. Our results suggested that participants' mean depression scores decreased significantly from 8.33 (SD 5.98) to 5.40 (SD 5.14) after the intervention (P=.01). In addition, most participants (14/15, 93%) believed that the iVR exergame has good usability. Furthermore, most participants (14/15, 93%) were satisfied with the iVR gameplay experience and would play the iVR exergame again in the future. Of the 15 participants, 11 (73%) would recommend the iVR exergame to their friends.

Conclusions: The results gained from this study show that the iVR exergame has good usability, is highly acceptable, and has the potential to reduce depression levels among university students.

(JMIR Serious Games 2021;9(4):e29330) doi:10.2196/29330

KEYWORDS

university students; depression; anxiety; stress; immersive virtual reality; exergame

XSL•FO
RenderX

Introduction

Background

There is growing evidence that the number of university students who are struggling with mental health problems is increasing globally (eg, in North America [1], Japan [2], and the United Kingdom [3]). Among these mental health problems, university students are particularly at high risk for anxiety, depression, and stress [2,4-9]. These mental health problems can result in different consequences among students, such as university dropout, decreased academic performance and social functioning, and even suicidal behavior [10,11].

One solution that is widely accepted is to provide support via counseling centers located within university campuses. Consistent with the rising rates of depression and anxiety, the number of students seeking mental health services in campus counseling centers has increased. Research has shown that the percentage of students who have received treatment at a university counseling center has increased from 6.6% in 2007 to 11.8% in 2017 [12]. Consequently, these counseling centers often report that they are over capacity and are unable to immediately meet the needs of the large number of students who are requesting services. This is because counseling centers have limitations such as a limited number of available sessions, long waiting lists, and a limited amount of staff [13], which can result in unwanted outcomes for students with mental health problems.

Sports-based interventions are useful for achieving health benefits—both mental and physical benefits—and provide more flexibility and ease of access for patients than those provided by counseling centers. Exercises involving large muscle groups in the whole body, especially those that involve following rhythmic flow patterns, can effectively alleviate depression [14]. A typical example of this type of exercise is boxing. Using boxing as a form of therapy can help with stress and anger management, boost confidence and self-esteem, elevate mood, serve as a natural antianxiety activity, and improve focus and sleep quality [15]. Boxing training exercises may have a better therapeutic effect on obesity, cardiovascular, and health-related quality of life outcomes than an equivalent dose of brisk walking [16]. In addition, boxing may be the most beneficial to those who are going through the early to middle stages of Parkinson disease progression [17]. For instance, therapeutic boxing can positively affect the speech, social interaction skills, and mental health of individuals with Parkinson disease [18]. It can also improve older adults' gait, balance, daily living activities, and quality of life [19]. Further, one study suggests that shadow boxing (ie, the practice of committing repetitive boxing movements to muscle memory), together with psychosomatic relaxation, has a beneficial auxiliary therapeutic effect on depression and anxiety among people with type 2 diabetes [20].

In recent years, exergaming, which combines video games and physical exercises, has been widely used to promote both physical and mental health [21,22] in different population groups (ie, children [23], young individuals [24], and older adults [25]). However, most studies have been conducted with non–immersive virtual reality (iVR) exergames [26] (ie, those played on flat-screen televisions or computer monitors) [27-29]. On the other hand, the use of iVR exergame–based interventions is still underexplored.

iVR exergames [30-32] have been gaining attention rapidly due to the recent emergence of affordable iVR head-mounted displays. iVR exergames have many advantages over non-iVR exergames; they can provide more positive game experiences to players compared to those provided by non-iVR exergames [32,33]. Furthermore, exercising within iVR can result in higher increases in enjoyment and motivation compared to those resulting from playing exergames with standard televisions or computer monitors [34]. As such, exergaming in iVR might increase people's adherence to regular physical exercise in general [35-37].

Goal of This Study

The goal of this study was to evaluate the usability and acceptability of an iVR exergame (FitXR; FitXR Limited) [38] for university students. We also wanted to examine the feasibility and usability of the iVR exergame by conducting a 6-week pilot study on reducing anxiety, depression, and perceived stress levels among university students.

Methods

Recruitment

Students were recruited from a local university campus through physical and digital advertisements (ie, posters, social media platforms, and a mailing list). The inclusion criteria were as follows: (1) enrolled as a full-time student, (2) aged at least 18 years, and (3) was not pregnant (because of the physical exertion required to play the game). The exclusion criterion was a "yes" answer in the Physical Activity Readiness Questionnaire [39].

Intervention

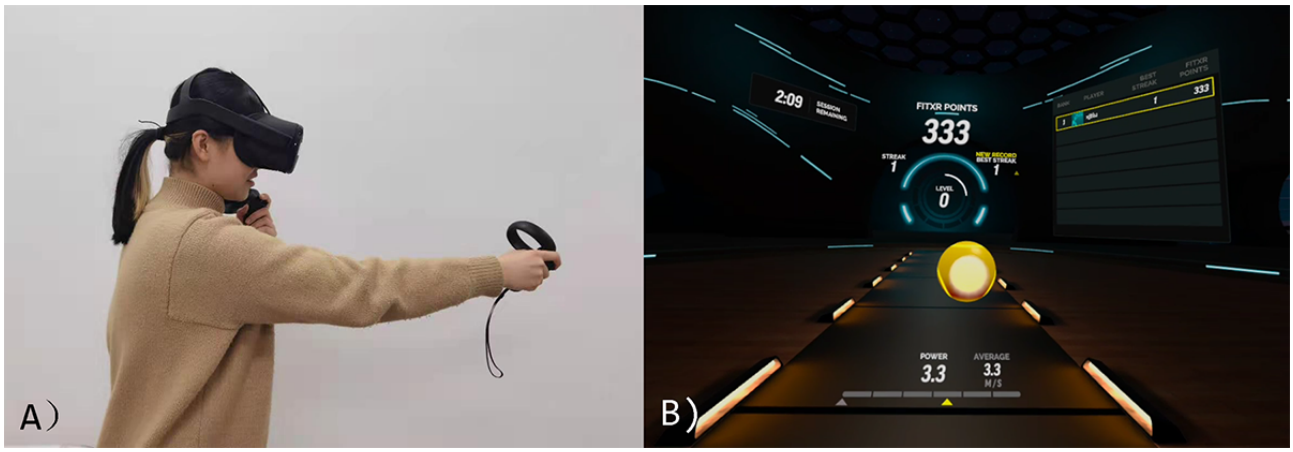
Eligible participants were invited to an indoor laboratory room that could not be seen from the outside. Participants first completed an in-person consent form. Then, they needed to fill in the preexperiment questionnaire, which was used to collect their demographic information and anxiety, depression, and perceived stress baselines (see *Outcome Measures* section for details on the questionnaires used in the experiment). An experimenter helped participants put on an Oculus Quest 1 (Oculus) when they completed the preexperiment questionnaire and instructed them on how to play the FitXR exergame (Figure 1). Once they were familiar with the device and the game, they started their first training session.

Eligible participants were scheduled for the iVR exergame intervention twice per week (not on the same day) for 6 weeks. Each session consisted of about 30 minutes of gameplay. The game provided 11 game levels that ranged in duration from 26 minutes to 31 minutes. We selected the game levels that had around 30 minutes of gameplay. Unlimited drinks and snacks were provided during each session. At the end of the last session, participants were asked to fill out the postexperiment questionnaire, which was used to measure participants' anxiety, depression, and perceived stress levels and the acceptability and usability of the game.

XSL•FO
RenderX

FitXR was selected because (1) it is a boxing-inspired iVR fitness game; (2) it involves a great number of jabs, uppercuts, defensive actions (ie, cover-ups), squats, and leftward and rightward movements; (3) it involves few rotational movements, which is helpful for potentially avoiding motion sickness [40]; and (4) it provides several unique training sessions from which players can choose. To play the game successfully, players need to perform several punching combos (jab or uppercut) on sphere objects and avoid getting hit by blocks by performing squats and lunges.

Figure 1. (A) An example of a participant playing the immersive virtual reality exergame by using the Oculus Quest 1 (Oculus). (B) A screenshot of the exergame.



Outcome Measures

Anxiety

The Beck Anxiety Inventory (BAI) [41] is a 21-item self-report scale that ranges from 0 to 63, with higher scores indicating more severe anxiety. The BAI has achieved a Cronbach α of .92, thereby demonstrating its internal consistency [41,42]. A review paper on anxiety questionnaires indicated that the BAI is suitable for the general population, has good reliability, and has moderate validity [43]. Further, the BAI has been used in studies dealing with exercise [44-46] and in exergame studies [22].

Depression

We measured depression levels by using the Beck Depression Inventory-II (BDI-II) [47], which consists of a 21-item self-report scale that ranges from 0 to 63, with higher scores indicating more severe depression. This scale has a Cronbach α coefficient of .92 and a test-retest reliability value of 0.93. A review paper on depression questionnaires indicated that the BDI-II is suitable for the general population, and it has excellent reliability and good validity [48]. The BDI-II has been used in studies dealing with exercise [44,45,49].

Perceived Stress

Perceived stress levels were measured via the widely used Perceived Stress Scale [50], which consists of 14 items questions with scores that range from 0 (never) to 4 (very often). This questionnaire has been used in studies dealing with exercise [45,51,52] and with exergame-based interventions [53,54].

Usability and Acceptability

Usability and acceptability were only tested at the end of the last session via self-reported questionnaires. Usability was measured by using the System Usability Scale (SUS). SUS scores range between 0 and 100; 100 represents the best usability, and a score of ≥68 is considered positive [55].

We measured acceptability and satisfaction with the following items, which were answered by using a 5-point Likert scale (1 indicated “extremely disagree” and 5 indicated “extremely agree”): “I was satisfied with FitXR experience,” “I would play FitXR again in the future,” and “I would recommend FitXR to a friend.” Finally, we asked the participants the following questions: “How often would you play FitXR per week” and “How long would you play FitXR each day?”

Statistical Power and Analysis

Analyses were performed by using SPSS version 24.0 (IBM Corporation). The normality of the data was tested by using Shapiro-Wilk tests. We used 2-tailed paired *t* tests, in which time (week 1: pretest; week 6: posttest) was used as the within-subjects variable, if the differences in the dependent variable between pretest and posttest times were normally distributed. Otherwise, the Wilcoxon signed-ranks test was used.

Results

Participants’ Characteristics

A total of 31 participants volunteered to take part in this study, and in the end, 15 completed it. The reasons for dropout included urgent family reasons (n=1); health-related reasons, such as discomfort due to menstrual periods (n=2); high coursework workloads (n=2); complaints that the selected game was not as good as expected (n=2); the belief that the exertion level of the game was too high (n=2); and physical reasons (n=1). Further, 6 people did not provide a reason for dropping out. Table 1 shows the demographics of the participants and their experience with iVR and exergames.

XSL•FO
RenderX

Table 1. Characteristics of the study participants.

Characteristic	Value
Students, n (number of males)	15 (8)
Age (years), mean (SD)	19.1 (0.96)
BMI, mean (SD)	21.7 (2.98)
Participants who had self-reported experience with exergames, n	7
Frequent user (daily or weekly)	1
Participants who had self-reported experience with iVR[a] HMDs[b], n	7
Frequent user (daily or weekly)	0

[a]iVR: immersive virtual reality.

[b]HMD: head-mounted display.

Evaluation Outcomes

BAI—Anxiety

The mean anxiety score for the pretest was 4.87 (SD 3.98), and the mean anxiety score for the posttest was 5.40 (SD 5.41). The results of the 2-tailed paired samples *t* test showed that there was no significant difference between the pretest and posttest scores for anxiety (t_{14}=−0.541; *P*=.60).

BDI-II—Depression

The mean depression score for the pretest was 8.33 (SD 5.98), and the mean depression score for the posttest was 5.40 (SD 5.14). The Wilcoxon signed-rank test indicated that the posttest depression scores were significantly lower than the pretest depression scores (*Z*=−2.526; *P*=.01) and that the intervention had a large effect on depression (effect size: *r*=−0.652). Overall, after playing the iVR boxing exergame for 6 weeks, the university students in this study reported being less depressed.

Perceived Stress

The mean perceived stress score for the pretest was 16.87 (SD 4.88), and the mean perceived stress score for the posttest was 16.13 (SD 5.81). A 2-tailed paired *t* test indicated that there was no significant difference between the pretest and posttest scores (t_{14}=0.564; *P*=.58).

Usability and Acceptability

The SUS has a range of 0 to 100 and is used to rate the usability of an application. A score of ≥68 is deemed to be positive (ie, scores of 68: okay; scores of 68-80.3: good; scores of >80.3: excellent). The participants rated FitXR with a mean SUS score of 79.5 (SD 9.51); 7 participants rated it as excellent, another 7 rated it as good, and 1 rated it as poor. The highest SUS score was 97.5, while the lowest SUS score was 57.5.

With regard to the three satisfaction items, participants gave a mean score of 4.33 (SD 0.62) to the “I was satisfied with FitXR experience” item; 6 participants gave an “extremely agree” rating, 8 participants gave an “agree” rating, and 1 participant gave a “neutral” rating. As for the “I would play FitXR again in the future” item, participants gave an average score of 4.27 (SD 0.59); 5 participants gave an “extremely agree” rating, 9 participants gave an “agree” rating, and 1 participant gave a “neutral” rating. With regard to the “I would recommend FitXR to a friend” item, participants gave an average score of 4 (SD 9.26); 5 participants gave an “extremely agree” rating, 6 participants gave an “agree” rating, 3 participants gave a “neutral” rating, and 1 participant gave a “disagree” rating.

Lastly, participants indicated that they would like to play FitXR for an average of 3.87 days (SD 1.06 days) per week and would play the game for an average of 35 minutes (SD 25.93 minutes) each day.

Discussion

Principal Results

This study investigated the usability and acceptability of a boxing iVR exergame–based training intervention among university students and evaluated the intervention’s feasibility in a 6-week training program for reducing anxiety, depression, and perceived stress levels. The results of this pilot study show (1) that the FitXR iVR exergame was perceived to have good usability (ie, good usability scores) and was highly acceptable among university students, and (2) that playing the iVR boxing exergame for 6 weeks (ie, 2 sessions per week; 30 minutes per session) would likely reduce depression levels among university students.

Comparison With Prior Work

There is evidence that exergames can significantly improve depression among people with depression (ie, among adults [56] and older adults [22,57,58]). The literature also suggests that exergames can improve depression in healthy subjects. For instance, depression among healthy older adults improved after playing Wii Fit (Nintendo) exergames twice per week for 4 weeks [59]; participants in the control group who participated in an education program for the same duration did not experience such improvements. A recent review also suggested that iVR therapies are effective in supporting the treatment of anxiety and depression [60].

Although university students are at high risk for mental issues [2,4-9], limited research has looked into using exergames to promote health among university students. A previous study suggested that playing a virtual reality exergame 3 times per week for 6 weeks could improve depression in healthy university students [61]. Our results also support this finding, as the depression scores measured by the BDI-II further decreased

XSL•FO
RenderX

after participants played the iVR head-mounted display–based exergame for 6 weeks (pretest score: 8.33; posttest score: 5.40). However, we failed to obtain results that reflected whether playing an exergame could improve anxiety among healthy university students [61].

In terms of perceived stress, Huang et al [54] conducted a study to explore how playing exergames impacts the mood states (including perceived stress) of healthy university students and staff. They found that playing exergames for 30 consecutive minutes each week for 2 weeks could reduce perceived stress levels. Cutter et al [53] conducted an 8-week trial in which methadone-maintained patients played an exergame 5 times per week (session duration ranged from 20 to 25 minutes per session). They found that the exergame could reduce perceived stress levels. However, we could not obtain similar results for perceived stress.

Viana et al [62] found that playing an exergame (even for only 1 session) seemed to be a useful method for reducing state anxiety in healthy women. However, this is also not supported by our study on an iVR exergame.

With regard to usability, FitXR, in line with other exergames such as SliverFit (SUS score: mean 87.0) [63] and the rehabilitation exergame (SUS score: mean 89.6) developed by Uzor and Baillie [64], was rated with a good usability score (mean 79.5). This indicates that players can learn how to play and use the game quickly and that the game is very easy to use.

In line with a study by Yunus et al [61], we observed that university students were willing to accept playing an iVR-based exergame as a form of exercise. FitXR was perceived to be acceptable, and the majority of our participants (11/15, 73%) were glad to recommend FitXR to their friends.

The retention rate of our study was 48% (15/31). Furthermore, even though only about half of the participants (15/31, 48%) completed this study, the participants who dropped out mainly did so due to personal reasons (n=5), and only a few (n=2) complained that the game was not as good as expected. These data provide further insight into the potential of using the iVR exergame as an intervention for studying mental health among university students.

Limitations and Future Work

There are 2 main limitations in this study. First, the lack of a control group did not allow for comparisons with people who undergo traditional therapy or do not undergo therapy at all. Since this pilot study has confirmed the feasibility of using the iVR exergame to reduce depression levels, it is worth conducting future studies on the impacts that iVR exergames have on health. Such studies could, for instance, use a more established study design, such as a randomized controlled trial involving people who undergo traditional therapy and a control group (people who do not undergo therapy). Second, no follow-up tests were conducted after the experiment. Therefore, it is unclear if and how long the observed benefits were maintained.

Like many other studies on usability and acceptability [65], we tested the exergame in a supervised environment. Future work could test the game in an unsupervised setting. In addition, we aim to develop our own iVR exergame prototype involving individualized and personalized features, such as avatars, environments, and behaviors, to improve gameplay engagement further [66,67].

Conclusions

In this study, we investigated whether playing an iVR exergame for over 6 weeks could improve university students' mental health (anxiety, depression, and perceived stress). We also explored the acceptability of the game and its usability. Our results indicated that the exergame has a good usability score and is highly acceptable among university students as a form of exercise. In addition, our findings showed that playing the exergame twice per week for 6 weeks can reduce depression scores among healthy university students.

Acknowledgments

The authors thank the participants for their time and the reviewers for their insightful comments that helped improve this paper. This work was supported in part by the Xi'an Jiaotong-Liverpool University (XJTLU) Key Special Fund (grant KSF-A-03), XJTLU Research Development Fund, and Massey University Strategic Research Excellence Fund Grant.

Conflicts of Interest

None declared.

References

1. LeViness P, Bershad C, Gorman K, Braun L, Murray T. The Association for University and College Counseling Center Directors Annual Survey 2018. Association for University and College Counseling Center Directors. URL: https://www.aucccd.org/assets/documents/Survey/2018%20aucccd%20survey-public-revised.pdf [accessed 2020-01-26]
2. Tomoda A, Mori K, Kimura M, Takahashi T, Kitamura T. One-year prevalence and incidence of depression among first-year university students in Japan: a preliminary study. Psychiatry Clin Neurosci 2000 Oct;54(5):583-588 [FREE Full text] [doi: 10.1046/j.1440-1819.2000.00757.x] [Medline: 11043810]
3. Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer H. Psychiatric morbidity among adults living in private households, 2000. Int Rev Psychiatry 2003;15(1-2):65-73. [doi: 10.1080/0954026021000045967] [Medline: 12745312]

XSL•FO
RenderX

4. Nerdrum P, Rustøen T, Rønnestad MH. Student psychological distress: A psychometric study of 1750 Norwegian 1st - year undergraduate students. Scandinavian Journal of Educational Research 2007 Jan 24;50(1):95-109. [doi: 10.1080/00313830500372075]
5. Ovuga E, Boardman J, Wasserman D. Undergraduate student mental health at Makerere University, Uganda. World Psychiatry 2006 Feb;5(1):51-52 [FREE Full text] [Medline: 16757997]
6. Stewart-Brown S, Evans J, Patterson J, Petersen S, Doll H, Balding J, et al. The health of students in institutes of higher education: an important and neglected public health problem? J Public Health Med 2000 Dec;22(4):492-499. [doi: 10.1093/pubmed/22.4.492] [Medline: 11192277]
7. Rith-Najarian LR, Boustani MM, Chorpita BF. A systematic review of prevention programs targeting depression, anxiety, and stress in university students. J Affect Disord 2019 Oct 01;257:568-584. [doi: 10.1016/j.jad.2019.06.035] [Medline: 31326690]
8. Voelker R. Mounting student depression taxing campus mental health services. JAMA 2003 Apr;289(16):2055-2056. [doi: 10.1001/jama.289.16.2055] [Medline: 12709447]
9. Wong JGWS, Cheung EPT, Chan KKC, Ma KKM, Tang SW. Web-based survey of depression, anxiety and stress in first-year tertiary education students in Hong Kong. Aust N Z J Psychiatry 2006 Sep;40(9):777-782. [doi: 10.1080/j.1440-1614.2006.01883.x] [Medline: 16911753]
10. Keyes CLM, Eisenberg D, Perry GS, Dube SR, Kroenke K, Dhingra SS. The relationship of level of positive mental health with current mental disorders in predicting suicidal behavior and academic impairment in college students. J Am Coll Health 2012;60(2):126-133. [doi: 10.1080/07448481.2011.608393] [Medline: 22316409]
11. Salzer MS. A comparative study of campus experiences of college students with mental illnesses versus a general college sample. J Am Coll Health 2012;60(1):1-7. [doi: 10.1080/07448481.2011.552537] [Medline: 22171723]
12. Lipson SK, Lattie EG, Eisenberg D. Increased rates of mental health service utilization by U.S. college students: 10-year population-level trends (2007-2017). Psychiatr Serv 2019 Jan 01;70(1):60-63 [FREE Full text] [doi: 10.1176/appi.ps.201800332] [Medline: 30394183]
13. Lattie E, Cohen KA, Winquist N, Mohr DC. Examining an app-based mental health self-care program, IntelliCare for college students: Single-arm pilot study. JMIR Ment Health 2020 Oct 10;7(10):e21075 [FREE Full text] [doi: 10.2196/21075] [Medline: 33037874]
14. Steptoe A, Bolton J. The short-term influence of high and low intensity physical exercise on mood. Psychol Health 1988 Apr;2(2):91-106. [doi: 10.1080/08870448808400346]
15. Shosha M. A brief introduction to therapeutic boxing. International Journal of Physiology, Nutrition and Physical Education 2020;5(2):29-31 [FREE Full text] [doi: 10.22271/journalofsport.2020.v5.i2a.1977]
16. Cheema BS, Davies TB, Stewart M, Papalia S, Atlantis E. The feasibility and effectiveness of high-intensity boxing training versus moderate-intensity brisk walking in adults with abdominal obesity: a pilot study. BMC Sports Sci Med Rehabil 2015 Jan 16;7:3 [FREE Full text] [doi: 10.1186/2052-1847-7-3] [Medline: 25973207]
17. Morris ME, Ellis TD, Jazayeri D, Heng H, Thomson A, Balasundaram AP, et al. Boxing for Parkinson's disease: Has implementation accelerated beyond current evidence? Front Neurol 2019 Dec 04;10:1222 [FREE Full text] [doi: 10.3389/fneur.2019.01222] [Medline: 31866923]
18. Meinert R, Hatkevich B. The effect of community-based therapeutic boxing on the speech, social interaction skills, and mental health of individuals with Parkinson's disease. Am J Occup Ther 2019 Aug 01;73(4_Supplement_1):7311515312p1. [doi: 10.5014/ajot.2019.73s1-po5059]
19. Long T, Robertson T. Foundations of Therapeutic Recreation. Champaign, Illinois: Human Kinetics; 2007.
20. Zheng Y, Zhou Y, Lai Q. Effects of twenty-four move shadow boxing combined with psychosomatic relaxation on depression and anxiety in patients with type-2 diabetes. Psychiatr Danub 2015 Jun;27(2):174-179 [FREE Full text] [Medline: 26057313]
21. Pluchino A, Lee SY, Asfour S, Roos BA, Signorile JF. Pilot study comparing changes in postural control after training using a video game balance board program and 2 standard activity-based balance intervention programs. Arch Phys Med Rehabil 2012 Jul;93(7):1138-1146. [doi: 10.1016/j.apmr.2012.01.023] [Medline: 22414490]
22. Rosenberg D, Depp CA, Vahia IV, Reichstadt J, Palmer BW, Kerr J, et al. Exergames for subsyndromal depression in older adults: a pilot study of a novel intervention. Am J Geriatr Psychiatry 2010 Mar;18(3):221-226 [FREE Full text] [doi: 10.1097/JGP.0b013e3181c534b5] [Medline: 20173423]
23. Hernandez HA, Ye Z, Graham TCN, Fehlings D, Switzer L. Designing action-based exergames for children with cerebral palsy. 2013 Apr 27 Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; April 27 to May 2, 2013; Paris, France p. 1261-1270. [doi: 10.1145/2470654.2466164]
24. Street TD, Lacey SJ, Langdon RR. Gaming your way to health: A systematic review of exergaming programs to increase health and exercise behaviors in adults. Games Health J 2017 Jun;6(3):136-146. [doi: 10.1089/g4h.2016.0102] [Medline: 28448175]
25. Xu W, Liang HN, Baghaei N, Berberich BW, Yue Y. Health benefits of digital videogames for the aging population: A systematic review. Games Health J 2020 Dec;9(6):389-404. [doi: 10.1089/g4h.2019.0130] [Medline: 32589482]
26. Zeng N, Pope Z, Lee JE, Gao Z. Virtual reality exercise for anxiety and depression: A preliminary review of current research in an emerging field. J Clin Med 2018 Mar 04;7(3):42 [FREE Full text] [doi: 10.3390/jcm7030042] [Medline: 29510528]

XSL•FO
RenderX

27. Huang HC, Nguyen HV, Cheng TCE, Wong MK, Chiu HY, Yang YH, et al. A randomized controlled trial on the role of enthusiasm about exergames: Players' perceptions of exercise. Games Health J 2019 Jun;8(3):220-226. [doi: 10.1089/g4h.2018.0057] [Medline: 30418041]
28. da Silva Alves R, Iunes DH, de Carvalho JM, Menezes FDS, Silva AM, Borges JBC, et al. Effects of exergaming on quality of life in cancer patients. Games Health J 2018 Nov 30;7(6):385-392. [doi: 10.1089/g4h.2017.0174]
29. Gao Y, Mandryk RL. GrabApple: The design of a casual exergame. In: Entertainment Computing – ICEC 2011. 2011 Oct Presented at: 10th International Conference on Entertainment Computing; October 5-8, 2011; Vancouver, Canada p. 35-46. [doi: 10.1007/978-3-642-24500-8_5]
30. Xu W, Liang HN, Yu Y, Monteiro D, Hasan K, Fleming C. Assessing the effects of a full-body motion-based exergame in virtual reality. 2019 Jun Presented at: Chinese CHI 2019: The seventh International Symposium of Chinese CHI; June 27-30, 2019; Xiamen, China p. 1-6. [doi: 10.1145/3332169.3333574]
31. Xu W, Liang HN, He Q, Li X, Yu K, Chen Y. Results and guidelines from a repeated-measures design experiment comparing standing and seated full-body gesture-based immersive virtual reality exergames: Within-subjects evaluation. JMIR Serious Games 2020 Jul 27;8(3):e17972 [FREE Full text] [doi: 10.2196/17972] [Medline: 32716004]
32. Xu W, Liang HN, Yu K, Baghaei N. Effect of gameplay uncertainty, display type, and age on virtual reality exergames. 2021 May Presented at: CHI '21: CHI Conference on Human Factors in Computing Systems; May 8-13, 2021; Yokohama, Japan p. 1-14. [doi: 10.1145/3411764.3445801]
33. Xu W, Liang HN, Zhang Z, Baghaei N. Studying the effect of display type and viewing perspective on user experience in virtual reality exergames. Games Health J 2020 Dec;9(6):405-414. [doi: 10.1089/g4h.2019.0102] [Medline: 32074463]
34. Farrow M, Lutteroth C, Rouse PC, Bilzon JLJ. Virtual-reality exergaming improves performance during high-intensity interval training. Eur J Sport Sci 2019 Jul;19(6):719-727. [doi: 10.1080/17461391.2018.1542459] [Medline: 30403927]
35. Frederick CM, Ryan RM. Differences in motivation for sport and exercise and their relations with participation and mental health. J Sport Behav 1993;16(3):124-146 [FREE Full text]
36. Ryan RM, Frederick CM, Lepes D, Rubio N, Sheldon KM. Intrinsic motivation and exercise adherence. Int J Sport Psychol 1997;28:335-354 [FREE Full text]
37. Alexandris K, Tsorbatzoudis C, Grouios G. Perceived constraints on recreational sport participation: Investigating their relationship with intrinsic motivation, extrinsic motivation and amotivation. J Leis Res 2017 Dec 13;34(3):233-252. [doi: 10.1080/00222216.2002.11949970]
38. FitXR - Exercise you'll love. FitXR Limited. URL: https://fitxr.com/ [accessed 2020-04-02]
39. Thomas S, Reading J, Shephard RJ. Revision of the physical activity readiness questionnaire (PAR-Q). Can J Sport Sci 1992 Dec;17(4):338-345. [Medline: 1330274]
40. Keshavarz B, Hecht H. Axis rotation and visually induced motion sickness: the role of combined roll, pitch, and yaw motion. Aviat Space Environ Med 2011 Nov;82(11):1023-1029. [doi: 10.3357/asem.3078.2011] [Medline: 22097636]
41. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988 Dec;56(6):893-897. [doi: 10.1037//0022-006x.56.6.893] [Medline: 3204199]
42. Halfaker DA, Akeson ST, Hathcock DR, Mattson C, Wunderlich TL. 3 - Psychological aspects of pain. In: Lennard TA, Walkowski S, Singla AK, Vivian DG, editors. Pain Procedures in Clinical Practice (Third Edition). Philadelphia, Pennsylvania: Elsevier Inc; 2011:13-22. [doi: 10.1016/B978-1-4160-3779-8.10003-X]
43. Julian LJ. Measures of anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11(0 11):S467-S472 [FREE Full text] [doi: 10.1002/acr.20561] [Medline: 22588767]
44. Ólafsdóttir KB, Kristjánsdóttir H, Saavedra JM. Effects of exercise on depression and anxiety. A comparison to transdiagnostic cognitive behavioral therapy. Community Ment Health J 2018 Aug;54(6):855-859. [doi: 10.1007/s10597-017-0213-9] [Medline: 29168106]
45. Paolucci EM, Loukov D, Bowdish DME, Heisz JJ. Exercise reduces depression and inflammation but intensity matters. Biol Psychol 2018 Mar;133:79-84. [doi: 10.1016/j.biopsycho.2018.01.015] [Medline: 29408464]
46. Rawson RA, Chudzynski J, Gonzales R, Mooney L, Dickerson D, Ang A, et al. The impact of exercise on depression and anxiety symptoms among abstinent methamphetamine-dependent individuals in a residential treatment setting. J Subst Abuse Treat 2015 Oct;57:36-40 [FREE Full text] [doi: 10.1016/j.jsat.2015.04.007] [Medline: 25934458]
47. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio, Texas: Psychological Corporation; 1996.
48. Smarr KL, Keefer AL. Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PHQ-9). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S454-S466 [FREE Full text] [doi: 10.1002/acr.20556] [Medline: 22588766]
49. Legrand FD. Effects of exercise on physical self-concept, global self-esteem, and depression in women of low socioeconomic status with elevated depressive symptoms. J Sport Exerc Psychol 2014 Aug;36(4):357-365. [doi: 10.1123/jsep.2013-0253] [Medline: 25226604]

XSL•FO
RenderX

50. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404]
51. Joshi AR, Pendse TN, Vaidya SM. Effect of moderate aerobic exercise on perceived stress during luteal phase of menstrual cycle in students pursuing professional course. Natl J Physiol Pharm Pharmacol 2018;8(5):705-708 [FREE Full text] [doi: 10.5455/njppp.2018.8.1146213122017]
52. Starkweather AR. The effects of exercise on perceived stress and IL-6 levels among older adults. Biol Res Nurs 2007 Jan;8(3):186-194. [doi: 10.1177/1099800406295990] [Medline: 17172317]
53. Cutter CJ, Schottenfeld RS, Moore BA, Ball SA, Beitel M, Savant JD, et al. A pilot trial of a videogame-based exercise program for methadone maintained patients. J Subst Abuse Treat 2014 Oct;47(4):299-305 [FREE Full text] [doi: 10.1016/j.jsat.2014.05.007] [Medline: 25012555]
54. Huang HC, Wong MK, Yang YH, Chiu HY, Teng CI. Impact of playing exergames on mood states: A randomized controlled trial. Cyberpsychol Behav Soc Netw 2017 Apr;20(4):246-250. [doi: 10.1089/cyber.2016.0322] [Medline: 28394215]
55. Brooke J. SUS: A 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London: Taylor & Francis Group; 1996:6.
56. Yuen HK, Holthaus K, Kamen DL, Sword DO, Breland HL. Using Wii Fit to reduce fatigue among African American women with systemic lupus erythematosus: a pilot study. Lupus 2011 Oct;20(12):1293-1299 [FREE Full text] [doi: 10.1177/0961203311412098] [Medline: 21700656]
57. Li J, Theng YL, Foo S, Xu X. Exergames vs. traditional exercise: investigating the influencing mechanism of platform effect on subthreshold depression among older adults. Aging Ment Health 2018 Dec;22(12):1634-1641. [doi: 10.1080/13607863.2017.1385722] [Medline: 28984486]
58. Li J, Theng YL, Foo S. Exergames for older adults with subthreshold depression: Does higher playfulness lead to better improvement in depression? Games Health J 2016 Jun;5(3):175-182. [doi: 10.1089/g4h.2015.0100] [Medline: 27135146]
59. Chao YY, Scherer YK, Montgomery CA, Wu YW, Lucke KT. Physical and psychosocial effects of Wii Fit exergames use in assisted living residents: a pilot study. Clin Nurs Res 2015 Dec;24(6):589-603. [doi: 10.1177/1054773814562880] [Medline: 25488422]
60. Baghaei N, Chitale V, Hlasnik A, Stemmet L, Liang HN, Porter R. Virtual reality for supporting the treatment of depression and anxiety: Scoping review. JMIR Ment Health 2021 Sep 23;8(9):e29681 [FREE Full text] [doi: 10.2196/29681] [Medline: 34554097]
61. Yunus FW, Tan XZ, Romli MH. Investigating the feasibility of exergame on sleep and emotion among university students. Games Health J 2020 Dec;9(6):415-424. [doi: 10.1089/g4h.2019.0077] [Medline: 33301386]
62. Viana RB, Alves CL, Vieira CA, Vancini RL, Campos MH, Gentil P, et al. Anxiolytic effects of a single session of the exergame Zumba Fitness on healthy young women. Games Health J 2017 Dec;6(6):365-370. [doi: 10.1089/g4h.2017.0085] [Medline: 29077484]
63. Nawaz A, Skjæret N, Ystmark K, Helbostad JL, Vereijken B, Svanæs D. Assessing seniors' user experience (UX) of exergames for balance training. 2014 Oct Presented at: NordiCHI '14: The 8th Nordic Conference on Human-Computer Interaction; October 26-30, 2014; Helsinki, Finland p. 578-587. [doi: 10.1145/2639189.2639235]
64. Uzor S, Baillie L. Investigating the long-term use of exergames in the home with elderly fallers. 2014 Apr Presented at: CHI '14: CHI Conference on Human Factors in Computing Systems; April 26 to May 1, 2014; Toronto, Ontario, Canada p. 2813-2822. [doi: 10.1145/2556288.2557160]
65. Nawaz A, Skjæret N, Helbostad JL, Vereijken B, Boulton E, Svanaes D. Usability and acceptability of balance exergames in older adults: A scoping review. Health Informatics J 2016 Dec;22(4):911-931 [FREE Full text] [doi: 10.1177/1460458215598638] [Medline: 26303810]
66. Baghaei N, Stemmet L, Khaliq I, Ahmadi A, Halim I, Liang HN, et al. Designing individualised virtual reality applications for supporting depression: A feasibility study. 2021 Jun Presented at: EICS '21: ACM SIGCHI Symposium on Engineering Interactive Computing Systems; June 8-11, 2021; Netherlands p. 6-11. [doi: 10.1145/3459926.3464761]
67. Baghaei N, Ahmadi A, Khaliq I, Liang HN. Individualised virtual reality for supporting depression: Feedback from mental health professionals. 2021 Oct Presented at: 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct); October 4-8, 2021; Bari, Italy p. 63-67. [doi: 10.1109/ISMAR-Adjunct54149.2021.00022]

Abbreviations

BAI: Beck Anxiety Inventory
BDI-II: Beck Depression Inventory-II
iVR: immersive virtual reality
SUS: System Usability Scale
XJTLU: Xi'an Jiaotong-Liverpool University

XSL•FO
RenderX

Edited by N Zary; submitted 02.04.21; peer-reviewed by E Bellei, A Szpak; comments to author 14.05.21; revised version received 07.07.21; accepted 25.08.21; published 22.11.21.

Please cite as:
Xu W, Liang HN, Baghaei N, Ma X, Yu K, Meng X, Wen S
Effects of an Immersive Virtual Reality Exergame on University Students' Anxiety, Depression, and Perceived Stress: Pilot Feasibility and Usability Study
JMIR Serious Games 2021;9(4):e29330
URL: https://games.jmir.org/2021/4/e29330
doi:10.2196/29330
PMID:34813487

©Wenge Xu, Hai-Ning Liang, Nilufar Baghaei, Xiaoyue Ma, Kangyou Yu, Xuanru Meng, Shaoyue Wen. Originally published in JMIR Serious Games (https://games.jmir.org), 22.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Nationwide Deployment of a Serious Game Designed to Improve COVID-19 Infection Prevention Practices in Switzerland: Prospective Web-Based Study

Melanie Suppan[1], MD; Loric Stuby[2], CAS; Stephan Harbarth[3], MD, MSc; Christophe A Fehlmann[4], MD; Sophia Achab[5,6], MD, PhD; Mohamed Abbas[3], MD; Laurent Suppan[4], MD

[1]Division of Anesthesiology, Department of Anesthesiology, Clinical Pharmacology, Intensive Care, and Emergency Medicine, University of Geneva Hospitals and Faculty of Medicine, Geneva, Switzerland

[2]Genève TEAM Ambulances, Geneva, Switzerland

[3]Infection Control Programme and WHO Collaborating Centre on Patient Safety, University of Geneva Hospitals and Faculty of Medicine, Geneva, Switzerland

[4]Division of Emergency Medicine, Department of Anesthesiology, Clinical Pharmacology, Intensive Care, and Emergency Medicine, University of Geneva Hospitals and Faculty of Medicine, Geneva, Switzerland

[5]Specialized Facility in Behavioral Addictions ReConnecte, Geneva University Hospitals, Geneva, Switzerland

[6]WHO Collaborating Center in Training and Research in Mental Health, University of Geneva, Geneva, Switzerland

Corresponding Author:
Laurent Suppan, MD
Division of Emergency Medicine
Department of Anesthesiology, Clinical Pharmacology, Intensive Care, and Emergency Medicine
University of Geneva Hospitals and Faculty of Medicine
Rue Gabrielle-Perret-Gentil 4
Geneva, 1211
Switzerland
Phone: 41 223723311
Email: laurent.suppan@hcuge.ch

Abstract

Background: Lassitude and a rather high degree of mistrust toward the authorities can make regular or overly constraining COVID-19 infection prevention and control campaigns inefficient and even counterproductive. Serious games provide an original, engaging, and potentially effective way of disseminating COVID-19 infection prevention and control guidelines. *Escape COVID-19* is a serious game for teaching COVID-19 infection prevention and control practices that has previously been validated in a population of nursing home personnel.

Objective: We aimed to identify factors learned from playing the serious game Escape COVID-19 that facilitate or impede intentions of changing infection prevention and control behavior in a large and heterogeneous Swiss population.

Methods: This fully automated, prospective web-based study, compliant with the Checklist for Reporting Results of Internet E-Surveys (CHERRIES), was conducted in all 3 main language regions of Switzerland. After creating an account on the platform, participants were asked to complete a short demographic questionnaire before accessing the serious game. The only incentive given to the potential participants was a course completion certificate, which participants obtained after completing the postgame questionnaire. The primary outcome was the proportion of participants who reported that they were willing to change their infection prevention and control behavior. Secondary outcomes were the infection prevention and control areas affected by this willingness and the presumed evolution in the use of specific personal protective equipment items. The elements associated with intention to change infection prevention and control behavior, or lack thereof, were also assessed. Other secondary outcomes were the subjective perceptions regarding length, difficulty, meaningfulness, and usefulness of the serious game; impression of engagement and boredom while playing the serious game; and willingness to recommend its use to friends or colleagues.

Results: From March 9 to June 9, 2021, a total of 3227 accounts were created on the platform, and 1104 participants (34.2%) completed the postgame questionnaire. Of the 1104 respondents, 509 respondents (46.1%) answered that they intended to change their infection prevention and control behavior after playing the game. Among the respondents who answered that they did not intend to change their behavior, 86.1% (512/595) answered that they already apply these guidelines. Participants who followed

XSL•FO
RenderX

the German version were less likely to intend to change their infection prevention and control behavior (odds ratio [OR] 0.48, 95% CI 0.24-0.96; *P*=.04) and found the game less engaging (*P*<.001). Conversely, participants aged 53 years or older had stronger intentions of changing infection prevention and control behavior (OR 2.07, 95% CI 1.44-2.97; *P*<.001).

Conclusions: Escape COVID-19 is a useful tool to enhance correct infection prevention and control measures on a national scale, even after 2 COVID-19 pandemic waves; however, the serious game's impact was affected by language, age category, and previous educational training, and the game should be adapted to enhance its impact on specific populations.

(JMIR Serious Games 2021;9(4):e33003) doi:10.2196/33003

KEYWORDS

COVID-19; serious game; infection prevention; SARS-CoV-2; prospective; web-based; deployment; prevention; gaming; public health; dissemination; health information; behavior; survey

Introduction

Background and Importance

Vaccination campaigns against SARS-CoV-2 have been gathering momentum, but most countries remain months away from reaching group immunity [1-4], provided that such immunity can be attained [5,6]. The almost incessant emergence of new variants is concerning [7]—all the more so as some variants have been shown to evade the immune response acquired by vaccination or by prior infection [8,9]. Viral dissemination may even be facilitated, if normal social interactions are restored and social distancing measures are eliminated, as countries face ever increasing social and economic pressures [10-12]. Moreover, many people experience a rather intense state of tiredness [13-15] coupled with a rather high degree of mistrust toward public health authorities [2,16-18], which can result in a rapidly progressive slackening of infection prevention and control procedures [19-21] and which may even affect health care workers [22]. Switzerland is no exception, and conspiracy theories also thrive in this country [23]. Strengthening infection prevention and control messages and promoting adequate behavior is, therefore, more important than ever; however, regular or aggressive information campaigns might prove counterproductive [21]. These latter strategies were intensively used during the early waves of the pandemic; however, new methods of communication should be considered to increase adherence to infection prevention and control guidelines.

To disseminate COVID-19 prevention messages in an original and engaging way, a transdepartmental and multidisciplinary development team created Escape COVID-19 [24], a serious game based on Nicholson's concept for meaningful gamification [25]. By increasing the players' engagement, serious games, which also enhance knowledge, satisfaction, and skills [26], can help promote adequate behaviors [27,28]. The impact of Escape COVID-19 was assessed in a population of nursing home employees [29]: In a triple-blind controlled trial, the self-reported likelihood of changing infection prevention and control practices was almost 4 times higher in the group of participants who had followed the serious game [29]. However, Escape COVID-19 was only available in French at the time of this prior study, and the population in which it was tested did not allow us to determine whether socioeconomic status or cultural differences influenced its uptake. We hypothesized that receptiveness to messages promoting COVID-19 infection prevention and control behaviors might depend on socioeconomic status and cultural specificities [30].

Objective

The objective of this study was to identify factors learned from playing Escape COVID-19 that favor or impede intentions of changing infection prevention and control behavior in a larger and more heterogeneous population than that used in the previous study [29].

Methods

The Escape COVID-19 Serious Game

The development and features of the Escape COVID-19 serious game have been described previously [24,29]. Briefly, this game was created using Storyline 3 (Articulate Global LLC) and developed using the SERES framework [31], an iterative development approach, to ensure that scientific and design foundations were evidence-based and adapted to the target audience. The different types of players described by Bartle [32] were taken into consideration during the development of Escape COVID-19; therefore, some graphics, game mechanics, and narratives were developed to target achievers, explorers, and socializers.

There is little consensus on what actually defines a serious game; however, according to the conceptual framework created by Tan and Zary [33], Escape COVID-19 fits the definition of a serious game. This framework [33], which describes the criteria required to determine whether a specific material is indeed a serious game, comprises 3 clusters: user experience, play, and learning. Each of these clusters includes 6 base markers, and a minimum of 4 markers per cluster is required to declare that the material under consideration is indeed a serious game. Escape COVID-19 complies with all 18 markers.

Escape COVID-19 contains 4 different levels and takes approximately 15 minutes to complete. Each level represents settings that health care workers typically encounter during a work day (at home, on the road, in communal areas, and in the ward). Players are faced with meaningful choices and are consistently provided with customized, relevant feedback [34]. Adequate infection prevention and control behaviors are rewarded (by an increase in their "thumbs-up" count), while dangerous behaviors lead to an increase in the viral count (Figure 1).

XSL•FO
RenderX

If the virus count reaches a value of 5, a game over screen is displayed (Figure 2). The player can then choose to restart the level or to exchange their thumbs-up count for an equivalent reduction in virus count.

An abridged version, in which the fourth and hardest level (the ward—Figure 3) was not available to players, was also developed (at the request of the Swiss National Science Foundation).

Figure 1. Screenshot showing an example—the player has given a wrong answer, and their viral count rose accordingly.

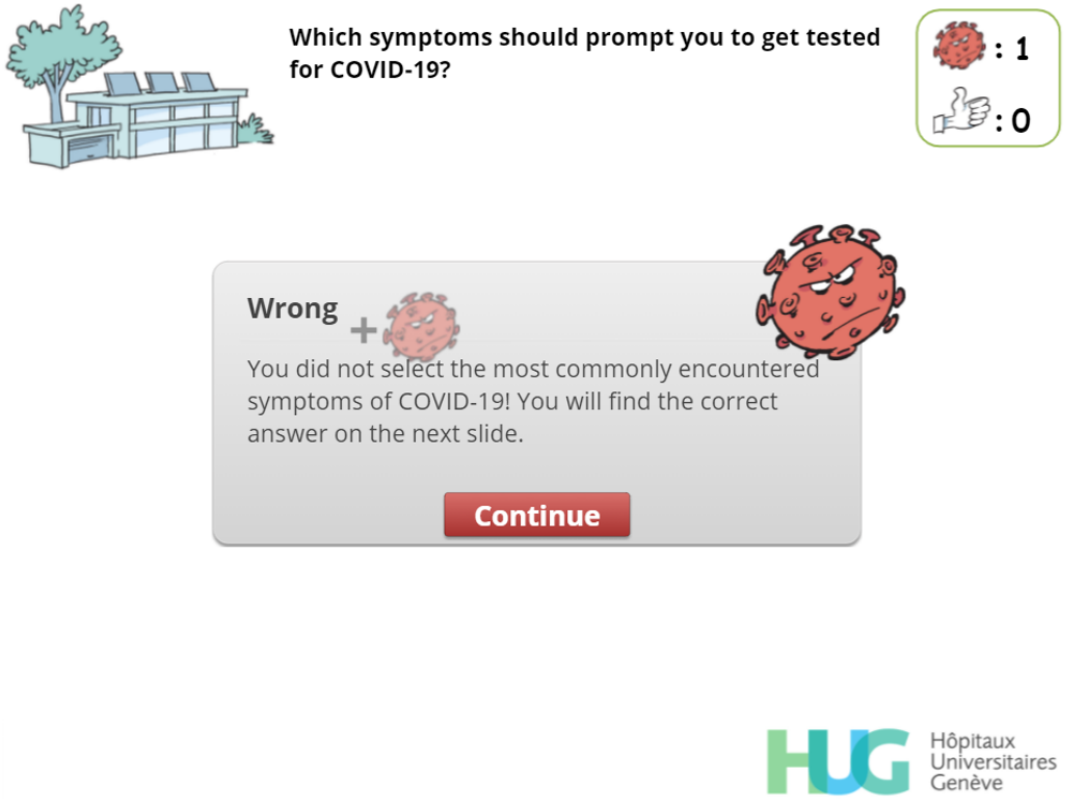
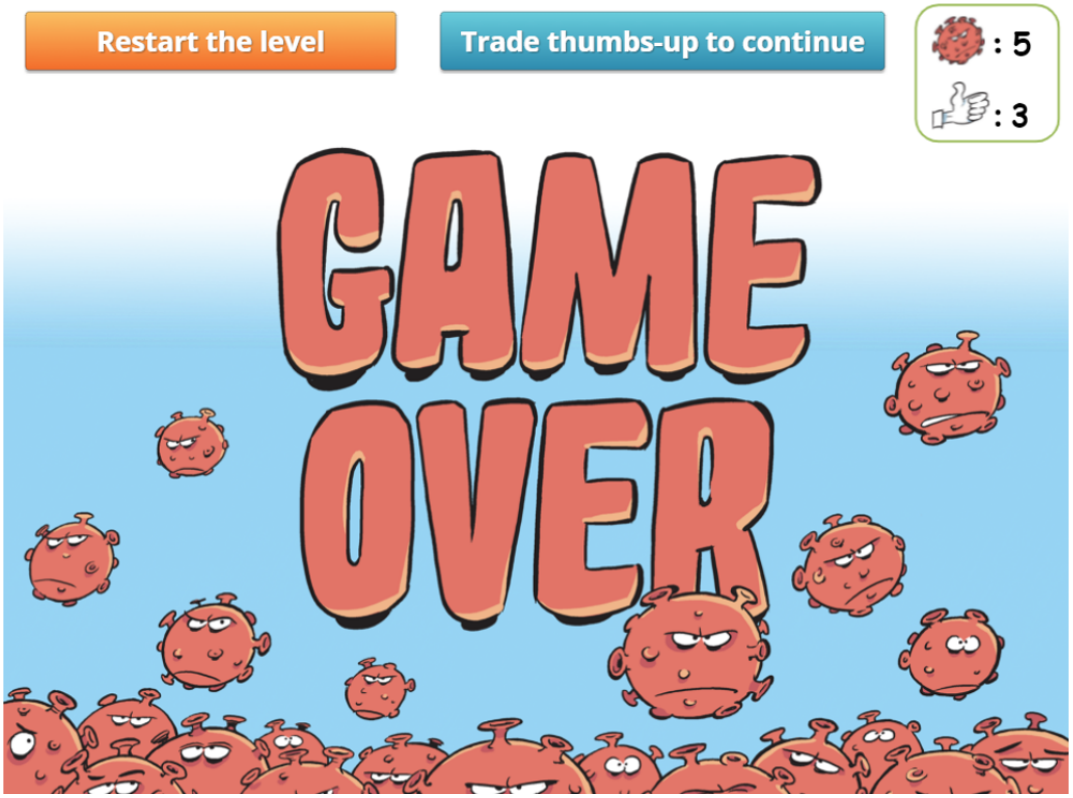
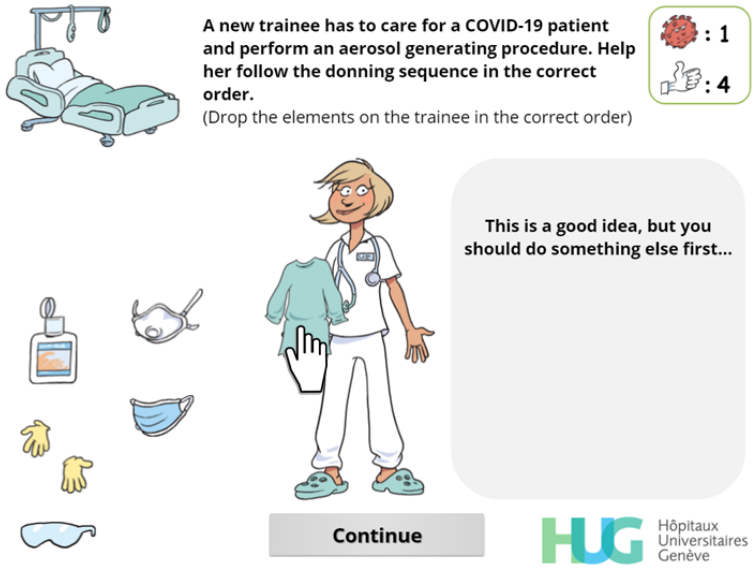


Figure 2. Game over screenshot.



XSL•FO
RenderX

Figure 3. In the donning sequence interaction, the player is asked to drag and drop the appropriate personal protective equipment items in the correct order, and live feedback is displayed in the grey area on the right.



Study Design, Setting, and Participants

We conducted a web-based, fully-automated prospective cohort study in accordance with the Checklist for Reporting Results of Internet E-Surveys (Figure 4) [35]. A declaration of *no objection* was issued by the regional research ethics committee (Req-2021-00600) as this project did not fall within the scope of the Swiss Act on Research involving Human Beings [36]. A disclaimer containing a data policy statement was displayed on the registration form. Data were collected between March and June 2021.

Switzerland is a federal country composed of 26 cantons. There are 4 official languages in Switzerland: German and Swiss German, which are usually grouped as one (62.1%), French (22.8%), Italian (8.0%), and Romansh (0.5%). English is the main language for 5.7% of the permanent resident population [37]. We developed the original version of the Escape COVID-19 serious game in French [24], which we then translated into English. Translations in German and Italian were requested and paid for by the Swiss National Science Foundation (National Research Program 78 framework [38]). A separate translation in Romansh was not requested because Romansh-speaking people in Switzerland generally also speak either German or Italian.

While health care workers were the primary target of this serious game [24], because many of its infection prevention and control messages also apply, we included non–health care workers in this study.

To recruit participants, the Swiss National Science Foundation financed the dissemination of the link to the Escape COVID-19 website [39]. Public and private health care organizations in Switzerland were contacted, and managers were asked to spread the link by different means (such as by email and newsletters). Ten organizations, including the Swiss Red Cross and the Swiss Society for Public Health, decided to endorse the game and allow their logo to be displayed on the website and on the course completion certificate. To enhance participation even further, the Geneva University Hospitals published a press release [40], which was relayed by *Agence France-Presse* [41].

Figure 4. Study design. The European Union flag (on the left at the account creation step) signifies that account and data management were in compliance with the General Data Protection Regulation.

Web Platform and Study Sequence

A multilingual website [39] was developed using a content management system (Joomla!, version 3.9; Open Source Matters Inc). The front page displayed a disclaimer to address concerns regarding potentially contradicting guidelines immediately above the language-specific links used for account creation (Figure 5). Dedicated links were available to allow non–health care workers to create specific accounts.

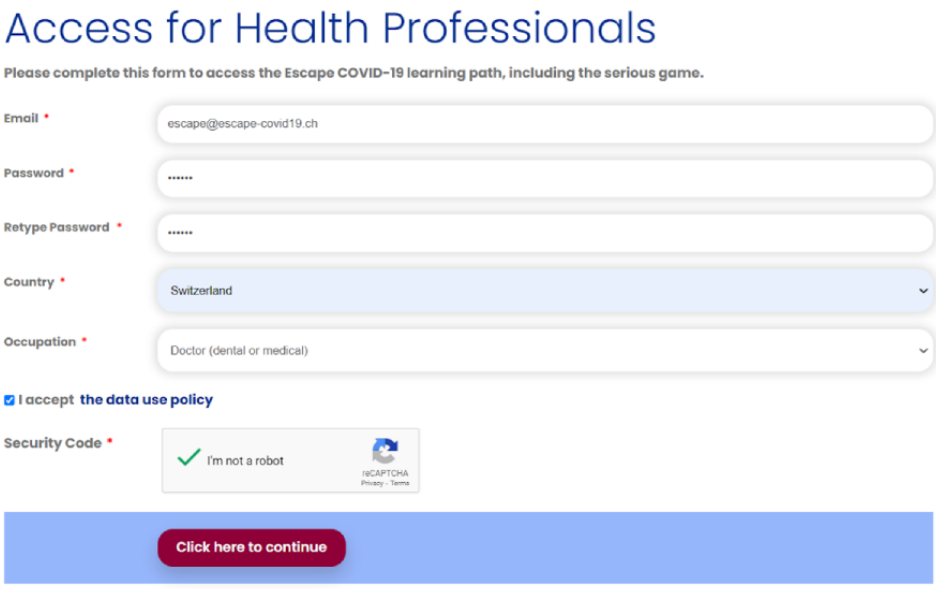
Participants registered using Joom Donation MembershipPro (version 2.1; Joomla Extensions by JoomDonation), which allows an extensive customization of the registration fields and the addition of specific fields. For instance, the occupation field

XSL•FO
RenderX

differed for those who were health care workers and those who were not health care workers. In addition, respondents who were not health care workers were asked to select the highest degree or level of school that they had completed. All fields, including CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) and a checkbox to accept the data use policy, were required before participants were automatically logged in (Figure 6).

Figure 5. Front page of the website.

Figure 6. Registration form example.



Participants were immediately redirected to a demographic questionnaire (gender, age, and COVID status). All questionnaires were created using Community Surveys (version 5.5; Bulasikku Technologies Pvt Ltd), which allows for the use of branching logic. Regular expressions were used to avoid invalid entries. Respondents who were not health care workers were asked whether they wanted to follow the full or abridged version of the Escape COVID-19 serious game.

After completing the demographic questionnaire, participants were able to launch and play Escape COVID-19. After completing the game, a second questionnaire was displayed (Table 1).

XSL•FO
RenderX

Table 1. Second questionnaire.

Questions, response options, and response-dependent questions
After playing this serious game, are you going to change any of your infection prevention practices?
Yes
What areas will these changes affect?[a]
Not going to work if you have symptoms compatible with COVID-19
Protecting yourself from both your colleagues and your patients[b]
The donning sequence[c] when dealing with procedures CARRYING a risk of aerosolization[b]
The donning sequence when dealing with procedures NOT CARRYING a risk of aerosolization[b]
Changing nonsterile gloves more frequently[b]
Practicing hand hygiene more frequently
Disinfecting your workplace
Handling the face mask more carefully
Protecting yourself from asymptomatic people as well as from symptomatic ones
You are now going to use:[d]
Face masks
N95 respirator masks
Eye protection[b]
Nonsterile gloves[b]
Which of these elements greatly contributed to your intention to modify your practices?[e]
The information given in the serious game
The feeling of playing an important role in the common effort against the epidemic
The probability of infecting a relative
One should follow the procedures
Another reason[f]
No
Why will your practices not change?[d]
I already apply all these guidelines
The information given in this serious game does not apply to my situation
The information given in the serious game was not helpful
I do not believe these measures to be useful
I disagree with these measures
Another reason[f]
What could have motivated you to change your practices?[e]
Better understand the reasons behind the recommendations
A greater probability of infecting a relative
The feeling of having an important role in the common effort against the epidemic
Nothing—I could not have been convinced by any argument
Other[f]
Difficulty of the serious game[g]
Please rate the following items from 1 (strongly disagree) to 5 (strongly agree):

XSL•FO
RenderX

Questions, response options, and response-dependent questions
This serious game is engaging
This serious game is meaningful
This serious game is useful
This serious game is boring
I will recommend this serious game to my friends and relatives
I will recommend this serious game to my colleagues
Duration of the serious game[h]
Do you have any further comments to make regarding this serious game?

[a]Response options were in the form of a 5-point Likert scale from 1 (not at all) to 5 (very much).

[b]This option was not shown to non–health care workers who had elected to follow the abridged version of the serious game.

[c]The *donning sequence* refers to the order in which the personal protective equipment items should be put on.

[d]Response options were in the form of a 5-point Likert scale from 1 (much less) to 5 (much more).

[e]Multiple choice question (more than one possible answer).

[f]A free-text field was displayed when this option was selected.

[g]Response options were in the form of a 5-point Likert scale from 1 (too easy) to 5 (too difficult).

[h]Response options were in the form of a 5-point Likert scale from 1 (much too short) to 5 (much too long).

Participants were unable to bypass any step of the study path by virtue of the Access Control List feature in Joomla!. Sourcerer (version 8, Regular Labs) was used to embed the appropriate PHP (Hypertext Preprocessor) functions (JUserHelper::removeUserFromGroup and JUserHelper::addUserToGroup) to ensure that participants were redirected to the appropriate step when resuming the study path. The Conditional Content component (version 3, Regular Labs) was used to display text and links allowing participants to easily resume the study path.

Data Collection

All data were automatically stored on an encrypted MySQL-compatible database (MariaDB, version 5.5.5; MariaDB Corporation Ab) located on a Swiss server. Participants were able to access and delete their accounts and data at any time. Data from all the participants who created an account on the study platform between March 9 and June 9, 2021 were included regardless of their professional status. Participants were allowed to delete their accounts after playing the game, in which case all data associated with their account were automatically removed from the database and could, therefore, be neither retrieved nor included in analysis.

Given the design of this study, there was no predetermined sample size, and no sampling scheme was used. Therefore, the participants who did not delete their accounts represented a convenience sample.

Outcomes

The primary outcome was the proportion of participants reporting that they were willing to change their infection prevention and control behavior (ie, by answering “Yes” to the first question of the second questionnaire). Some secondary outcomes depended on the participants’ willingness to change infection prevention and control behavior: In participants willing to change behavior, secondary outcomes were the infection prevention and control areas affected and the intensity of the participants’ willingness to change. The presumed evolution in the use of specific personal protective equipment items was also assessed, as were the elements motivating this willingness to change behavior. In participants unwilling to change infection prevention and control behavior, secondary outcomes were assessment of the reasons for refusing to change and of the potential motivators which could have induced a willingness to change. Other secondary outcomes such as length, difficulty, and willingness to recommend the serious game to either friends or colleagues were assessed regardless of the participants’ intention to change behavior. Perceived meaningfulness and usefulness of the game and impression of engagement and boredom while playing the serious game were also assessed in the whole sample.

Data Curation and Statistical Analysis

Data were extracted from the database and imported into Stata (version 15; StataCorp LLC) for data curation. Records of participants who did not complete the first questionnaire were excluded. The curated DTA file is available in Multimedia Appendix 1. Descriptive statistics (frequency, relative percentage, mean, and standard deviation) were used when appropriate. As the serious game has already proven its usefulness in health care workers [29], the primary outcome, intention of changing infection prevention and control behavior, was first analyzed using univariate logistic regression according to professional status (health care workers versus non–health care workers). Multivariable logistic regression was then performed to further adjust for gender, COVID status, language, and age. The assumption of log-linearity of age was assessed graphically. As age was not log-linear, it was categorized according to its quartiles. Goodness of fit was checked using the Hosmer-Lemeshow test. Among non–health care workers, univariate and multivariable logistic regression were performed to determine whether language or other specific factors such as gender, age category, occupation, level of education, or playing the abridged version of the serious game rather than the full

XSL•FO
RenderX

game were associated with answering the second questionnaire after completing the serious game. The actual values, from 1 (not at all) to 5 (very much) based on Likert scales, were first used to compute the secondary outcomes. No weighting was applied. For these outcomes, inconsistent answers (ie, those from participants who gave either the minimum or maximum rating for both boredom and engagement) were excluded.

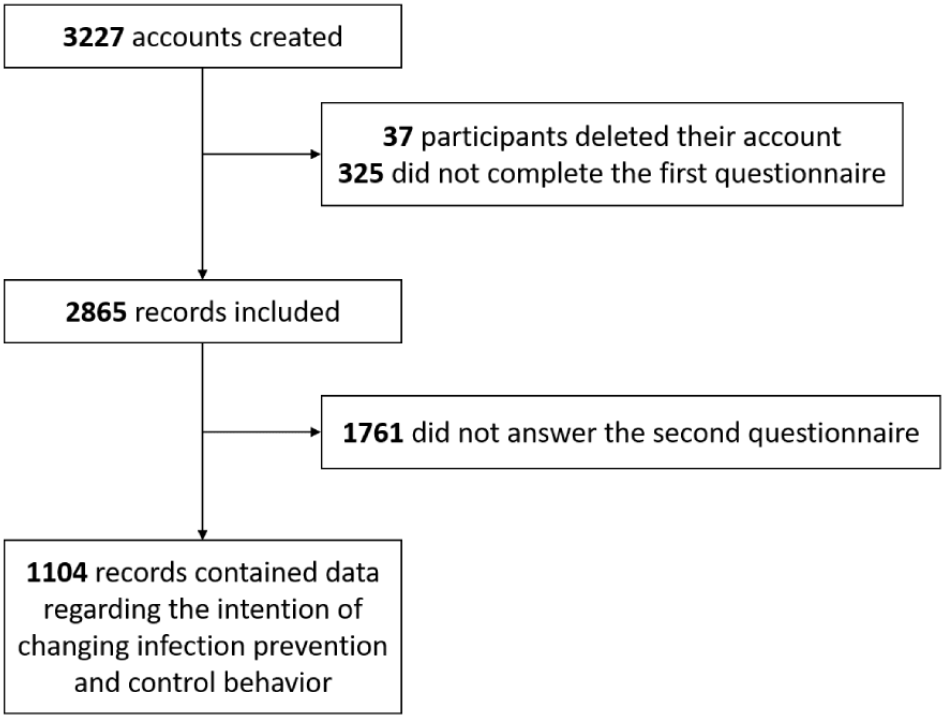
The likelihood ratio test was used to determine whether the parallel lines or proportional odds assumption was met. When met, multivariable ordered logistic regression was used to search for an association between participants' characteristics and secondary outcomes related to the perception of the serious game (engagement and meaningfulness). Because the aforementioned assumption was not met, 2 generalized ordered logistic models were generated using 2 different secondary outcomes as dependent variables: boredom and usefulness. Willingness to change, language, gender, age category, COVID status, occupation, and level of education were used as predictive variables in these models.

Results

From March 9 to June 9, 2021, a total of 3227 accounts were created on the platform: 37 participants chose to delete their accounts, and 325 were excluded as they did not complete the first questionnaire (Figure 7). The characteristics of the 2865 participants included in the analysis are described in Tables 2 and 3.

Most health care workers were nurses (53.0%, 967/1823). Physicians accounted for 11.9% of participants (216/1823), and nursing assistants accounted for 8.1% (148/1823). Other health care professions accounted for 27.0% of the health care worker population.

Figure 7. Study flowchart.



XSL•FO
RenderX

Table 2. Participant characteristics.

Characteristic	Non–health care worker (n=1042)	Health care worker (n=1823)
Language, n (%)[a]		
English	86 (8.3)	83 (4.6)
French	515 (49.4)	646 (35.4)
German	415 (39.8)	1064 (58.4)
Italian	26 (2.5)	30 (1.7)
Gender, n (%)		
Male	395 (37.9)	404 (22.2)
Female	630 (60.5)	1404 (77.0)
Other	17 (1.6)	13 (0.7)
Missing	0 (0)	2 (0.1)
Age, mean (SD)	40.9 (13.8)	41.6 (13.6)[b]
COVID status, n (%)[a]		
Negative or not tested	905 (86.9)	1447 (79.4)
Positive or isolated	3 (0.3)	6 (0.3)
Cured	84 (8.1)	232 (12.7)
Refused to answer	50 (4.8)	136 (7.5)
Missing	0 (0)	2 (0.1)

[a]Percentage totals may exceed 100% due to rounding.
[b]n=2 values were missing.

Table 3. Additional characteristics of non–health care workers.

Characteristic	Non–health care workers, n (%)
Occupation	
Health or social	251 (24.1)
Business	49 (4.7)
Hospitality	20 (1.9)
Manufacturing	35 (3.4)
Public sector or education	252 (24.2)
Transport or retail	28 (2.7)
Other	335 (32.1)
Missing	72 (6.9)
Degree of education	
Mandatory school	49 (4.7)
Secondary education	39 (3.7)
Professional diploma	289 (27.7)
High school graduate	149 (14.3)
University graduate	430 (41.3)
Other	60 (5.8)
Missing	26 (2.5)
Version played	
Abridged	576 (55.3)
Full	466 (44.7)

XSL•FO
RenderX

Of the 2865 participants included in the analysis, 1104 (38.5%) provided information regarding their intentions of changing infection prevention and control behavior, and 1061 fully completed the second questionnaire. Of those, 477 (45.0%) generated a course completion certificate.

Among non–health care workers, the only factor associated with the probability of answering the second questionnaire was having completed the abridged version, which was significant after adjusting for playing the abridged version rather than the full one and for gender, age, occupation, and level of education (odds ratio [OR] 1.96, 95% CI 1.48-2.59). In health care workers, an association between language and the probability of answering the questionnaire was present after univariate analysis (P<.001) and remained after adjusting for gender, age category, profession, and COVID status (P=.002). In this group, participants who refused to detail their COVID status were less likely to answer this questionnaire (OR 0.49, 95% CI 0.32-0.75).

Among participants who answered the second questionnaire, less than half (509/1104, 46.1%) intended to change their infection prevention and control behavior after playing the game. However, among those who did not intend to change their infection prevention and control behavior, the vast majority answered that they already apply these guidelines (Table 4). Few participants (13/595, 2.2%) answered that they disagreed with these measures or that they did not believe such measures to be useful.

Participants intending to change infection prevention and control behavior were generally radical in their willingness to change certain aspects of the practices (not at all: 1003/4085 answers, 24.6%; very much: 1795/4085 answers, 43.9%).

Most of them were not likely to change their use of specific personal protective equipment items, the only exception being related to the use of protective goggles (156/383, 40.7% answered that they would use goggles more or much more after playing the serious game).

Participants who answered that they intended to change their infection prevention and control behavior were mostly motivated by the information contained in the serious game and by the feeling of playing an important role in the common effort against the epidemic (Table 5).

Table 4. Reasons given for not changing infection prevention and control behavior.

Reason	Non–health care worker (n=309), n (%)[a]	Health care worker (n=286), n (%)[a]
Already applies these guidelines	288 (93.2)	224 (78.3)
The information given in the serious game did not apply to the participant's situation	33 (10.7)	78 (27.3)
Disagrees with these measures	3 (1.0)	10 (3.5)
The information given in the serious game was not considered helpful	5 (1.6)	2 (0.7)
Did not believe these measures to be useful	2 (0.6)	2 (0.7)
Another reason	6 (1.9)	19 (6.6)

[a]Multiple responses are possible; therefore, percentages do not add to 100%.

Table 5. Reasons associated with the willingness of changing infection prevention and control behavior.

Reason	Non–health care worker (n=124), n (%)[a]	Health care worker (n=385), n (%)[a]
The information given in the serious game	70 (56.5)	208 (54.0)
The feeling of playing an important role in the common effort against the epidemic	76 (61.3)	200 (51.9)
One should follow the procedures	43 (34.7)	144 (37.4)
The probability of infecting a relative	57 (46.0)	129 (33.5)
Another reason	3 (2.4)	8 (2.1)

[a]Multiple responses are possible; therefore, percentages do not add to 100%.

Univariate logistic regression showed that, compared to non–health care workers, health care workers were more likely to intend to change their infection prevention and control behavior (OR 3.35, 95% CI 2.59-4.34). This association was reinforced after adjusting for age category, gender, language, and COVID status (OR 3.88, 95% CI 2.94-5.13). The 2 elements with significant associations were German language (OR 0.48, 95% CI 0.24-0.96) compared to reference category (English) and being aged 53 years or older (OR 2.07, 95% CI 1.44-2.97) compared to reference category (11-30 years).

Among non–health care workers, belonging to the highest age category (53 years or older: OR 3.33, 95% CI 1.66-6.68) and playing the full rather than the abridged version of the serious game (OR 4.45, 95% CI 2.63-7.55) were associated with a more significant willingness to change infection prevention and control behavior. There was no significant difference in this

XSL•FO
RenderX

outcome related to either occupation (*P*=.42) or educational status (*P*=.11).

Thirty-three records (33/1061, 3.1%) were excluded because they contained either minimal or maximal ratings for both the engagement and boredom items. The serious game was generally considered engaging (mean 3.9, SD 1.0), meaningful (mean 4.4, SD 0.9), and useful (mean 4.3, SD 0.9). Boredom was rated rather low (mean 1.9, SD 1.1).

Perceived meaningfulness was affected by 4 different factors. Non–health care workers who intended to change their infection prevention and control behavior were more likely to find the game meaningful (OR 2.40, 95% CI 1.44-3.99). Those who had chosen to play the full version of the game were also more likely to find the game meaningful (OR 1.74, 95% CI 1.05-2.88). Participants who did not identify themselves as belonging to the male or female gender found the game to be less meaningful (OR 0.14, 95% CI 0.03-0.68).

Likewise, non–health care workers who intended to change their infection prevention and control behavior, and those who had chosen to play the full version of the game were more likely to find it engaging (OR 2.73, 95% CI 1.71-4.33 and OR 2.24 95% CI 1.39-3.60, respectively). Non–health care workers who identified themselves as belonging to the female gender also found the game more engaging (OR 1.58, 95% CI 1.02-2.45), while participants who did not identify themselves as either male or female found it less engaging (OR 0.19, 95% CI 0.05-0.76). Participants who played the German version of the game also found it less engaging (OR 0.15, 95% CI 0.06-0.40).

Health care workers were more likely than non–health care workers to recommend this serious game to their colleagues (OR 1.87, 95% CI 1.49-2.35). They were also more likely to recommend it to their relatives than non–health care workers were (OR 1.43, 95% CI 1.14-1.79). Regarding difficulty, Escape COVID-19 was considered to be well balanced by 71.3% (756/1061), and more health care workers (534/655, 81.5%) to found it to be well balanced (*P*<.001), with 13.6% (89/655) finding it either easy or too easy, versus 54.7% of non–health care workers (222/406), with 39.4% (160/406) finding it either easy or too easy. Regarding length, the serious game was considered as well balanced by 89.4% of participants (health care workers: 597/655, 91.1%; non–health care workers: 351/406, 86.5%; *P*=.002).

Given the differences observed in the uptake of the French and German versions, we carried out a posthoc qualitative analysis of the final comments recorded by participants to determine whether translation problems could be involved. There were 158 comments, 81 of which (51.3%) were recorded by participants who played the German version of Escape COVID-19. Of these, 8 (9.9%) were linked with translation issues. Two examples can be found below:

> *The game is very good, so please refrain from using unnecessary anglicisms that are hard to understand for anyone who is not a native English speaker...*
>
> *Please pay attention to correct gendering. The job title "Krankenschwester" is wrong...*

All original comments are included in the curated data file (Multimedia Appendix 1).

Discussion

Principal Results

This nationwide campaign based on an innovative educational tool shows that giving players the impression of having an important role in the fight against the pandemic is a potent motivator for enhancing intentions of changing infection prevention and control behavior. Future serious games could use this element to promote vaccination.

Slightly less than half (509/1104, 46.1%) of participants who provided information on their intentions responded that they intended to change their infection prevention and control behavior after playing Escape COVID-19. While this figure might seem rather low at first glance, the vast majority of those not willing to change behavior answered that they already apply the guidelines outlined in the serious game. Even more reassuringly, the proportion of participants who did not believe infection prevention and control measures to be useful was extremely low. These findings must, however, be mitigated by a probably important selection bias, which is further detailed below.

The only element associated with a lesser probability of answering the second questionnaire for non–health care worker participants was choosing to play the full rather than the abridged version of the serious game. Two hypotheses might explain this finding. First, the full version of the game is longer and is much more difficult because it includes the fourth level (the ward), and non–health care workers are neither used to this setting nor to the specific personal protective equipment items and procedures depicted there. Some participants might, therefore, have chosen to abandon the game without completing this level. As the serious game was not embedded in a learning management system, we were unable to define the exact moment at which participants chose to drop out, and we were unable to gather information related to specific questions or to record the final score obtained by participants. Second, some of the participants who completed this full version of the game might have felt ill at ease completing the second questionnaire and might have thought that only health care workers were actually qualified to answer it. Nevertheless, findings from a recent study [42] strengthen support for the importance of including non–health care workers in the target population. Indeed, a secondary analysis of quantitative data collected in Ghana showed that nonclinical staff, midwives and pharmacists demonstrated lower adherence to infection prevention and control guidelines than that demonstrated by other health care workers [42]. Thus, designing infection prevention and control promotion materials that can reach many different professions, including non–health care workers in regular contact with vulnerable populations, is essential to limit viral transmission.

Older participants reported that they were more likely to change their infection prevention and control behavior after following the serious game. Three theories could explain this association. First, older age is associated with an increased risk of

XSL•FO
RenderX

complications after SARS-CoV-2 infection [43], and older individuals might therefore be wary of such complications. Second, regular information material such as flyers, posters, or other text documents might be too dull and too hard to grasp [44]. A more engaging and unambiguous way of providing information might, therefore, be more appropriate to convey critical messages to this population. Indeed, even though there might be a tendency to think that games are only for young individuals, factual data bely this somewhat prejudiced impression [45-48]. Third, an effect that is linked to different digital generations (ie, X, Y and Z) could also be present. Indeed, entering an e-learning process has been found to be linked to the strongest level of contextualization of learned web-based content in generation X (individuals born between 1965 and 1979) [49].

As a federal and multilinguistic country, Switzerland represents a unique challenge. In this nationwide study, the probability of gathering health care worker data regarding the intention of changing infection prevention and control behavior and concerning perceptions about the serious game depended on language, gender, and COVID status. This last fact is hardly surprising, as it can be hypothesized that participants unwilling to detail their COVID status before playing the game might already have been skeptical regarding its content. Indeed, it does not seem unrealistic to assume that these participants did not trust the investigators and the data protection policy displayed on the website or that they were simply unwilling to provide any useful data. Though not conclusive, an element supporting this hypothesis was found in a randomized controlled trial [50] that assessed the effect of an e-learning module on attitude and knowledge regarding personal protective equipment, in which an inconsistent answer set had to be excluded and was the only one in which the respondent answered that they were unwilling to disclose their COVID status.

This study was carried out after the peak of the second pandemic wave had been reached in Switzerland, and most health care workers had already accessed a wide array of infection prevention and control educational tools, thereby substantially decreasing the number of potentially naïve participants. While gender has already been shown to influence engagement in video game activities [51], the impact of language on messages conveyed by serious game has scarcely been studied and deserves attention. Indeed, apart from the lesser probability of answering the second questionnaire, playing the game in a specific language was associated with a different probability of changing infection prevention and control behavior. One aspect of this issue is that the German and Italian translations were performed by third parties who did not take into account the multidisciplinary aspect which had driven the creation of the serious game [24]. Taking learners' profiles into account is, however, essential [52] and might have enhanced message uptake, and one participant commented on the fact that an outdated word was used to identify nurses (ie, *Krankenschwester* rather than *Pflegefachperson*).

Many studies [53,54] have already explored the adherence of particular subsets of health care workers to specific COVID-19 transmission-reducing behaviors. Recently, a web-based survey carried out in Saudi Arabia showed that respiratory therapists were more likely to adhere to infection prevention and control guidelines at home than at work [53]. Thus, behavioral change should be promoted not only in the professional setting but also in the private sphere, and motivation appears to be a key factor to promote such change. Strengthening this message even further, a systematic review [54] assessed the impact of interventions aimed at promoting eHealth capability or motivation among health care professionals, reported that most interventions aimed to promote capability rather than increasing motivation, and concluded that evidence-based developments should be carried out to enhance this last factor. Escape COVID-19, which was generated by virtue of a theory-driven process, seems to provide adequate messages, if our results are to be trusted. In addition, this serious game also helps health care workers reflect on their infection prevention and control behavior outside of the hospital environment through its first 2 levels (at home and on the road).

Cultural differences and perceptions regarding serious games, their graphics, player and nonplayer characters, and the game environment might also affect the impact of serious games. In this study, the participants who followed the German version of Escape COVID-19 were less likely to find it engaging. The reasons underlying this difference could not be thoroughly assessed, and further data should be gathered to determine how this version should be modified to increase engagement. This is critical, as the probability of actually carrying out an action relies upon this parameter [55]. In any case, including native speakers and ensuring the quality and validity of the translation is essential.

In addition, cultural and regional differences can alter whether a specific material should be considered a game, and the boundaries between serious games and gamification are often unclear [56]. However, using McCallum's [57] definition, there is little doubt that Escape COVID-19 is indeed a serious game as it is the result of the “creation of a whole new experience to achieve some change in the player.”

Obtaining a course completion certificate might have convinced some participants to complete the second questionnaire, even though more than half of those who completed it chose not to generate a certificate. Many reasons could explain this fact. First, in Switzerland, some health care workers are not required to obtain continuing medical education credits. This decreases the potential impact of the incentive of obtaining a course completion certificate. Second, official federal organizations are less used to issuing continuing medical education credits for web-based courses, particularly serious games. This prevented us from granting official continuing medical education credits to participants, which further decreased the value of this particular incentive.

Limitations

The main limitation of this study was the use of convenience sampling, which led to a 2-tiered selection bias. First, it is to be expected that the participants who decided to register to play the game were already interested in COVID-19 protective procedures and were probably motivated to enhance their infection prevention and control behavior. Second, even after registering and playing the game, the only incentive to complete

XSL•FO
RenderX

the second questionnaire was the possibility of acquiring a course completion certificate. This incentive was rather limited, as this certificate did not grant any continuing medical education credits, and many Swiss health care workers are not required to attend continuous education. In addition, this study took place after the peak of the second pandemic wave in Switzerland, which potentially decreased the interest some might have had in playing such a game. Furthermore, we were not able to determine the reasons that prevented a rather high proportion of participants from completing the second questionnaire. Attrition is to be expected in such studies [58,59], and we, therefore, strove to keep all questionnaires as short and straightforward as possible [60]. Despite our efforts, the second questionnaire might, nevertheless, have been considered too long by participants who dropped out. It is also possible that some participants who dropped out found the game ill-suited to their particular situation, unnecessary, or even boring. The high level of satisfaction among respondents is, however, relatively reassuring in this matter.

As an abridged version of the game had not been considered during the initial development process and because it was urgently requested, considerations specific to non–health care workers were neither sought nor taken into account. Therefore, even though infection prevention and control procedures are simpler and easier to follow for non–health care workers, this lack of customization might explain, at least in part, the lower willingness to change infection prevention and control behavior observed in non-health care workers. Rather than modifying Escape COVID-19 even further, the use of other serious games, such as "COVID-19-Did You Know?," which was designed to target non–health care workers from inception, should be considered [61].

Perspectives

The cultural differences in message uptake identified in this study are intriguing. Even though such differences have already been identified [62], the extent of these differences was quite surprising, and their causes should now be studied in order to be able to adapt future serious games to these cultural characteristics. While the original, French version of Escape COVID-19 was developed taking feedback from potential end users into account [24], the translated versions were not. Therefore, in line with the co.LAB framework, which recommends identifying learners' profiles and adapting the game accordingly [52], we are considering carrying out focus groups to determine the shortcomings of the current version of this game and the elements that could make Escape COVID-19 more engaging. Different focus groups would be required to take into consideration elements specific to particular social, cultural, and linguistic backgrounds. In addition, such focus groups could be used to determine whether a new level or a new module could help promote vaccination. The main issues such a development must address should be explored during these sessions, which might also shed light on cultural differences regarding vaccine hesitancy [1,2,63,64].

Conclusion

Despite the incentive of obtaining a course completion certificate, the postgame questionnaire was completed by less than half of all participants. Even though more than half of those who filled in this questionnaire reported that they were not going to change their infection prevention and control behavior, they attributed this answer to the fact that they already apply the infection prevention and control guidelines presented in the serious game. The impact of Escape COVID-19 on participants who intended to change their infection prevention and control behavior after playing was, however, substantial as these users were generally radical in their willingness to alter their practice. They reported that their willingness to change their infection prevention and control behavior was equally motivated by the feeling of playing an important role in the common effort against the pandemic and by the information contained in the serious game.

In this study, older age was associated with the intention of changing infection prevention and control behavior. Conversely, playing the German version of the game decreased the likelihood of intending to change behavior. Different hypotheses could explain these findings and should now be explored to help adapt the game and enhance the uptake of infection prevention and control messages.

Acknowledgments

First and foremost, we would like to thank Patrick Sauter from EK Biosciences GmbH who has tirelessly worked to promote, support, and disseminate this serious game. We would also like to extend their warmest thanks to Markus Ehrat and Pascal Walther for their unfailing support throughout the course of this project.

Many thanks to Gaud Catho, Monica Perez, Valérie Sauvan, and Tomás Robalo Nunes, from the Infection Prevention and Control division at Geneva University Hospitals, who provided essential input for developing the content of Escape COVID-19, and most particularly to Mr. Eric Buche who designed the graphics of the serious game. The German and Italian translations of the Escape COVID-19 serious game, as well as its promotion and dissemination among Swiss health care institutions, were financed by the Swiss National Science Foundation by means of a knowledge transfer project linked to National Research Programme 78.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Curated data file in CSV and DTA formats.
[ZIP File (Zip Archive), 248 KB - games_v9i4e33003_app1.zip]

References

1. Griffith J, Marani H, Monkman H. COVID-19 vaccine hesitancy in Canada: content analysis of tweets using the theoretical domains framework. J Med Internet Res 2021 Apr 13;23(4):e26874 [FREE Full text] [doi: 10.2196/26874] [Medline: 33769946]
2. Hou Z, Tong Y, Du F, Lu L, Zhao S, Yu K, et al. Assessing COVID-19 vaccine hesitancy, confidence, and public engagement: a global social listening study. J Med Internet Res 2021 Jun 11;23(6):e27632 [FREE Full text] [doi: 10.2196/27632] [Medline: 34061757]
3. Li J, Giabbanelli P. Returning to a normal life via COVID-19 vaccines in the United States: a large-scale agent-based simulation study. JMIR Med Inform 2021 Apr 29;9(4):e27419. [doi: 10.2196/27419] [Medline: 33872188]
4. Clemente-Suárez VJ, Hormeño-Holgado A, Jiménez M, Benitez-Agudelo JC, Navarro-Jiménez E, Perez-Palencia N, et al. Dynamics of population immunity due to the herd effect in the COVID-19 pandemic. Vaccines 2020 May 19;8(2):236. [doi: 10.3390/vaccines8020236] [Medline: 32438622]
5. Aschwanden C. Five reasons why COVID herd immunity is probably impossible. Nature 2021 Mar 18;591(7851):520-522. [doi: 10.1038/d41586-021-00728-2] [Medline: 33737753]
6. Kwok SWH, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: machine learning analysis. J Med Internet Res 2021 May 19;23(5):e26953. [doi: 10.2196/26953] [Medline: 33886492]
7. Chrysostomou AC, Vrancken B, Koumbaris G, Themistokleous G, Aristokleous A, Masia C, et al. A comprehensive molecular epidemiological analysis of SARS-CoV-2 infection in Cyprus from April 2020 to January 2021: evidence of a highly polyphyletic and evolving epidemic. Viruses 2021 Jun 09;13(6):1098. [doi: 10.3390/v13061098] [Medline: 34207490]
8. Lazarevic I, Pravica V, Miljanovic D, Cupic M. Immune evasion of SARS-CoV-2 emerging variants: what have we learnt so far? Viruses 2021 Jun 22;13(7):1192. [doi: 10.3390/v13071192] [Medline: 34206453]
9. McCallum M, Bassi J, De Marco A, Chen A, Walls AC, Di Iulio J, et al. SARS-CoV-2 immune evasion by the B.1.427/B.1.429 variant of concern. Science 2021 Jul 01;373(6555):648-654. [doi: 10.1126/science.abi7994] [Medline: 34210893]
10. Garnier R, Benetka JR, Kraemer J, Bansal S. Socioeconomic disparities in social distancing during the COVID-19 pandemic in the United States: observational study. J Med Internet Res 2021 Jan 22;23(1):e24591. [doi: 10.2196/24591] [Medline: 33351774]
11. Krumer-Nevo M, Refaeli T. COVID-19: a poverty-aware perspective. Am J Orthopsychiatry 2021;91(3):423-431. [doi: 10.1037/ort0000566] [Medline: 34138631]
12. Kavanagh NM, Goel RR, Venkataramani AS. County-level socioeconomic and political predictors of distancing for COVID-19. Am J Prev Med 2021 Jul;61(1):13-19 [FREE Full text] [doi: 10.1016/j.amepre.2021.01.040] [Medline: 33947527]
13. Houghton C, Meskell P, Delaney H, Smalle M, Glenton C, Booth A, et al. Barriers and facilitators to healthcare workers' adherence with infection prevention and control (IPC) guidelines for respiratory infectious diseases: a rapid qualitative evidence synthesis. Cochrane Database Syst Rev 2020 Apr 21;4:CD013582. [doi: 10.1002/14651858.CD013582] [Medline: 32315451]
14. Zou X, Liu S, Li J, Chen W, Ye J, Yang Y, et al. Factors associated with healthcare workers' insomnia symptoms and fatigue in the fight against COVID-19, and the role of organizational support. Front Psychiatry 2021;12:652717 [FREE Full text] [doi: 10.3389/fpsyt.2021.652717] [Medline: 33841214]
15. Matsuo T, Taki F, Kobayashi D, Jinta T, Suzuki C, Ayabe A, et al. Health care worker burnout after the first wave of the coronavirus disease 2019 (COVID-19) pandemic in Japan. J Occup Health 2021 Jan;63(1):e12247 [FREE Full text] [doi: 10.1002/1348-9585.12247] [Medline: 34375497]
16. Hardy L, Mana A, Mundell L, Benheim S, Morales K, Sagy S. Living in opposition: how women in the United States cope in spite of mistrust of federal leadership during the pandemic of Covid-19. J Community Psychol 2021 Aug;49(6):2059-2070 [FREE Full text] [doi: 10.1002/jcop.22544] [Medline: 33729586]
17. Freeman D, Waite F, Rosebrock L, Petit A, Causier C, East A, et al. Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. Psychol Med 2020 May 21:1-13. [doi: 10.1017/s0033291720001890] [Medline: 32436485]
18. Cheung ATM, Parent B. Mistrust and inconsistency during COVID-19: considerations for resource allocation guidelines that prioritise healthcare workers. J Med Ethics 2021 Feb;47(2):73-77. [doi: 10.1136/medethics-2020-106801] [Medline: 33106381]
19. Varella MAC, Luoto S, Soares RBDS, Valentova JV. COVID-19 pandemic on fire: evolved propensities for nocturnal activities as a liability against epidemiological control. Front Psychol 2021;12:646711 [FREE Full text] [doi: 10.3389/fpsyg.2021.646711] [Medline: 33828510]
20. MacIntyre CR, Nguyen P, Chughtai AA, Trent M, Gerber B, Steinhofel K, et al. Mask use, risk-mitigation behaviours and pandemic fatigue during the COVID-19 pandemic in five cities in Australia, the UK and USA: a cross-sectional survey. Int J Infect Dis 2021 May;106:199-207 [FREE Full text] [doi: 10.1016/j.ijid.2021.03.056] [Medline: 33771668]

XSL•FO
RenderX

21. Ball H, Wozniak TR. Why do some Americans resist COVID-19 prevention behavior? an analysis of issue importance, message fatigue, and reactance regarding COVID-19 messaging. Health Commun 2021 May 03:1-8. [doi: 10.1080/10410236.2021.1920717] [Medline: 33941005]
22. Etafa W, Gadisa G, Jabessa S, Takele T. Healthcare workers' compliance and its potential determinants to prevent COVID-19 in public hospitals in Western Ethiopia. BMC Infect Dis 2021 May 19;21(1):454 [FREE Full text] [doi: 10.1186/s12879-021-06149-w] [Medline: 34011263]
23. Lugon L, Revello S. Les complotistes s'enflamment autour du coronavirus. Le Temps. 2020 May 11. URL: https://www.letemps.ch/suisse/complotistes-senflamment-autour-coronavirus [accessed 2021-09-13]
24. Suppan M, Catho G, Robalo Nunes T, Sauvan V, Perez M, Graf C, et al. A serious game designed to promote safe behaviors among health care workers during the COVID-19 pandemic: development of "Escape COVID-19". JMIR Serious Games 2020 Dec 03;8(4):e24986 [FREE Full text] [doi: 10.2196/24986] [Medline: 33242312]
25. Nicholson S. A recipe for meaningful gamification. In: Reiners T, Wood L, editors. Gamification in Education and Business. Cham: Springer; 2015.
26. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: systematic review. J Med Internet Res 2019 Mar 28;21(3):e12994 [FREE Full text] [doi: 10.2196/12994] [Medline: 30920375]
27. Esmaeilzadeh P. The Influence of Gamification and Information Technology Identity on Postadoption Behaviors of Health and Fitness App Users: Empirical Study in the United States. JMIR Serious Games 2021 Jul 5;9(3):e28282 [FREE Full text] [doi: 10.2196/28282]
28. Epstein DS, Zemski A, Enticott J, Barton C. Tabletop board game elements and gamification interventions for health behavior change: realist review and proposal of a game design framework. JMIR Serious Games 2021 Mar 31;9(1):e23302 [FREE Full text] [doi: 10.2196/23302] [Medline: 33787502]
29. Suppan L, Abbas M, Catho G, Stuby L, Regard S, Achab S, et al. Impact of a Serious game (Escape COVID-19) on the intention to change COVID-19 control practices among employees of long-term care facilities: web-based randomized controlled trial. J Med Internet Res 2021 Mar 25;23(3):e27443 [FREE Full text] [doi: 10.2196/27443] [Medline: 33685854]
30. Hernandez M, Gibb J. Culture, behavior and health. Evol Med Public Health 2020;2020(1):12-13 [FREE Full text] [doi: 10.1093/emph/eoz036] [Medline: 31976074]
31. Verschueren S, Buffel C, Vander Stichele G. Developing theory-driven, evidence-based serious games for health: framework based on research community insights. JMIR Serious Games 2019 May 02;7(2):e11565 [FREE Full text] [doi: 10.2196/11565] [Medline: 31045496]
32. Bartle R. Hearts, clubs, diamonds, spades: players who suit MUDs. Richard A Bartle website. 1996. URL: https://mud.co.uk/richard/hcds.htm [accessed 2021-09-13]
33. Tan JW, Zary N. Diagnostic markers of user experience, play, and learning for digital serious games: a conceptual framework study. JMIR Serious Games 2019 Jul 16;7(3):e14620 [FREE Full text] [doi: 10.2196/14620] [Medline: 31313660]
34. Kopp V, Stark R, Fischer M. Fostering diagnostic knowledge through computer-supported, case-based worked examples: effects of erroneous examples and feedback. Med Educ 2008 Aug;42(8):823-829. [doi: 10.1111/j.1365-2923.2008.03122.x] [Medline: 18564096]
35. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
36. Swiss Confederation. Federal act on research involving human beings. Fedlex. 2011 Sep 30. URL: https://www.admin.ch/opc/en/classified-compilation/20061313/index.html [accessed 2020-05-12]
37. Languages. Federal Statistical Office. 2019. URL: https://www.bfs.admin.ch/bfs/en/home/statistics/population/languages-religions/languages.html [accessed 2021-09-23]
38. National Research Programme. Swiss National Science Foundation. URL: https://www.nfp78.ch/en/ [accessed 2021-11-22]
39. Suppan M, Catho G, Perez M, Sauvan V, Robalo Nunes T, Abbas M, et al. Escape COVID-19. URL: https://escape-covid19.ch [accessed 2021-09-13]
40. Hôpitaux Universitaires de Genève. COVID-19: Le « jeu sérieux » plus efficace pour prévenier le COVID-19 à l'hôpital. 2021 May 03. URL: https://www.hug.ch/medias/communique-presse/covid-19-jeu-serieux-plus-efficace-pour-prevenir-covid-19-lhopital [accessed 2021-09-13]
41. Euronews. Health workers in Switzerland are using videogames to beat COVID-19. 2021 May 13. URL: https://www.euronews.com/next/2021/05/13/health-workers-in-switzerland-are-using-videogames-to-beat-covid-19 [accessed 2021-09-13]
42. Ashinyo ME, Dubik SD, Duti V, Amegah KE, Ashinyo A, Asare BA, et al. Infection prevention and control compliance among exposed healthcare workers in COVID-19 treatment centers in Ghana: a descriptive cross-sectional study. PLoS One 2021 Mar 9;16(3):e0248282 [FREE Full text] [doi: 10.1371/journal.pone.0248282] [Medline: 33690699]
43. Levin A, Hanage W, Owusu-Boaitey N, Cochran K, Walsh S, Meyerowitz-Katz G. National Bureau of Economic Research. 2020 Jul. URL: https://www.nber.org/system/files/working_papers/w27597/w27597.pdf [accessed 2021-11-22]
44. Meyer BJF, Marsiske M, Willis SL. Text processing variables predict the readability of everyday documents read by older adults. Read Res Q 1993 Jul 01;28(3):235 [FREE Full text] [Medline: 20169005]

XSL•FO
RenderX

45. Marston HR. Older adults as 21st century game designers. Comput Game J 2012 May 15;1(1):90-102. [doi: 10.1007/bf03392330]
46. Chu CH, Biss RK, Cooper L, Quan AML, Matulis H. Exergaming platform for older adults residing in long-term care homes: user-centered design, development, and usability study. JMIR Serious Games 2021 Mar 09;9(1):e22370 [FREE Full text] [doi: 10.2196/22370] [Medline: 33687337]
47. 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]
48. Misconceptions—Video Game Dissection. Sites at Penn State. URL: https://sites.psu.edu/videogamebreakdown/stereotypes/ [accessed 2021-09-13]
49. Yawson DE, Yamoah FA. Understanding satisfaction essentials of e-learning in higher education: a multi-generational cohort perspective. Heliyon 2020 Nov;6(11):e05519 [FREE Full text] [doi: 10.1016/j.heliyon.2020.e05519] [Medline: 33251368]
50. Suppan L, Abbas M, Stuby L, Cottet P, Larribau R, Golay E, et al. Effect of an e-learning module on personal protective equipment proficiency among prehospital personnel: web-based randomized controlled trial. J Med Internet Res 2020 Aug 21;22(8):e21265 [FREE Full text] [doi: 10.2196/21265] [Medline: 32747329]
51. Ogletree SM, Drake R. College students' video game participation and perceptions: gender differences and implications. Sex Roles 2007 Mar 21;56(7-8):537-542. [doi: 10.1007/s11199-007-9193-5]
52. Jaccard D, Suppan L, Sanchez E, Huguenin A, Laurent M. The co.LAB generic framework for collaborative design of serious games: development study. JMIR Serious Games 2021 Jul 02;9(3):e28674 [FREE Full text] [doi: 10.2196/28674] [Medline: 34255655]
53. Al Khathlan N, Padhi BK. Adherence to COVID-19 appropriate behavior among respiratory therapists: a cross-sectional study in the Kingdom of Saudi Arabia. Front Public Health 2021 Aug 18;9:715982 [FREE Full text] [doi: 10.3389/fpubh.2021.715982] [Medline: 34490196]
54. Virtanen L, Kaihlanen A, Laukka E, Gluschkoff K, Heponiemi T. Behavior change techniques to promote healthcare professionals' eHealth competency: a systematic review of interventions. Int J Med Inform 2021 May;149:104432 [FREE Full text] [doi: 10.1016/j.ijmedinf.2021.104432] [Medline: 33684712]
55. Ajzen I. Action Control. In: Kuhl J, Beckmann J, editors. From Intentions to Actions: A Theory of Planned Behavior. Berlin: Springer; 1985:11-39.
56. Warsinsky S, Schmidt-Kraepelin M, Rank S, Thiebes S, Sunyaev A. Conceptual Ambiguity surrounding gamification and serious games in health care: literature review and development of game-based intervention reporting guidelines (GAMING). J Med Internet Res 2021 Sep 10;23(9):e30390 [FREE Full text] [doi: 10.2196/30390] [Medline: 34505840]
57. McCallum S. Gamification and serious games for personalized health. Stud Health Technol Inform 2012;177:85-96. [Medline: 22942036]
58. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
59. Hochheimer CJ, Sabo RT, Perera RA, Mukhopadhyay N, Krist AH. Identifying attrition phases in survey data: applicability and assessment study. J Med Internet Res 2019 Aug 23;21(8):e12811 [FREE Full text] [doi: 10.2196/12811] [Medline: 31444875]
60. McCambridge J, Kalaitzaki E, White IR, Khadjesari Z, Murray E, Linke S, et al. Impact of length or relevance of questionnaires on attrition in online trials: randomized controlled trial. J Med Internet Res 2011 Nov 18;13(4):e96 [FREE Full text] [doi: 10.2196/jmir.1733] [Medline: 22100793]
61. Gaspar JDS, Lage EM, Silva FJD, Mineiro ?, Oliveira IJRD, Oliveira I, et al. A mobile serious game about the pandemic (COVID-19 - did you know?): design and evaluation study. JMIR Serious Games 2020 Dec 22;8(4):e25226 [FREE Full text] [doi: 10.2196/25226] [Medline: 33301416]
62. Harbarth S, Albrich W, Brun-Buisson C. Outpatient antibiotic use and prevalence of antibiotic-resistant pneumococci in France and Germany: a sociocultural perspective. Emerg Infect Dis 2002 Dec;8(12):1460-1467 [FREE Full text] [doi: 10.3201/eid0812.010533] [Medline: 12498664]
63. Dror AA, Eisenbach N, Taiber S, Morozov NG, Mizrachi M, Zigron A, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. Eur J Epidemiol 2020 Aug;35(8):775-779. [doi: 10.1007/s10654-020-00671-y] [Medline: 32785815]
64. Wang J, Jing R, Lai X, Zhang H, Lyu Y, Knoll MD, et al. Acceptance of COVID-19 vaccination during the COVID-19 pandemic in China. Vaccines (Basel) 2020 Aug 27;8(3) [FREE Full text] [doi: 10.3390/vaccines8030482] [Medline: 32867224]

XSL•FO
RenderX

Edited by N Zary; submitted 18.08.21; peer-reviewed by P Giabbanelli, A White, M Szeto, J Anderson; comments to author 09.09.21; revised version received 18.09.21; accepted 12.10.21; published 25.11.21.

Please cite as:
Suppan M, Stuby L, Harbarth S, Fehlmann CA, Achab S, Abbas M, Suppan L
Nationwide Deployment of a Serious Game Designed to Improve COVID-19 Infection Prevention Practices in Switzerland: Prospective Web-Based Study
JMIR Serious Games 2021;9(4):e33003
URL: https://games.jmir.org/2021/4/e33003
doi:10.2196/33003
PMID:34635472

©Melanie Suppan, Loric Stuby, Stephan Harbarth, Christophe A Fehlmann, Sophia Achab, Mohamed Abbas, Laurent Suppan. Originally published in JMIR Serious Games (https://games.jmir.org), 25.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Virtual Reality in Health Care: Bibliometric Analysis

Christian Matthias Pawassar[1*], MSc; Victor Tiberius[1*], MBA, Prof Dr rer pol, Dr phil

Faculty of Economics and Social Sciences, University of Potsdam, Potsdam, Germany

[*]all authors contributed equally

Corresponding Author:
Victor Tiberius, MBA, Prof Dr rer pol, Dr phil
Faculty of Economics and Social Sciences
University of Potsdam
August-Bebel-Str 89
Potsdam, 14482
Germany
Phone: 49 3319773593
Email: tiberius@uni-potsdam.de

Abstract

Background: Research into the application of virtual reality technology in the health care sector has rapidly increased, resulting in a large body of research that is difficult to keep up with.

Objective: We will provide an overview of the annual publication numbers in this field and the most productive and influential countries, journals, and authors, as well as the most used, most co-occurring, and most recent keywords.

Methods: Based on a data set of 356 publications and 20,363 citations derived from Web of Science, we conducted a bibliometric analysis using BibExcel, HistCite, and VOSviewer.

Results: The strongest growth in publications occurred in 2020, accounting for 29.49% of all publications so far. The most productive countries are the United States, the United Kingdom, and Spain; the most influential countries are the United States, Canada, and the United Kingdom. The most productive journals are the Journal of Medical Internet Research (JMIR), JMIR Serious Games, and the Games for Health Journal; the most influential journals are Patient Education and Counselling, Medical Education, and Quality of Life Research. The most productive authors are Riva, del Piccolo, and Schwebel; the most influential authors are Finset, del Piccolo, and Eide. The most frequently occurring keywords other than "virtual" and "reality" are "training," "trial," and "patients." The most relevant research themes are communication, education, and novel treatments; the most recent research trends are fitness and exergames.

Conclusions: The analysis shows that the field has left its infant state and its specialization is advancing, with a clear focus on patient usability.

(JMIR Serious Games 2021;9(4):e32721) doi:10.2196/32721

KEYWORDS

virtual reality; healthcare; bibliometric analysis; literature review; citation analysis; VR; usability; review; health care

Introduction

There is no definite date for the first virtual reality (VR) application but the Sensorama device described by Morton Heilig in 1955 can be seen as a possible starting point [1,2]. His invention is one of the earliest known examples of immersive technology incorporating vision, sound, smell, as well as the sensation of touch, thereby letting users experience an illusory form of reality [2,3]. VR can be defined as a technology that makes users believe that they are in another place, based on heavily influencing the primary sensory inputs with computer-generated data [4]. What started as a pure form of entertainment involving immersion and interactivity has now become a thriving market with a wide range of use cases including recreation, communication, research, education, and health care [5-7]. VR initially started out as a luxury but has recently become more accessible, with companies like Facebook, Google, Samsung, and Sony making large investments in this technology [8,9]. The health sector benefits from this development as VR technology for entertainment purposes can also improve medical training [5,10,11], leading to a fusion of gaming and educational training called serious gaming [12]. Professionals, researchers, and students get early exposure to equipment, procedures, and clinical settings, as well as feedback,

XSL•FO
RenderX

without putting anyone at risk [5,11]. The relevance of VR in the health sector is also increasing due to breakthroughs such as O'Keefe and Moser's discovery of the brain's "GPS" (for which they received the Nobel Prize), highlighting the possibilities of VR-supported studies [13]. While analyzing the brain activity of rodents in a VR setting, these researchers found "cells that constitute a positioning system in the brain" [14,15]. Additionally, a wide array of VR engines and VR applications for experimental and computational use is becoming more accessible for everyone [14,16-18]. Lastly, VR is a driving factor in the development of new treatments or the revamping of older ones [19,20]. Novel approaches include phobia therapy [21], rehabilitation [22], and many more. These new forms of psychological and physical treatment help patients and health professionals to achieve better results in care, healing, and comfort [19,20,23,24].

The widespread use of VR makes it an increasingly relevant topic for research, leading to a growing body of scholarly literature in recent years [25], which has become more and more complex and fragmented. As a consequence, a systematic overview of research on VR in health care is needed. Our research goal is to provide such a comprehensive overview, based on a bibliometric analysis. We aim to identify current trends to promote and guide future research.

Our study supplements previous bibliometric analyses on VR in the health sector, which have a narrower focus on specific use cases, such as dementia [26], autism [27], or rehabilitation [28]. Other bibliometric analyses are outdated and do not adequately represent the current state of available material [29]. Some bibliometric analyses are not limited to VR but also include augmented reality (AR) [25,30]. Cipresso et al [31] provided a large-scale network and cluster analysis for both VR and AR across all scientific disciplines. In contrast to all these analyses, this study aims to provide a current and comprehensive overview of VR research in the health sector via a bibliometric analysis. It therefore contributes to research on VR in health care by measuring and mapping the body of literature on this topic.

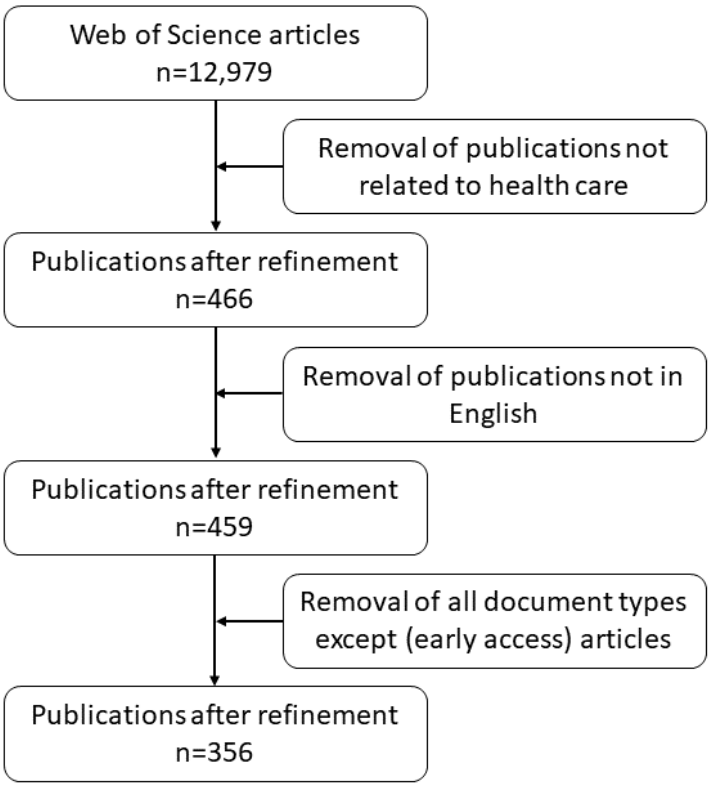
Methods

Search and Screening Strategy

We conducted a title search for "virtual reality" or "VR" on the Web of Science Core Collection on April 19, 2021. The title search ensured that only publications focused on VR were included in the data set, while publications touching on VR as a side aspect were excluded. The search yielded 12,979 publications from 1994-2021.

The data set was narrowed down by focusing on publications written in the predominant scientific language (English). We selected only publications in health care–related categories: Health Care Science Services, Public Environmental Occupational Health, Health Policy Services, and Primary Health Care. Finally, we excluded all document types other than articles ("early access" articles were included). This elimination process helped guarantee that only peer-reviewed research was included in the bibliometric analysis. The final data set consists of a total of 356 publications (Figure 1).

Figure 1. Data collection process.



XSL•FO
RenderX

Bibliometric Analysis

Bibliometrics focuses on the analysis of quantifiable publication data [32], which can be used to create objective and reproducible results [33] and offer insights into relationships between analyzed documents [34]. The basic measures for bibliometric performance analyses are the number of articles (representing productivity) and the number of citations (representing impact or influence) [35]. We focused on performance data regarding productivity and impact by year as well as by country, journal, and author. Frequencies and pro rata percentages in each category were based on publication years, total share of records, or total global citation scores. Country scores were based on the affiliation data provided by the authors themselves. Additionally, the most used, most co-occurring, and most recent keywords were examined to find and structure the main research themes. Furthermore, the twenty most important keywords were determined. These bibliometric indicators are regularly used to provide a bibliometric overview of a field [36] and to determine the most influential countries, authors, and journals of a field [37]. Based on bibliometric findings, future research trends can be derived [37].

This analysis involved the use of total global citation scores and a computation of popularity based on appearance frequency. The underlying data for the bibliometric analysis was analyzed with BibExcel (version 2016-02-20), HistCite (version 12.03.17), and VOSviewer (version 1.6.16).

Results

Annual Productivity

Table 1 depicts the annual numbers of publications on VR in health care from 1994 to April 2021. From 1994 to 2020, the average growth rate of scientific publications about VR in health care was 20.46%. Including the partial year 2021, the calculated average growth rate is currently 13.98%. The average growth rate from 1994-2010 was 10.58%, which increased to 20.54% and 60.05% for 2011-2015 and 2016-2020, respectively. The growth rate from 2016-2021 is currently 13.40%. There was a surge of scientific publications starting around 2014, accounting for 80.62% of all eligible articles. The biggest increase was in 2020, accounting for 29.49% of all analyzed scholarly works.

Overall, 88.8% (315/356) of the analyzed scientific articles were published by researchers from 20 countries (Table 1). The top three countries are the United States (94/356, 26.4%), the United Kingdom (28/356, 7.9%), and Spain (24/356, 6.7%). Another 24 countries not shown in Table 2 contributed to VR research in health care. The percentages shown in Table 1 relate to the total number of publications by the depicted countries (N=315), not the overall data set.

The country with the highest global citation score is the United States, accounting for nearly one-third of all citations in the data set (1995/6701, 29.8%). Canada is ranked second, with about one-tenth of all citations (659/6701, 9.8%), followed by the United Kingdom in third place, with about 9% of all citations (613/6701, 9.1%). The percentages in Table 3 refer to the total number of citations of the top 20 countries, not the complete data set.

XSL•FO
RenderX

Table 1. Country productivity and country impact.

Year	Number of publications
1994	1
1996	1
1997	1
2000	3
2001	1
2002	1
2003	2
2004	3
2005	4
2006	1
2007	4
2008	5
2009	8
2010	5
201	9
2012	10
2013	8
2014	15
2015	19
2016	16
2017	28
2018	31
2019	43
2020	105
2021	30
Early access	2

XSL•FO
RenderX

Table 2. Country productivity (N=315).

Rank	Country	Output, n (%)
1	United States of America	94 (29.8)
2	United Kingdom	28 (8.9)
3	Spain	24 (7.6)
4	Canada	21 (6.7)
5	People's Republic of China	18 (5.7)
6	Australia	16 (5.1)
6	Italy	16 (5.1)
7	South Korea	15 (4.8)
8	Germany	13 (4.1)
8	Taiwan	13 (4.1)
9	Netherlands	12 (3.8)
10	France	8 (2.5)
11	Brazil	6 (1.9)
12	Finland	5 (1.6)
12	Israel	5 (1.6)
12	Turkey	5 (1.6)
13	Norway	4 (1.3)
13	Poland	4 (1.3)
13	Portugal	4 (1.3)
23	Sweden	4 (1.3)

XSL•FO
RenderX

Table 3. Country impact (N=6432).

Rank	Country	Output, n (%)
1	United States	1995 (31.0)
2	Canada	659 (10.3)
3	United Kingdom	613 (9.5)
4	Italy	493 (7.7)
5	Netherlands	383 (6.0)
6	Norway	362 (5.6)
7	Germany	338 (5.3)
8	Australia	303 (4.7)
9	Spain	267 (4.2)
10	Switzerland	256 (4.0)
11	Belgium	147 (2.3)
12	Israel	112 (1.7)
13	Taiwan	100 (1.6)
14	People's Republic of China	98 (1.5)
15	Singapore	63 (1.0)
16	France	60 (0.9)
17	Mexico	58 (0.9)
18	New Zealand	47 (0.7)
19	India	43 (0.7)
20	Finland	35 (0.5)

Journal Productivity and Impact

The highest-ranking journal regarding productivity is the Journal of Medical Internet Research, with 34 of the 356 total publications (9.6%). In second place is JMIR Serious Games (29/356, 8.2%), followed by the Games for Health Journal in third position (28/356, 7.9%). The percentages in Table 3 relate to the total number of articles in the top 20 most productive journals; the articles in the full data set were published in 70 additional journals not shown in Table 4.

The most cited journal is Patient Education and Counselling, with a total global citation score of 428 of the 3993 citations in the entire data set (10.7%). In second place is Medical Education (312/3993, 7.8%), followed by Quality of Life Research (281/3993, 7%). Again, out of the 90 journals included in the data set, 70 are not shown in Table 5. The percentages given in the table relate to the 3105 citations received by the top 20 journals, not the whole data set.

The three authors with the highest productivity are Riva (9/356, 3%), del Piccolo (6/356, 2%), and Schwebel (6/356, 2%). Out of 1698 authors in the data set, 1678 are not shown in Table 6. Most authors (1532, 90.2%) only published one article on VR in health care, while 129 (7.6%) published two qualifying articles.

The author with the highest impact as determined by overall citations is Finset, with a total of 340 citations of 20,363 (1.67%) in the whole data set. The next most influential authors are del Piccolo and Eide, both in second place with 308 citations (308/20,363, 1.51%). Taking all citations and authors into account, the average is 20.36 citations per researcher. The computed h-index of an author (shown in Table 7) relates to the subject of this bibliometric analysis, and does not include other publications by the author. Once again, the displayed percentage relates to the total number of citations represented in the table, not the whole data set.

XSL•FO
RenderX

Table 4. Journal productivity (N=235).

Rank	Journal	Output, n (%)
1	Journal of Medical Internet Research	34 (14.5)
2	JMIR Serious Games	29 (12.3)
3	Games for Health Journal	28 (11.9)
4	International Journal of Environmental Research and Public Health	27 (11.5)
5	Patient Education and Counselling	16 (6.8)
6	Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare	11 (4.7)
7	Journal of Healthcare Engineering	10 (4.3)
8	Technology and Healthcare	10 (4.3)
9	Accident Analysis and Prevention	8 (3.4)
9	Annual Review of Cybertherapy and Telemedicine 2015: Virtual Reality in Healthcare: Medical Simulation and Experiential Interface	8 (3.4)
10	Annual Review of Cybertherapy and Telemedicine 2014: Positive Change: Connecting the Virtual and the Real	7 (3.0)
11	Frontiers in Public Health	6 (2.6)
11	Journal of Medical Systems	6 (2.6)
12	Aerospace Medicine and Human Performance	5 (2.1)
12	Aviation Space and Environmental Medicine	5 (2.1)
12	Injury Prevention	5 (2.1)
12	International Journal of Occupational Safety and Ergonomics	5 (2.1)
12	JMIR Research Protocols	5 (2.1)
12	Methods of Information in Medicine	5 (2.1)
12	Work: A Journal of Prevention Assessment & Rehabilitation	5 (2.1)

XSL•FO
RenderX

Table 5. Journal impact (N=3105).

Rank	Journal	Citations, n (%)
1	Patient Education and Counselling	428 (13.8)
2	Medical Education	312 (10.0)
3	Quality of Life Research	281 (9.1)
4	Aviation Space and Environmental Medicine	247 (8.0)
5	Games for Health Journal	234 (7.5)
6	Academic Medicine	213 (6.9)
7	Accident Analysis and Prevention	196 (6.3)
8	Journal of Medical Internet Research	142 (4.6)
9	Nicotine & Tobacco Research	111 (3.6)
10	Supportive Care in Cancer	106 (3.4)
11	Telemedicine Journal and e-Health	101 (3.3)
12	Methods of Information in Medicine	100 (3.2)
13	Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare	90 (2.9)
14	International Journal on Disability and Human Development	89 (2.9)
15	Journal of Medical Systems	80 (2.6)
15	Simulation in Healthcare	80 (2.6)
16	Psychology & Health	75 (2.4)
16	Technology and Health Care	75 (2.4)
17	International Journal of Medical Informatics	74 (2.4)
18	JMIR Serious Games	71 (2.3)

XSL•FO
RenderX

Table 6. Author productivity (N=87).

Rank	Author	Publications, n (%)
1	G Riva	9 (10)
2	L del Piccolo	6 (7)
2	DC Schwebel	6 (7)
3	A Finset	5 (6)
3	J Gutierrez-Maldonado	5 (6)
3	G Humphris	5 (6)
3	BK Wiederhold	5 (6)
3	SC Yeh	5 (6)
4	H Eide	4 (5)
4	M Ferrer-Garcia	4 (5)
4	L Smith	4 (5)
4	MD Wiederhold	4 (5)
4	C Zimmermann	4 (5)
5	SAW Andersen	3 (3)
5	WP Brinkman	3 (3)
5	R Cano de la Cuerda	3 (3)
5	P Cipresso	3 (3)
5	A Fisher	3 (3)
5	I Fletcher	3 (3)
5	C Goss	3 (3)

XSL•FO
RenderX

Table 7. Author impact (N=5459).

Rank	Author	Citations, n (%)	H-index
1	A Finset	340 (6.2)	5
2	L del Piccolo	308 (5.6)	5
2	H Eide	308 (5.6)	4
3	G Humphris	300 (5.5)	5
4	C Zimmermann	297 (5.4)	4
5	W Rogers	281 (5.2)	3
6	M Rimondini	273 (5.0)	3
7	C Goss	273 (5.0)	3
8	I Fletcher	271 (5.0)	3
9	YM Kim	256 (4.7)	2
9	S Bergvik	256 (4.7)	2
9	P Salmon	256 (4.7)	2
9	J Bensing	256 (4.7)	2
9	C Heaven	256 (4.7)	2
9	L Zandbelt	256 (4.7)	2
9	W Langewitz	256 (4.7)	2
9	S van Dulmen	256 (4.7)	2
9	H de Haes	256 (4.7)	2
10	L Wissow	256 (4.7)	2
11	S Qian	248 (4.5)	2

Most Used and Co-Occurring Keywords

As shown in Table 8, the most used keyword is “virtual” (322/356, 90.4%), followed by “reality” (321/356, 90.2%). The third most used keyword is “training” (60/356, 16.9%).

The author keyword co-occurrence analysis led to 11 clusters, with a minimum citation threshold of 3 and 94 qualifying keywords. Figure 2 depicts the keyword co-occurrence map generated by VOSviewer. The 11 clusters are distinguished by different colors. The clusters comprise keywords, which are often mentioned together in the keyword lists of the articles in the literature sample and therefore have the tendency to form groups that represent common research themes. Of these clusters, 3 are relatively small, with only 2-3 keywords. Half of the clusters are made up of 10 or more keywords. Being presented as a cluster does not necessarily mean that publications deal with the same topic, albeit all qualifying publications within a respective cluster complement each other. Furthermore, 3 clusters have an everyday focus, 3 other ones address the specifics of VR training, 4 have diverse health care–related fields as their common topic, and one is centered around known issues in VR settings. The clusters are described in greater detail below.

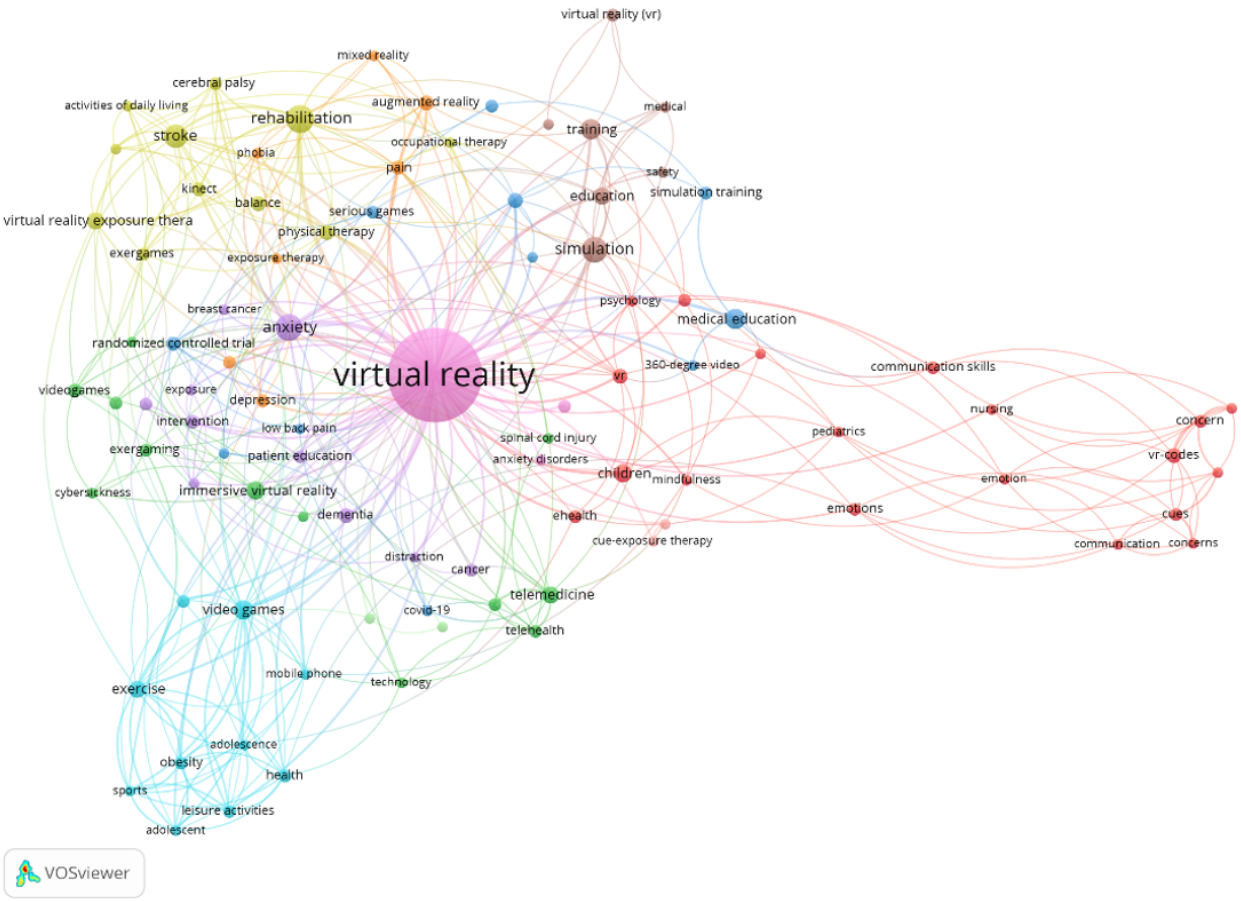
XSL•FO
RenderX

Table 8. Most used keywords.

Rank	Keyword	Occurrences (N=1168), n (%)	Overall percentage of occurrences (n=356)
1	Virtual	322 (27.6)	90.4
2	Reality	321 (27.5)	90.2
3	Training	60 (5.1)	16.9
4	Based	55 (4.7)	15.4
5	Using	51 (4.4)	14.3
6	Trial	36 (3.1)	10.1
7	Randomized	32 (2.7)	9.0
8	Patients	30 (2.6)	8.4
9	Therapy	29 (2.5)	8.1
10	Simulation	28 (2.4)	7.9
11	Controlled	26 (2.2)	7.3
12	Health	24 (2.0)	6.7
13	Children	23 (2.0)	6.5
14	Pilot	22 (1.9)	6.2
15	Rehabilitation	20 (1.7)	5.6
16	Effects	19 (1.6)	5.3
16	Immersive	19 (1.6)	5.3
17	Care	17 (1.5)	4.8
17	Effect	17 (1.5)	4.8
27	Treatment	17 (1.5)	4.8

XSL•FO
RenderX

Figure 2. Keyword co-occurrence map.



Clusters

Cluster 1: Communication, Especially in Pediatrics

This large cluster covers 19 keywords revolving around promoting a more rounded caretaker/patient dialogue, with a focus on children. Some papers address issues regarding hindered communication with children and how VR can help with that [38,39], as well as the development of communication skills for patients [40,41] and medical staff [42,43].

Cluster 2: On-site or Telemedicine Health Care

This cluster containing 12 items highlights some positive aspects in caretaking via VR. A rather large focus lies with the enrichment and possibilities of interventions and care from a distance offering new and effective methods [21,44,45]. A smaller portion sees VR in use for on-site interventions [46].

Cluster 3: Physical Health Training

This cluster with 12 items specifies similarities and discrepancies between fitness training in real-life and VR settings. Findings of articles within this cluster are that VR-based training is effective but not more effective than real-life trainings [47,48]. VR-supported physical exercise can be applied to both private fitness endeavors and physical therapy, as the next cluster shows.

Cluster 4: Physical Rehabilitation

This group of keywords containing 11 items addresses the flexibility of VR in addition to conventional rehabilitation programs [49,50], and the possibility of expanding the reach of treatments beyond clinically controlled settings to achieve better results [22]. In addition, this cluster includes the adaption of VR games detached from any clinical setting and found the VR games had a positive effect on patient rehabilitation comparable to trainings conducted by medical professionals [51,52].

Cluster 5: Geriatric Care

This section containing 10 items centers around the care of older adults. VR was used to tackle dementia and memory loss [53-55]. Other application areas include VR interventions in life-threatening situations to improve patients' moods [56], as well as patient education and training.

Cluster 6: Motivation, Health, and Adolescents

This cluster, also containing 10 items, addresses the need to motivate adolescents to partake in physical activities to tackle motivational and health issues. The focus is on cultivating fascination and enjoyment in young adults regarding the possibilities of VR and educating parents to promote the consideration of VR where antipathy might be prevalent [57-59]. This cluster underscores the need for awareness training.

XSL•FO
RenderX

Cluster 7: Psychological Treatments

This 7-item cluster focuses on new treatment possibilities, particularly for depression and stress disorders. Topics range from posttraumatic experience and chronic disorder treatments to preventive strategies [60-62], increasing the need for more training about the increased options for effective handling.

Cluster 8: Accident Prevention

This group of 7 keywords includes ophthalmology research and implements eye-tracking features to analyze behavior in dangerous situations. This is feasible because VR headsets are often very close to the eyes, and already analyze how and where the user moves in a predefined space [63-65], offering a safe environment for researchers to analyze test subjects in otherwise unworkable experimental situations.

Cluster 9: Palliative Care

This rather small cluster comprises only 3 keywords. The articles in this cluster focus on anxiety in people with terminal diseases and how to broaden the range of effective treatments [66]. It is closely related to cluster 7, with a focus on providing relief for the patient. For effective use, more training for caretakers needs to take place.

Cluster 10: Everyday Health Support

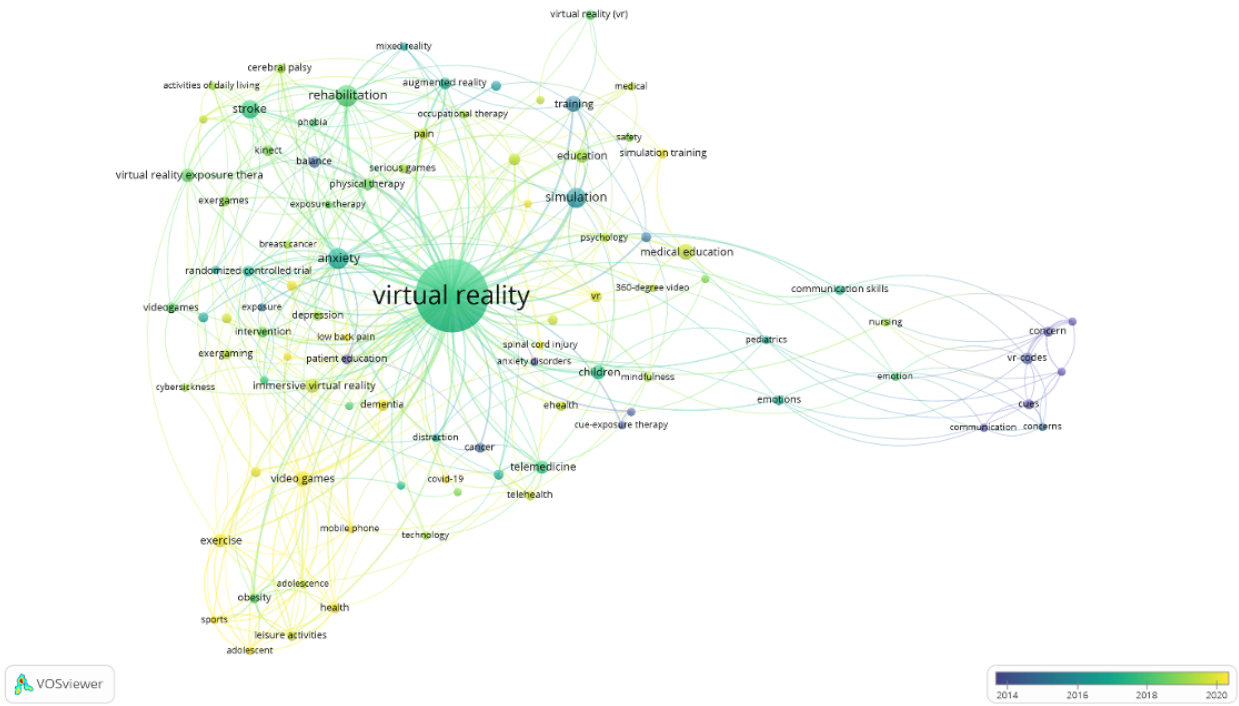
This small cluster with 2 keywords promotes interventions in relation to binge eating [67,68], but future fields of application include other interventions for working toward a healthier life every day via small reminders.

Cluster 11: Aftereffects of VR Environments

The last cluster also includes 2 keywords and addresses issues that may occur when people experience or cease using VR [69,70]. This is relevant for everyday life as well as trainings to increase awareness.

Figure 3 depicts the trend evolution of keywords. The map shows how the importance of keywords changed for VR in health care starting around 2013-2014 (blue) to the current date (yellow). Earlier focal points were about concerns with the new technology and possible fields of use. A selection of representative keywords is “concern,” “communication,” “anxiety disorders,” and “patient education.” A few years later, these application fields broadened and started to include medicine-related trends. This is seen in keywords such as “telemedicine,” “physical therapy,” “exposure therapy,” and “randomized controlled trial.” The latest trend sees VR in health care become centered around education and exergames, so-called exercise games for mental and physical health, as well as mobility for virtual realities. Representative keywords are “sport(s),” “video games,” “simulation training,” and “mobile phone.”

Figure 3. Keyword trend from 2013-2021.



Discussion

Overview

In this study, research focusing on VR in health care from 1994 to 2021 was analyzed by bibliometric measures. The focal points for this analysis were publication and author scores, as well as a trend analysis. Overall, research has been steadily increasing, with a recent spike in publications, underlining the relevance of research on VR applications in health care.

Growth Rate of Publications

Although VR emerged in the 20th century and has been a topic in health care–related research for over 25 years, it was initially

XSL•FO
RenderX

a niche topic [2]. An increase in the number of VR publications can be seen between 2015 and 2017, with a notable spike in 2020. The growth rate nearly tripled to 60.05% when compared to the previous 5 years (20.54%). What may have been a decisive factor in this change was the sudden increase in investments and availability of high-quality technology, which increased scientific interest and use [8]. The gaming industry had a major impact, showcasing the manifold (though mostly unintentional) benefits and possibilities via development of new headsets and programs, which could be appropriated and modified for new applications [11,14,17]. Since then, an increasing number of medical fields saw the potential of VR for data collection, treatments, or trainings [10,41,65]. Following the past growth pattern of health care–related VR research, it can be expected that this field will continue to grow.

Publication Pattern

VR research is being conducted around the globe, with over one-fifth of all countries contributing to scientific progress in this field. The United States, as in many other fields, tends to lead research in this field. The top countries are consistently similar across different rankings, at most just swapping placements [71,72]. Germany, Spain, and the United Kingdom are also driving forces in the development of other digital health interventions such as telemedicine and artificial intelligence [72,73]. However, there is a notable absence of emerging countries, which could reap the benefits of the digitalization of their health sectors [74], especially as VR technology becomes more affordable [9]. The recent positive VR market development could help these countries to overcome obstacles including funding, distribution prioritization, and language support, promoting a future shift in research hotspots and more exponential growth in publications [75]. After all, research on VR and related topics of digitalization could have a large combined positive effect for developed and developing countries, which may result in cooperation and dialogue toward future progress.

The journals publishing VR-related health care content are mostly high-impact and well-known journals but lower-impact journals also add to the increasing reach and availability of research. With an average of 5 authors per publication (1698/256), VR in health care has seen growth in co-authorship over time. In terms of citations, the high-impact journals take the lead.

Scientists' productivity and impact do not always go hand in hand. Due to the novelty and rapid development of the field, well-established scientists or articles do not exist yet. This could change over time, as rapid growth based on novelty is usually finite.

Trends and Research Themes

After clustering the author keywords, clear distinctions were apparent, with the currently highest trending area of research involving communication and the integration of VR for educational purposes. This topic spans over manifold instances like patient and caretaker interactions, communication training, and new forms of education [10,38,43]. The second most prominent field combines telemedicine and VR, with a foray into the development of VR-based "medical professionals." The focus is on the implementation of a VR setting in which health care professionals do not necessarily need to be physically present, or even involved in the procedure at all [45,46]. This also includes the development of new treatments using VR, such as for phobias and mental disorders [21,61]. The third most important area is centered around physical activities for both younger and older people [47,59]. The findings here see great potential and marketing opportunities for a healthy lifestyle and better life expectancy by using VR.

When looking at the prominence of keywords from 2013 to 2021, the ranking is reversed. Modern research focuses on the fitness and health opportunities of VR and on (smaller) interventions to promote better health behavior, while the communication-related and educational aspects of VR in health care tend to be less researched. In particular, the surge of interest in this topic started at the end of 2019, which could be directly related to the COVID-19 outbreak, during which VR fitness may have been one of the few viable options for implementing physical activity into daily life. Another reason for the reversed relevance of keywords might be the latest developments in VR technology, such as omnidirectional treadmills, optimized gaming gear, and newly released games using new technologies that represent untapped research potential for fitness and entertainment [76,77].

Health care–related VR research has seen strong growth in the past decades. Current contexts and possibilities underline the potential of this field to gain more traction in the near future. The cross-sectional methods for implementation and simple ways of integrating VR into current systems are what makes the technology so interesting and worth researching. As a field, VR has passed the early first steps of development and is no longer under the radar. It can be expected that, with further technological progress, the availability of VR will increase, so that emerging countries can increasingly benefit from this technology.

However, VR in health care also has some obstacles, such as motion sickness, a lost sense of presence, eye strain, or inappropriate responses in the real world [78,79]. Future research can be expected to focus on how to tackle these challenges. In particular, holographic projections have the potential to alleviate many of the negative symptoms of VR [80]. Another issue is telemedicine, which is currently limited to algorithms or an AI responding to medical personnel or patients [42], rather than focusing on VR. Future integrations could see cybernetics become an essential part of VR to probe into autonomous health care with the help of robotics [81]. Even though this raises many questions and involves barriers [82], it could become a necessity to tackle the problem of a lack of health care professionals who can take care of the increasingly aging population [83]. Older adults' acceptance of technology can also be challenging in this context. However, recent research seems to suggest these claims may be unfounded [83]. Future research could therefore provide guidelines to increase or secure older patients' acceptance. Only a few standardized guidelines have been developed so far [84].

XSL•FO
RenderX

Limitations

This bibliometric analysis has some limitations. First, our search was limited to the Web of Science, which is a widely used academic database. However, the use of other databases, such as Scopus or Google Scholar, may have provided slightly different results. Second, our analysis only included articles published in English, the lingua franca of science. The inclusion of other languages, grey literature, and books might also have led to different outcomes, especially as scholars from different cultures might have different perspectives on the use of VR in health care. Third, searching article titles alone and using only two search terms was very limiting. As our goal was to focus on literature that deals with VR as the main research theme, a title search was more appropriate than a topic search. However, articles dealing with VR as a side aspect might also enrich the body of knowledge in the field. Therefore, we encourage future research to also use topic searches including abstracts and keywords, and extend the range of search terms. Fourth, the inclusion of the specified disciplines alone might have excluded relevant articles that have been published in more technologically oriented journals rather than heath care–related ones. Future research might therefore use "health*" as an additional search term rather than using a filter based on disciplines. Fifth, whereas publication and citation numbers are objective, the interpretation of keyword clusters has a subjective character; other researchers may have drawn different conclusions.

An area which should also be pursued in future studies, which has not seen enough attention thus far, is the possible interconnection between AR and VR, which are closely related. However, mix-ups between these two occur, and AR might be as feasible for use in health care as VR. In addition, they might offer benefits upon their combined use, which should be analyzed further.

Conclusion

This bibliometric analysis aimed to give an overview of VR-related research in health care. It comprises 356 publications across about 27 years, from 1994 to 2021 (partial year). The main results are the following: (1) VR-related publications in health care have seen increased growth, (2) developed countries are the driving force in health care–related VR research but the topic has already been researched around the world, (3) the three predominant research themes center around communication, education, and novel treatments, and (4) the most recent research trends cover fitness and exergames for VR in health care. The analysis shows that the field has left its infant state and the research is becoming increasingly specialized, with a clear focus on patient usability. Future research should broaden the range of involved countries, industries, and companies.

Conflicts of Interest

None declared.

References

1. Jones S, Dawkins S. The Sensorama revisited: evaluating the application of multi-sensory input on the sense of presence in 360-degree immersive film in virtual reality. In: Augmented reality and virtual reality. Cham: Springer; 2018:183-197.
2. Morton Heilig: Inventor VR. USC School of Cinematic Arts. URL: https://www.uschefnerarchive.com/morton-heilig-inventor-vr/ [accessed 2021-04-23]
3. Yoh MS. The reality of virtual reality. 2001 Presented at: Seventh International Conference on Virtual Systems and Multimedia; 2001; Berkeley p. 666-674. [doi: 10.1109/vsmm.2001.969726]
4. Heim M. Virtual Realism. Oxford: Oxford University Press; 1998.
5. Zheng JM, Chan KW, Gibson I. Virtual reality. IEEE Potentials 1998;17(2):20-23. [doi: 10.1109/45.666641]
6. Pedroli E, Serino S, Pallavicini F, Cipresso P, Riva G. Exploring virtual reality for the assessment and rehabilitation of executive functions. International Journal of Virtual and Augmented Reality 2018;2(1):32-47. [doi: 10.4018/IJVAR.2018010103]
7. Liou HH, Yang SJH, Chen SY, Tarng W. The influences of the 2D image-based augmented reality and virtual reality on student learning. Educational Technology & Society 2017;20(3):110-121.
8. Castelvecchi D. Low-cost headsets boost virtual reality's lab appeal. Nature 2016;533(7602):153-154. [doi: 10.1038/533153a] [Medline: 27172022]
9. Ebert C. Looking into the Future. IEEE Software 2015;32(6):92-97. [doi: 10.1109/ms.2015.142]
10. Rubino F, Soler L, Marescaux J, Maisonneuve H. Advances in virtual reality are wide ranging. BMJ 2002;324(7337):612 [FREE Full text] [doi: 10.1136/bmj.324.7337.612] [Medline: 11884336]
11. McCloy R, Stone R. Science, medicine, and the future. Virtual reality in surgery. BMJ 2001;323(7318):912-915 [FREE Full text] [doi: 10.1136/bmj.323.7318.912] [Medline: 11668138]
12. Checa D, Bustillo A. A review of immersive virtual reality serious games to enhance learning and training. Multimed Tools Appl 2019 Dec 05;79(9-10):5501-5527. [doi: 10.1007/s11042-019-08348-9]
13. Burgess N. The 2014 Nobel Prize in Physiology or Medicine: a spatial model for cognitive neuroscience. Neuron 2014;84(6):1120-1125 [FREE Full text] [doi: 10.1016/j.neuron.2014.12.009] [Medline: 25521374]
14. Minderer M, Harvey CD, Donato F, Moser EI. Neuroscience: Virtual reality explored. Nature 2016 May 19;533(7603):324-325. [doi: 10.1038/nature17899] [Medline: 27193673]

XSL•FO
RenderX

15. The 2014 Nobel Prize in Physiology or Medicine. Nobel Prize. 2014. URL: https://www.nobelprize.org/prizes/medicine/2014/press-release [accessed 2021-05-07]
16. Riva G, Gaggioli A, Grassi A, Raspelli P, Cipresso P, Pallavicini F, et al. NeuroVR 2 – A free virtual reality platform for the assessment and treatment in behavioral health care. Medicine Meets Virtual Reality 2011;18(1):493-495. [doi: 10.3233/978-1-60750-706-2-493]
17. Brown A, Green T. Virtual Reality: Low-Cost Tools and Resources for the Classroom. TechTrends 2016 Jun 30;60(5):517-519. [doi: 10.1007/s11528-016-0102-z]
18. Cipresso P, Serino S, Riva G. Psychometric assessment and behavioral experiments using a free virtual reality platform and computational science. BMC Med Inform Decis Mak 2016;16(1):37 [FREE Full text] [doi: 10.1186/s12911-016-0276-5] [Medline: 26992890]
19. Botella C, Fernández-Álvarez J, Guillén V, García-Palacios A, Baños R. Recent progress in virtual reality exposure therapy for phobias: A systematic review. Curr Psychiatry Rep 2017;19(7):42. [doi: 10.1007/s11920-017-0788-4] [Medline: 28540594]
20. Alahmari KA, Sparto PJ, Marchetti GF, Redfern MS, Furman JM, Whitney SL. Comparison of Virtual Reality Based Therapy With Customized Vestibular Physical Therapy for the Treatment of Vestibular Disorders. IEEE Trans Neural Syst Rehabil Eng 2014 Mar;22(2):389-399. [doi: 10.1109/tnsre.2013.2294904]
21. Levy F, Leboucher P, Rautureau G, Jouvent R. E-virtual reality exposure therapy in acrophobia: A pilot study. J Telemed Telecare 2015 Aug 06;22(4):215-220. [doi: 10.1177/1357633x15598243]
22. Tunney N, Perlow E. Expanding Rehabilitation Beyond the Clinic—Strategies to Increase Total Restorative Therapy Time for Adults with Hemiplegia. IJAHSP 2021;19(1):9. [doi: 10.46743/1540-580x/2021.1951]
23. Carl E, Stein AT, Levihn-Coon A, Pogue JR, Rothbaum B, Emmelkamp P, et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. J Anxiety Disord 2019;61:27-36. [doi: 10.1016/j.janxdis.2018.08.003] [Medline: 30287083]
24. Hoffman HG, Patterson DR, Carrougher GJ. Use of virtual reality for adjunctive treatment of adult burn pain during physical therapy: a controlled study. Clin J Pain 2000 Sep;16(3):244-250. [doi: 10.1097/00002508-200009000-00010] [Medline: 11014398]
25. Cipresso P, Giglioli IAC, Raya MA, Riva G. The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature. Front Psychol 2018;9:2086 [FREE Full text] [doi: 10.3389/fpsyg.2018.02086] [Medline: 30459681]
26. Sobral M, Pestana M. Virtual reality and dementia: A bibliometric analysis. The European Journal of Psychiatry 2020 Jul;34(3):120-131. [doi: 10.1016/j.ejpsy.2020.04.004]
27. Fernández-Herrero J, Lorenzo-Lledó G, Carreres AL. A bibliometric study on the use of virtual reality (VR) as an educational tool for high-functioning autism spectrum disorder (ASD) children. In: Contemporary Perspective on Child Psychology and Education. London: IntechOpen; 2017.
28. Huang Y, Huang Q, Ali S, Zhai X, Bi X, Liu R. Rehabilitation using virtual reality technology: a bibliometric analysis, 1996–2015. Scientometrics 2016 Oct 05;109(3):1547-1559. [doi: 10.1007/s11192-016-2117-9]
29. Zhang T, Liu A. The research status on virtual reality by using of bibliometric analysis. 2010 Presented at: 2010 2nd International Conference on Computer Engineering and Technology; 2010; Chengdu p. 289-292. [doi: 10.1109/iccet.2010.5485415]
30. Calabuig-Moreno F, González-Serrano MH, Fombona J, García-Tascón M. The Emergence of Technology in Physical Education: A General Bibliometric Analysis with a Focus on Virtual and Augmented Reality. Sustainability 2020 Mar 30;12(7):2728. [doi: 10.3390/su12072728]
31. Cipresso P, Giglioli IAC, Raya MA, Riva G. The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature. Front Psychol 2018;9:2086 [FREE Full text] [doi: 10.3389/fpsyg.2018.02086] [Medline: 30459681]
32. Schaer P. Applied informetrics for digital libraries: an overview of foundations, problems and current approaches. Historical Social Research 2013;38(3):267-281.
33. Glänzel W. The need for standards in bibliometric research and technology. Scientometrics 1996 Feb;35(2):167-176. [doi: 10.1007/bf02018475]
34. Bibliometrics and citation analysis. University of Wisconsin-Madison Libraries. URL: https://researchguides.library.wisc.edu/c.php?g=178134&p=1169079 [accessed 2021-04-25]
35. Gutiérrez-Salcedo M, Martínez M, Moral-Munoz JA, Herrera-Viedma E, Cobo MJ. Some bibliometric procedures for analyzing and evaluating research fields. Appl Intell 2018;48:1275-1287. [doi: 10.1007/s10489-017-1105-y]
36. Narin F, Hamilton KS. Bibliometric performance measures. Scientometrics 1996 Jul;36(3):293-310. [doi: 10.1007/bf02129596]
37. Marczewska M, Kostrzewski M. Sustainable Business Models: A Bibliometric Performance Analysis. Energies 2020 Nov 19;13(22):6062. [doi: 10.3390/en13226062]

XSL•FO
RenderX

38. Gilardi F, De Falco F, Casasanta D, Andellini M, Gazzellini S, Petrarca M, et al. Robotic Technology in Pediatric Neurorehabilitation. A Pilot Study of Human Factors in an Italian Pediatric Hospital. Int J Environ Res Public Health 2020 May 17;17(10):3503 [FREE Full text] [doi: 10.3390/ijerph17103503] [Medline: 32429562]
39. Ahmadpour N, Weatherall AD, Menezes M, Yoo S, Hong H, Wong G. Synthesizing Multiple Stakeholder Perspectives on Using Virtual Reality to Improve the Periprocedural Experience in Children and Adolescents: Survey Study. J Med Internet Res 2020 Jul 17;22(7):e19752 [FREE Full text] [doi: 10.2196/19752] [Medline: 32706671]
40. Romero-Ayuso D, Alcántara-Vázquez P, Almenara-García A, Nuñez-Camarero I, Triviño-Juárez JM, Ariza-Vega P, et al. Self-Regulation in Children with Neurodevelopmental Disorders "SR-MRehab: Un Colegio Emocionante": A Protocol Study. Int J Environ Res Public Health 2020 Jun 12;17(12):4198 [FREE Full text] [doi: 10.3390/ijerph17124198] [Medline: 32545534]
41. Kron FW, Fetters MD, Scerbo MW, White CB, Lypson ML, Padilla MA, et al. Using a computer simulation for teaching communication skills: A blinded multisite mixed methods randomized controlled trial. Patient Educ Couns 2017;100(4):748-759. [doi: 10.1016/j.pec.2016.10.024] [Medline: 27939846]
42. Daher S, Hochreiter J, Schubert R, Gonzalez L, Cendan J, Anderson M, et al. The Physical-Virtual Patient Simulator. Sim Healthcare 2020;15(2):115-121. [doi: 10.1097/sih.0000000000000409]
43. Frost J, Delaney L, Fitzgerald R. Exploring the application of mixed reality in Nurse education. BMJ STEL 2019 Aug 21;6(4):214-219. [doi: 10.1136/bmjstel-2019-000464]
44. Wiederhold BK, Riva G, Wiederhold MD. Annual Review of Cybertherapy and Telemedicine 2015: Virtual Reality in Healthcare: Medical Simulation and Experiential Interface. Amsterdam, Berlin, Washington DC: IOS Press; 2016.
45. Knepley KD, Mao JZ, Wieczorek P, Okoye FO, Jain AP, Harel NY. Impact of Telerehabilitation for Stroke-Related Deficits. Telemed J E Health 2021 Mar;27(3):239-246. [doi: 10.1089/tmj.2020.0019] [Medline: 32326849]
46. Tielman ML, Neerincx MA, Bidarra R, Kybartas B, Brinkman W. A Therapy System for Post-Traumatic Stress Disorder Using a Virtual Agent and Virtual Storytelling to Reconstruct Traumatic Memories. J Med Syst 2017 Aug;41(8):125 [FREE Full text] [doi: 10.1007/s10916-017-0771-y] [Medline: 28699083]
47. Sarig Bahat H, Hadar D, Treleaven J. Predictors for Positive Response to Home Kinematic Training in Chronic Neck Pain. J Manipulative Physiol Ther 2020 Oct;43(8):779-790. [doi: 10.1016/j.jmpt.2019.12.008] [Medline: 32829943]
48. Gouveia E Silva EC, Lange B, Bacha JMR, Pompeu JE. Effects of the Interactive Videogame Nintendo Wii Sports on Upper Limb Motor Function of Individuals with Post-Polio Syndrome: A Randomized Clinical Trial. Games Health J 2020 Dec;9(6):461-471. [doi: 10.1089/g4h.2019.0192] [Medline: 32673102]
49. Abbas RL, Cooreman D, Al Sultan H, El Nayal M, Saab IM, El Khatib A. The Effect of Adding Virtual Reality Training on Traditional Exercise Program on Balance and Gait in Unilateral, Traumatic Lower Limb Amputee. Games Health J 2021 Feb;10(1):50-56. [doi: 10.1089/g4h.2020.0028] [Medline: 33533682]
50. García-Bravo S, Cano-de-la-Cuerda R, Domínguez-Paniagua J, Campuzano-Ruiz R, Barreñada-Copete E, López-Navas MJ, et al. Effects of Virtual Reality on Cardiac Rehabilitation Programs for Ischemic Heart Disease: A Randomized Pilot Clinical Trial. Int J Environ Res Public Health 2020 Nov 16;17(22):8472 [FREE Full text] [doi: 10.3390/ijerph17228472] [Medline: 33207670]
51. Feyzioğlu Ö, Dinçer S, Akan A, Algun ZC. Is Xbox 360 Kinect-based virtual reality training as effective as standard physiotherapy in patients undergoing breast cancer surgery? Support Care Cancer 2020 Sep;28(9):4295-4303. [doi: 10.1007/s00520-019-05287-x] [Medline: 31907649]
52. Shahmoradi L, Almasi S, Ghotbi N, Gholamzadeh M. Learning promotion of physiotherapy in neurological diseases: Design and application of a virtual reality-based game. J Educ Health Promot 2020;9(1):234 [FREE Full text] [doi: 10.4103/jehp.jehp_736_19] [Medline: 33209926]
53. Torpil B, Şahin S, Pekçetin S, Uyanık M. The Effectiveness of a Virtual Reality-Based Intervention on Cognitive Functions in Older Adults with Mild Cognitive Impairment: A Single-Blind, Randomized Controlled Trial. Games Health J 2021 Apr;10(2):109-114. [doi: 10.1089/g4h.2020.0086] [Medline: 33058735]
54. Saredakis D, Keage HA, Corlis M, Loetscher T. Using Virtual Reality to Improve Apathy in Residential Aged Care: Mixed Methods Study. J Med Internet Res 2020 Jun 26;22(6):e17632 [FREE Full text] [doi: 10.2196/17632] [Medline: 32469314]
55. Coelho T, Marques C, Moreira D, Soares M, Portugal P, Marques A, et al. Promoting Reminiscences with Virtual Reality Headsets: A Pilot Study with People with Dementia. Int J Environ Res Public Health 2020 Dec 12;17(24):9301 [FREE Full text] [doi: 10.3390/ijerph17249301] [Medline: 33322679]
56. Baños RM, Espinoza M, García-Palacios A, Cervera JM, Esquerdo G, Barrajón E, et al. A positive psychological intervention using virtual reality for patients with advanced cancer in a hospital setting: a pilot study to assess feasibility. Support Care Cancer 2013;21(1):263-270. [doi: 10.1007/s00520-012-1520-x] [Medline: 22688373]
57. Farič N, Smith L, Hon A, Potts HWW, Newby K, Steptoe A, et al. A Virtual Reality Exergame to Engage Adolescents in Physical Activity: Mixed Methods Study Describing the Formative Intervention Development Process. J Med Internet Res 2021 Feb 04;23(2):e18161 [FREE Full text] [doi: 10.2196/18161] [Medline: 33538697]
58. McMichael L, Farič N, Newby K, Potts HWW, Hon A, Smith L, et al. Parents of Adolescents Perspectives of Physical Activity, Gaming and Virtual Reality: Qualitative Study. JMIR Serious Games 2020 Aug 25;8(3):e14920 [FREE Full text] [doi: 10.2196/14920] [Medline: 32840487]

XSL•FO
RenderX

59. Farič N, Yorke E, Varnes L, Newby K, Potts HW, Smith L, et al. Younger Adolescents' Perceptions of Physical Activity, Exergaming, and Virtual Reality: Qualitative Intervention Development Study. JMIR Serious Games 2019 Jun 17;7(2):e11960 [FREE Full text] [doi: 10.2196/11960] [Medline: 31210135]
60. Lyons KD, Slaughenhaupt RM, Mupparaju SH, Lim JS, Anderson AA, Stankovic AS, et al. Autonomous Psychological Support for Isolation and Confinement. Aerosp Med Hum Perform 2020 Nov 01;91(11):876-885. [doi: 10.3357/amhp.5705.2020]
61. Veling W, Lestestuiver B, Jongma M, Hoenders HJR, van Driel C. Virtual Reality Relaxation for Patients With a Psychiatric Disorder: Crossover Randomized Controlled Trial. J Med Internet Res 2021 Jan 15;23(1):e17233 [FREE Full text] [doi: 10.2196/17233] [Medline: 33448933]
62. Zhang Z, Qi B, Xu Y, Jin Y, Gao B. The Effect of Hyperbaric Oxygen Combined with Virtual Reality Training on Oxidative Stress Indicators and Inflammatory Factors of Swimming Athletes Suffering from Depression. Journal of Healthcare Engineering 2021 Mar 11;2021:1-8. [doi: 10.1155/2021/6645515]
63. Zeuwts LH, Vansteenkiste P, Deconinck FJ, Cardon G, Lenoir M. Hazard perception training in young bicyclists improves early detection of risk: A cluster-randomized controlled trial. Accid Anal Prev 2017 Nov;108:112-121. [doi: 10.1016/j.aap.2017.08.024] [Medline: 28865926]
64. 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;5(3):e17 [FREE Full text] [doi: 10.2196/games.7923] [Medline: 28842388]
65. Sipatchin A, Wahl S, Rifai K. Eye-Tracking for Clinical Ophthalmology with Virtual Reality (VR): A Case Study of the HTC Vive Pro Eye's Usability. Healthcare (Basel) 2021 Feb 09;9(2):180 [FREE Full text] [doi: 10.3390/healthcare9020180] [Medline: 33572072]
66. Niki K, Okamoto Y, Maeda I, Ueda M. Responses to Kako et al. (DOI: 10.1089/jpm.2019.0072) and Niki et al. (DOI: 10.1089/jpm.2018.0233): A Novel Palliative Care Approach Using Virtual Reality for Improving Various Symptoms of Terminal Cancer Patients: A Preliminary Prospective, Multicenter Study. J Palliat Med 2019 Dec;22(12):1490. [doi: 10.1089/jpm.2019.0487] [Medline: 31794368]
67. Pla-Sanjuanelo J, Ferrer-Garcia M, Gutiérrez-Maldonado J, Vilalta-Abella F, Andreu-Gracia A, Dakanalis A, et al. Trait and state craving as indicators of validity of VR-based software for binge eating treatment. In: Volume 219: Annual Review of Cybertherapy and Telemedicine 2015. Amsterdam: IOS Press Ebooks; 2015:141-146.
68. Ferrer-Garcia M, Gutiérrez-Maldonado J, Agliaro-López M, Lobera-Espi X, Pla J, Vilalta-Abella F. Validation of VR-based Software for Binge Eating Treatment: Preliminary Data. In: Volume 199: Annual Review of Cybertherapy and Telemedicine 2014. Amsterdam: IOS Press Ebooks; 2014:146-150.
69. Szpak A, Michalski SC, Loetscher T. Exergaming With Beat Saber: An Investigation of Virtual Reality Aftereffects. J Med Internet Res 2020 Oct 23;22(10):e19840 [FREE Full text] [doi: 10.2196/19840] [Medline: 33095182]
70. Pettijohn KA, Pistone DV, Warner AL, Roush GJ, Biggs AT. Postural Instability and Seasickness in a Motion-Based Shooting Simulation. Aerosp Med Hum Perform 2020 Sep 01;91(9):703-709. [doi: 10.3357/amhp.5539.2020]
71. Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial Intelligence in Health Care: Bibliometric Analysis. J Med Internet Res 2020 Jul 29;22(7):e18228 [FREE Full text] [doi: 10.2196/18228] [Medline: 32723713]
72. Şenel E, Demir E. A global productivity and bibliometric analysis of telemedicine and teledermatology publication trends during 1980–2013. Dermatologica Sinica 2015 Mar;33(1):16-20. [doi: 10.1016/j.dsi.2014.10.003]
73. Savage N. The race to the top among the world's leaders in artificial intelligence. Nature 2020 Dec 09;588(7837):S102-S104. [doi: 10.1038/d41586-020-03409-8]
74. Martinerie L, Rasoaherinomenjanahary F, Ronot M, Fournier P, Dousset B, Tesnière A, et al. Health Care Simulation in Developing Countries and Low-Resource Situations. J Contin Educ Health Prof 2018;38(3):205-212. [doi: 10.1097/ceh.0000000000000211]
75. Reis RS, Salvo D, Ogilvie D, Lambert EV, Goenka S, Brownson RC, Lancet Physical Activity Series 2 Executive Committee. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. Lancet 2016 Sep 24;388(10051):1337-1348 [FREE Full text] [doi: 10.1016/S0140-6736(16)30728-0] [Medline: 27475273]
76. Wehden L, Reer F, Janzik R, Tang WY, Quandt T. The Slippery Path to Total Presence: How Omnidirectional Virtual Reality Treadmills Influence the Gaming Experience. MaC 2021 Jan 06;9(1):5-16. [doi: 10.17645/mac.v9i1.3170]
77. Breitkreuz KR, Kardong-Edgren S, Gilbert GE, DeBlieck C, Maske M, Hallock C, et al. A multi-site study examining the usability of a virtual reality game designed to improve retention of sterile catheterization skills in nursing students. Simulation & Gaming 2020 Sep 25;52(2):169-184. [doi: 10.1177/1046878120954891]
78. Riva G. Virtual reality for health care: the status of research. Cyberpsychol Behav 2002 Jun;5(3):219-225. [doi: 10.1089/109493102760147213] [Medline: 12123244]
79. Lewis CH, Griffin MJ. Human factors consideration in clinical applications of virtual reality. In: Volume 44: Virtual Reality in Neuro-Psycho-Physiology. Amsterdam: IOS Press Ebooks; 1997:35-56.
80. Custură-Crăciun D, Cochior D, Constantinoiu S, Neagu C. Surgical virtual reality - highlights in developing a high performance surgical haptic device. Chirurgia (Bucur) 2013;108(6):757-763 [FREE Full text] [Medline: 24331310]

XSL•FO
RenderX

81. Graur F. Virtual Reality in Medicine - Going Beyond the Limits. In: Lányi CS, editor. The Thousand Faces of Virtual Reality. London: IntechOpen; 2014.
82. Turolla A, Dam M, Ventura L, Tonin P, Agostini M, Zucconi C, et al. Virtual reality for the rehabilitation of the upper limb motor function after stroke: a prospective controlled trial. J Neuroeng Rehabil 2013 Aug 01;10:85 [FREE Full text] [doi: 10.1186/1743-0003-10-85] [Medline: 23914733]
83. Mazurek J, Kiper P, Cieślik B, Rutkowski S, Mehlich K, Turolla A, et al. Virtual reality in medicine: a brief overview and future research directions. hm 2019;20(3):16-22. [doi: 10.5114/hm.2019.83529]
84. Moher D, Schulz KF, Altman D, CONSORT Group (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA 2001 Apr 18;285(15):1987-1991. [doi: 10.1001/jama.285.15.1987] [Medline: 11308435]

Abbreviations

AR: augmented reality
VR: virtual reality

Edited by N Zary; submitted 07.08.21; peer-reviewed by B Chaudhry, J Gutiérrez-Maldonado; comments to author 26.08.21; revised version received 20.09.21; accepted 24.09.21; published 01.12.21.

Please cite as:
Pawassar CM, Tiberius V
Virtual Reality in Health Care: Bibliometric Analysis
JMIR Serious Games 2021;9(4):e32721
URL: https://games.jmir.org/2021/4/e32721
doi:10.2196/32721
PMID:34855606

©Christian Matthias Pawassar, Victor Tiberius. Originally published in JMIR Serious Games (https://games.jmir.org), 01.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using the Behaviour Change Wheel Program Planning Model to Design Games for Health: Development Study

Michael C Robertson[1], PhD, MPH; Tom Baranowski[2], PhD; Debbe Thompson[2], PhD, RD; Karen M Basen-Engquist[3], PhD, MPH; Maria Chang Swartz[4], PhD, MPH, RD, LD; Elizabeth J Lyons[1], PhD, MPH

[1]Department of Nutrition, Metabolism & Rehabilitation Sciences, University of Texas Medical Branch at Galveston, Galveston, TX, United States

[2]U.S. Department of Agriculture/Agricultural Research Service Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX, United States

[3]Department of Behavioral Science, University of Texas MD Anderson Cancer Center, Houston, TX, United States

[4]Department of Pediatrics-Research, University of Texas MD Anderson Cancer Center, Houston, TX, United States

Corresponding Author:
Elizabeth J Lyons, PhD, MPH
Department of Nutrition, Metabolism & Rehabilitation Sciences
University of Texas Medical Branch at Galveston
301 University Blvd
Galveston, TX, 77555
United States
Phone: 1 409 772 2578
Email: ellyons@utmb.edu

Abstract

Background: Games for health are a promising approach to health promotion. Their success depends on achieving both experiential (game) and instrumental (health) objectives. There is little to guide game for health (G4H) designers in integrating the science of behavior change with the art of game design.

Objective: The aim of this study is to extend the Behaviour Change Wheel program planning model to develop Challenges for Healthy Aging: Leveraging Limits for Engaging Networked Game-Based Exercise (CHALLENGE), a G4H centered on increasing physical activity in insufficiently active older women.

Methods: We present and apply the G4H Mechanics, Experiences, and Change (MECHA) process, which supplements the Behaviour Change Wheel program planning model. The additional steps are centered on identifying target G4H player experiences and corresponding game mechanics to help game designers integrate design elements and G4H objectives into behavioral interventions.

Results: We identified a target behavior of increasing moderate-intensity walking among insufficiently active older women and key psychosocial determinants of this behavior from self-determination theory (eg, autonomy). We used MECHA to map these constructs to intervention functions (eg, persuasion) and G4H target player experiences (eg, captivation). Next, we identified behavior change techniques (eg, framing or reframing) and specific game mechanics (eg, transforming) to help realize intervention functions and elicit targeted player experiences.

Conclusions: MECHA can help researchers map specific linkages between distal intervention objectives and more proximal game design mechanics in games for health. This can facilitate G4H program planning, evaluation, and clearer scientific communication.

(JMIR Serious Games 2021;9(4):e29964) doi:10.2196/29964

KEYWORDS

physical activity; video games; eHealth; intervention; behavior and behavior mechanisms; psychological theory; serious games; gamification; older women; older adults; behavior change; behavioral interventions; mobile phone

XSL•FO
RenderX

Introduction

Background

Games for health (Gs4H) are a promising venue through which to increase motivation for, and enjoyment associated with, health-related behaviors. Its success is largely dependent on eliciting a playful, enjoyable, fun experience while simultaneously delivering efficacious behavior change techniques (BCTs) [1]. Gs4H seem to effect change in various contexts [2-5]. However, existing studies provide limited guidance for developers interested in designing novel Gs4H. Studies often have confusing or sparse discussions of their game design elements, and specific game mechanics are not often linked to theory-based behavior change mechanisms. This lack of clarity limits the accumulation of scientific knowledge and may be due in part to a lack of a common program planning method.

The extent to which a game for health (G4H) brings about behavior change defines its success. G4H developers have long recognized the potential utility of using insights from the field of behavioral science to achieve health-related behavior change [6]. BCTs have become a popular choice for parsing the *active ingredients* of behavioral interventions because they cut across diverse behavior change theories and help to explicate mechanisms of behavior change. Gs4H may be a particularly useful avenue through which to implement BCTs because they may be amenable to doing so with high fidelity to large numbers of people [7]. Explicitly including BCTs in G4H design will help G4H designers to identify the most effective strategies for fostering behavior change in various contexts and articulate hypothesized mechanisms of behavior change [7-9].

Detailed program planning methods exist for traditional health promotion programs. Systematic methods such as Intervention Mapping and the Behaviour Change Wheel (BCW) emphasize using behavior change theories and techniques to link modifiable determinants of behavior to health-related behaviors and outcomes [10-12]. However, these approaches do not readily extend to the intricacies of game design. For example, application of the BCW might indicate that the use of a BCT such as *framing* or *reframing* is warranted to influence health-related behavior [13], but researchers have little guidance on *how* to implement this BCT using game design. Research centered on using program planning methods in game design may facilitate more rigorous development and clear scientific communication of Gs4H.

Models of gamification or using game design elements to achieve nongame objectives emphasize the importance of designing game systems that are chiefly centered on the benefits to, and interests of, the user [14-16]. They highlight the importance of iterative prototyping and the potential utility of self-determination theory (SDT)—a health behavior change theory that posits that the satisfaction of one's core psychological needs can have a bearing on one's quality of motivation for engaging in health-related behavior [17,18]. The Mechanics, Dynamics, and Aesthetics Framework provides a formal approach to understanding games by helping designers to decompose games into their constituent design elements and map their interrelationships [18]. The Learning Mechanics-Game Mechanics Model enables researchers to map pedagogical principles featured in educational games to corresponding game design elements [19]. Existing models provide useful nuance but may not match the particular needs of behavioral health interventionists because they are not well suited for explicating mechanisms of behavior change.

Most Gs4H do not seem to be grounded in models of gamification or health behavior change theory [20]. Furthermore, although BCTs are commonly found in Gs4H (typically falling into the categories of feedback and monitoring, comparison of behavior, and reward and threat), their putative role in the experiences evoked by Gs4H are seldom specified [21]. Thus, at present there are not strong linkages between the core tenets of health behavior change theory and the experiential objectives of Gs4H (eg, fun or playfulness) [22]. This gap stymies progress in G4H design because it limits our ability to integrate key experiential and instrumental objectives and the precision with which we are able to evaluate the pathways through which Gs4H may effect change.

Step-by-Step Program Planning Model

The development of Gs4H typically demands considerable up-front costs [9], and potential funders need to see how game design elements purport to effect health-related outcomes [23]. Thus, it is especially important to ensure that key scientific and game design principles are adequately integrated early in design. The BCW specifies commonly accepted BCTs and how to select them for program design [11,12]. We present a step-by-step program planning model adapting the BCW [11,12] to the development of Challenges for Healthy Aging: Leveraging Limits for Engaging Networked Game-Based Exercise (CHALLENGE). CHALLENGE is a social media game, delivered through Facebook and centered on increasing physical activity in insufficiently active older women. This paper thereby illustrates a process to guide researchers in the development of Gs4H and also provides a template to facilitate clearer scientific communication for linking behavior change objectives to G4H design.

Methods

Overview

To develop the behavioral intervention core of CHALLENGE, we adapted the BCW program planning process enumerated by Michie et al [12] by adding 2 steps and removing 1. The first added step, termed *Identify target player experiences*, is centered on selecting target experiences that are conducive to intervention functions and influencing psychosocial constructs. The second added step, termed *Identify game mechanics*, is centered on selecting specific game design elements for eliciting target player experiences and implementing BCTs. We removed the BCW step centered on identifying policy categories because Gs4H are typically concerned with targeting intrapersonal- and interpersonal-level factors. Taken together, we call this G4H planning process the Mechanics, Experiences, and Change (MECHA) Model (Figure 1). It consists of 9 steps: (1) define the problem in behavioral terms, (2) select the target behavior, (3) specify the target behavior, (4) identify what needs to

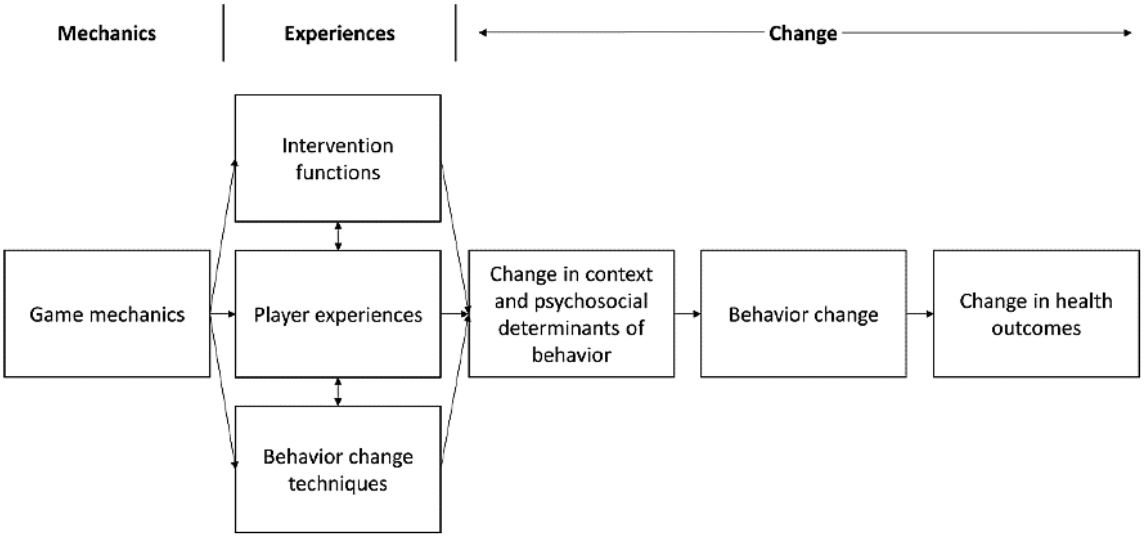
XSL•FO
RenderX

change, (5) identify intervention functions, (6) identify target player experiences, (7) select BCTs, (8) identify game mechanics, and (9) determine mode of delivery (Table 1).

Study Approval

This study was approved by the institutional review board of The University of Texas Medical Branch at Galveston (protocol number: 19-0158).

Figure 1. A Mechanics, Experiences, and Change Model for game-induced behavior change.



XSL•FO
RenderX

Table 1. Steps of game for health Mechanics, Experiences, and Change Model.

Step	Description	Example
Step 1: Define the problem in behavioral terms[a]	Identify the specific target population. Review the epidemiological evidence concerning the health-related outcomes of interest. Identify relevant behaviors linked to these outcomes and their location in that target population	We identified the target population as women aged 65-85 years who are not meeting nationally recommended physical activity guidelines
Step 2: Select a target behavior[a]	Select a target behavior from the relevant behaviors identified in step 1. Consider the relative impact of each behavior, its likelihood of change, the potential for spillover into other important behaviors, and ease of measurement of the behavior	The nationally recommended physical activity guidelines for older adults include several related behaviors (eg, aerobic physical activity, muscle-strengthening physical activity, and time spent sedentary) [24]. In this step we opted to focus on aerobic physical activity because of its well-documented health benefits, preference in older adults [25], and amenability to objective measurement
Step 3: Specify the target behavior[a]	Specify the target behavior identified in step 2 in detail. To do so, use the 6 template questions proposed by Michie et al [12]: (1) Who needs to perform the behavior? (2) What does the person need to do differently to achieve the desired change? (3) When will they do it? (4) Where will they do it? (5) How often will they do it? (6) With whom will they do it?	We specified details pertaining to the target behavior selected in step 2 (presented in the *Results* section)
Step 4: Identify what needs to change[a]	Conduct a behavioral analysis as recommended by Michie et al [12] (ie, focus groups, questionnaires, observations, documentary analysis, or review of relevant literature) to identify psychosocial constructs that need to change to achieve the desired behavior change. Select an appropriate health-related behavior change theory or model to guide this process	We identified key constructs that need to change to promote the target behavior identified in step 3 (the *Results* section provides an illustrative template). This process may be guided by, for example, SDT[b] [17,18] or the COM-B[c] Model [12]. For example, *competence* for engaging in physical activity predicts physical activity levels across a wide range of contexts [26]
Step 5: Identify intervention functions[a]	Identify the primary intervention functions of the G4H[d] intervention to target each of the constructs identified in step 4. The process of mapping intervention functions to psychosocial constructs can be guided by the BCW[e] [12] and its subcomponents [27]	*Modeling*, for example, is an intervention function that may be used to increase feelings of competence for engaging in physical activity (a psychosocial construct featured in SDT)
Step 6. Identify target player experiences	Identify experiential objectives that would facilitate compelling gameplay and can be integrated with the intervention functions (identified in step 5). Behavior change theory can facilitate this process	We used the Playful Experiences Framework to frame the identification of target player experiences [28,29]. As an example, SDT suggests that experiences of friendly *competition* with appropriately matched role models may bolster feelings of *competence* for engaging in physical activity
Step 7: Identify BCTs[a,f]	Select the BCTs to be featured in the intervention. BCTs are the "active ingredients" of a behavioral intervention, or "the observable, replicable, irreducible components of an intervention" designed to change behavior (eg, the provision of feedback). The taxonomy of 93 distinct BCTs developed by Michie et al [13] can be used to frame the selection of BCTs. Review the extant scientific literature to identify BCTs that may be most effective in the context of the target behavior and target population	Existing literature suggests that prompting self-monitoring of behavior, as an example, may be particularly useful for helping older adults increase physical activity [30]
Step 8: Identify game mechanics	Identify specific game mechanics designed to evoke target participant experiences [15,31]. Select a taxonomy of game mechanics to guide G4H game design [32-37]. Map individual game mechanics to the player experiences, BCTs, and psychosocial constructs identified in previous steps. These conceptual links can be evaluated and refined during iterative game development [15,31]	As an example, participants' physical activity performance (game mechanics) [33] may be used to evoke experiences of *competition* and implement the BCT of *social comparison*. These processes can be oriented toward increasing participants' feelings of *competence* for engaging in physical activity
Step 9: Identify mode of delivery[a]	Identify modes of delivery of the intervention that would be appropriate for the target population. These decisions can be informed by formative research and the scientific literature	For example, previous research may suggest that older adults tend to prefer computer-based Gs4H to those delivered through mobile devices

[a]Michie et al [12] and Michie et al [13].

[b]SDT: self-determination theory.

[c]COM-B: Capability-Opportunity-Motivation Behavior.

[d]G4H: game for health.

[e]BCW: Behaviour Change Wheel.

[f]BCT: behavior change technique.

XSL•FO
RenderX

Results

Application of MECHA

We applied MECHA to develop CHALLENGE. In doing so, we sought to refine our understanding of the context of the behavior in the target population and mapped the hypothesized linkages between health behavior change theory constructs, intervention functions, target participant experiences, BCTs, and game mechanics.

Step 1: Define the Problem in Behavioral Terms

Sustained moderate-to-vigorous intensity physical activity is beneficial for older adults [38]. Most older adults do not meet recommended levels of physical activity [39], and physical activity levels tend to decline with age [40,41]. Women tend to be less physically active than men [42-45] and can face decreased lean muscle mass after menopause [46] that can exacerbate the negative effects of inadequate activity [47]. Accordingly, the target population for this G4H includes being female, aged 65-85 years, owning a smartphone and having regular internet access, self-reported weekly moderate-to-vigorous intensity physical activity ≥150 minutes, and evidence suggesting that engaging in moderate-intensity physical activity would be safe (eg, answering *no* to all Physical Activity Readiness Questionnaire for Everyone items or having a physician's note).

Step 2: Select a Target Behavior

Moderate-intensity (ie, 3.0-6.0 metabolic equivalents) brisk walking has consistently emerged as a physical activity preference of older adults and can satisfy the nationally recommended aerobic physical activity guideline of engaging in at least 150 minutes of moderate to vigorous intensity physical activity per week [24,25]. Thus, we designed CHALLENGE to be primarily centered on increasing moderate-intensity walking.

Step 3: Specify the Target Behavior

Keeping in mind the existing literature norms for older adults and findings from pilot study research [48], we selected a default step goal recommendation of walking at least 8000 steps per day for at least 5 days per week (Textbox 1).

Textbox 1. Specify the target behavior.

Specify the target behavior identified in detail by using the 6 template questions proposed by Michie et al [12]

- Target behavior
- Walking at least 8000 steps per day for at least 5 days per week
- Who needs to perform the behavior?
- Women aged 65-85 years in southeast Texas who are not meeting nationally recommended aerobic physical activity guidelines
- What do they need to do differently to achieve the desired change?
- Increase walking, both as a lifestyle and for exercise. Increases should occur gradually with a target of ≥1000 daily steps per week until target goals are met
- When do they need to do it?
- Daily
- Where do they need to do it?
- Outdoors if possible, indoors at large venues, or at home
- How often do they need to do it?
- Daily
- With whom do they need to do it?
- Alone or in small groups

Step 4: Identify What Needs to Change

We chose to ground the intervention in SDT [17,18] because it has led to better understanding of physical activity patterns [26] and effective game design [49,50]. We selected 2 constructs from SDT as the primary targets of CHALLENGE: intrinsic regulation and integrated regulation. Rather than focus on intrinsic regulation per se, we opted to focus on the basic psychological needs that are theorized to predict this form of motivation: perceived autonomy, competence, and relatedness. Although these psychological needs also predict integrated regulation, we focused on the construct separately to account for its emphasis on personal values and identity (Table 2).

XSL•FO
RenderX

Table 2. Identify what needs to change[a].

Theoretical constructs	Requirements for the target behavior to occur	Should the intervention target this construct?
Perceived autonomy	Participant wants to engage in physical activity for autonomous reasons (ie, enjoyment, interest, identity, and values) [17,18,51]	Yes; autonomous motivations for physical activity predict long-term compliance to physical activity goals, and older women want autonomy-promoting interventions [26,52-54]
Perceived competence	Participant feels competent and able to engage in physical activity	Yes; self-efficacy is a strong predictor of physical activity, and many older women report low levels of self-efficacy for consistently meeting nationally recommended physical activity guidelines [55,56]
Perceived relatedness	Participant feels supported by others regarding her physical activity	Yes; social support is a strong predictor of physical activity, and older women want social physical activity interventions [55,57]
Intrinsic regulation	Participant perceives physical activity as fun and interesting	Yes; previous studies suggest that these factors predict physical activity, and this is an identified barrier in this population [54,55]
Integrated regulation	Participant perceives physical activity as being in line with her values and identity	Yes; previous studies suggest that integrated regulation predicts physical activity in this population [53-55,58]
Identified regulation	Participant perceives physical activity as associated with an outcome that is important to her	Yes; previous studies suggest that identified regulation predicts physical activity in this population [53-55,58]
Introjected regulation	Participant feels obligated to engage in physical activity	No; although in some cases this type of motivation may lead to behavior initiation, it is not conducive to long-term adherence to physical activity [26]
External regulation	Participant perceives physical activity as something that outside forces are encouraging her to do	No; although in some cases this type of motivation may lead to behavior initiation, it is not conducive to long-term adherence to physical activity [26]

[a]Target behavior: walking at least 8000 steps per day for at least 5 days per week.

Step 5: Identify Intervention Functions

We identified 6 functions of our intervention to bring about change in the key psychological needs enumerated in Table 2. These were persuasion, incentivization, environmental restructuring, modeling, training, and enablement (Multimedia Appendix 1 [11-13,17,18,28,29,33]). We arrived upon these intervention functions by first reviewing the various intervention functions outlined by Michie et al [11,12]. We drew from this list of potential intervention functions in an iterative process as we proceeded through the remaining steps of MECHA to map our G4H intervention. In doing so, we paired specific intervention functions with specific target playful experiences to ensure that our intervention's featured elements of game design would be integrated with these core tenets of behavior change.

Step 6: Identify Target Player Experiences

We selected the Playful Experiences Framework to guide our identification of target player experiences [28,29]. This framework was derived from studies of electronic art and video games and was created to help designers make more engaging and attractive participative systems. Using this framework, we selected target experiences for bolstering participants' psychological needs in accordance with SDT tenets (Multimedia Appendix 1). For example, because an individual's sense of autonomy is largely derived from their feelings of volitional control and authentic endorsement of their own behavior [17,18], we sought to evoke sensations of captivation, discovery, exploration, expression, and humor in the course of their gameplay. Note that the connections we make between participant experiences and SDT constructs are not necessarily exclusive; they are those that we believe are the strongest links conceptually, but it is likely that player experiences could target multiple theoretical constructs.

Step 7: Identify BCTs

We identified 13 BCTs to increase moderate-intensity walking in CHALLENGE by pairing the target player experiences identified in step 6 with the taxonomy of 93 distinct BCTs described by Michie et al [13] (Multimedia Appendix 1). These were as follows: adding objects to the environment, framing or reframing, goal setting for behavior, graded tasks, identity associated with changed behavior, information about others' approval, nonspecific reward, prompts or cues, reducing negative emotions, self-monitoring of behavior, self-talk, social comparison, and social reward.

Step 8: Identify Game Mechanics

We selected the library of game mechanics developed by Järvinen [33] to guide our intervention design (1) for its balance of complexity and parsimony and (2) because we felt that this conceptualization of game mechanics, defined as "what the player does in relation to the game state during a standard turn or sequence," would lend itself to autonomy-supportive, meaningful game design. We identified game mechanics from this library to elicit targeted player experiences and implement BCTs (Multimedia Appendix 1). Note that these connections are not necessarily exclusive; the mapped connections are those that we believe are the strongest links conceptually, but it is likely that many of these elements of gameplay could engender varied player experiences.

XSL•FO
RenderX

Step 9: Identify Mode of Delivery

We identified Facebook as an appropriate platform through which to deliver this intervention [59,60]. We also decided to provide participants with digital physical activity trackers. Other studies have shown high acceptability of these technologies for delivering physical activity interventions to insufficiently active older adults [61,62].

Developed G4H

Table 3 presents details of the behavioral intervention using the Template for Intervention Description and Replication Checklist [63]. Participants receive weekly challenges for 1 year. Challenges are centered on implementing the game mechanics identified previously (Multimedia Appendix 1; example challenges). At baseline, participants receive materials that can be used for some challenges throughout their participation in the study (eg, description cards for scavenger hunt–style challenges and cardboard frames, as well as masks and similar props that allow participants to obscure their faces in photographs). Weekly challenges are delivered by trained moderators, who post in a private Facebook group (Table 3; moderators are research staff trained by the principal investigator [EJL] on basic aspects of the Playful Experiences Framework [28,29] and the intended mechanisms of change of the intervention). Participants are encouraged to comment in response to the weekly challenges and interact with each other in this forum at their convenience throughout the week. All challenges were presented to a sample of individuals from the target population (n=20) in individual interviews, and we conducted an in-depth, qualitative evaluation of the interview transcripts and revised the intervention content according to participant feedback (manuscript in preparation).

Table 3. Template for Intervention Description and Replication Checklist for Challenges for Healthy Aging: Leveraging Limits for Engaging Networked Game-Based Exercise (CHALLENGE).

Item name	Item
Brief name	CHALLENGE
Why	Despite short-term benefits, older adults' adherence to physical activity and tracker use decrease sharply over time. Most existing intervention systems use a corrective frame: they are oriented toward *fixing* undesirable behaviors. Emerging theoretical frameworks indicate that this approach is unlikely to produce sustained behavior change. Instead, taking a celebratory approach that facilitates enjoyable and valued aspects of behavior may better promote longer-term adherence to physical activity. Games for health are conducive to this approach to physical activity promotion
What (materials)	Participants are provided a wrist-worn electronic physical activity tracker (Fitbit Inspire 2 [Google LLC]) and various props (eg, scavenger hunt bingo cards and sunglasses or masks to help obscure their identities). If participants do not already have Facebook and Fitbit accounts, study staff help the participants to create them
What (procedures)	Participants meet face-to-face with study staff for orientation procedures (eg, aiding with technology) and data collection. Participants engage in goal setting and action planning with study staff at baseline and are invited to join a private Facebook group. Through this private Facebook group, participants receive weekly challenges that are centered on encouraging walking behaviors and eliciting playful experiences (see examples in Multimedia Appendix 1). Participants are encouraged to directly respond to challenges through Facebook posts and like or comment on others' posts. Participants also receive weekly messages providing feedback on their physical activity levels and study engagement (ie, number of times participants posted in the Facebook group)
Who provided	Interventionists are trained by the principal investigator (EJL) on basic aspects of the Playful Experiences Framework [28,29] and the intended mechanisms of change of the intervention. Moderators are also trained by a key collaborator (MCS) who oversaw formative research for this intervention
How	Goal setting and action planning are conducted face-to-face or through videoconferencing at the start of the study. All other intervention content is delivered on the web. Challenges are posted weekly using social media to a single, private Facebook group. Participants also receive individual weekly emails presenting their device-measured physical activity levels, suggested goals, and engagement level
Where	Face-to-face meetings and data collection sessions are held at a large medical research university in southeast Texas. The intervention content is largely delivered through the internet
When and how much	Intervention content is sent weekly over the course of 1 year for participants, with participants being enrolled on a rolling basis until the target sample size is reached (estimated to be 2-3 years). Recruitment began in June 2021
Tailoring	At the beginning of the intervention, study staff meet with participants to establish physical activity goals (ie, target step count and number of days per week that participants aim to meet that target step count). In this meeting, participants also select their target weekly improvement rate (eg, participants may indicate that if they did not meet their target step count one week, then they would like their goal for the next week to be to increase their daily average step count by 1000). Weekly emails accordingly present feedback on the previous week and provide a suggested step count goal for the upcoming week
How well (planned)	Moderators' weekly posts are based on a set schedule and accompanying scripts. The only communication with participants that is not heavily based on scripts are responses to direct messages or SMS text messages sent regarding scheduling, reporting unacceptable content, and so on. We will extract information from the Facebook group and Fitbit app regularly to track participant engagement

XSL•FO
RenderX

Human involvement in the G4H intervention is conducted by moderators, who facilitate all web-based proceedings. They post all weekly challenges with example responses, review the Facebook group for adverse events or inappropriate comments, provide supportive and clarifying comments to study participants as necessary, and publicly recognize consistent participation (eg, badges awarded to *power users*). Moderators also send weekly emails to participants with basic weekly feedback on their device-measured step count data. Trained interventionists perform face-to-face follow-up data collection procedures at 6, 12 (intervention end), and 18 months after baseline. After intervention end, participants can elect to stay in the Facebook group if they wish (to facilitate an active, supportive environment). Participants are compensated for their time with a US $25 gift card at each of the 3 follow-up appointments.

To evaluate the G4H, we aim to recruit 300 participants reflective of the target population detailed in step 1. We will recruit participants on a rolling basis by using several strategies, including in-person recruitment at gerontology and primary care clinics, web-based recruitment methods, and flyers and brochures placed at locations frequented by members of the target population. Outcome measures for the developed G4H include objectively measured step count and moderate to vigorous physical activity levels [64], as well as measures of physical functioning (eg, 6-minute walk test [65,66]). We will also investigate potentially mediating variables (eg, the psychosocial constructs presented in Table 2 [67-69]), the role of process variables such as those reflecting participant engagement (eg, participants' weekly number of posts, comments, reactions, days engaged, and total number of engagements), and the degree to which the experimental group, compared with the comparison group, experienced the targeted playful experiences over the course of the study period [29]. We will investigate predictors of participant engagement (eg, age and fitness level) and how engagement is associated with study outcomes.

Discussion

Principal Findings

We extended the BCW to create MECHA, a step-by-step program planning model for designing Gs4H, and applied it to the development of a behavioral intervention centered on increasing physical activity in insufficiently active older women. MECHA may help researchers map specific linkages between distal instrumental and experiential objectives to the more proximal elements of game design. This may facilitate program planning, evaluation, and clearer communication of results to expedite scientific accumulation of knowledge.

Comparison With Prior Work

MECHA shares similarities with other models that can help researchers parse G4H game design elements. The Mechanics, Dynamics, and Aesthetics Framework emphasizes the importance of mapping game mechanics to targeted emotional responses in game design [31]. This is analogous to the process of mapping game mechanics to targeted player experiences presented in this study. The Mechanics, Dynamics, and Aesthetics Framework can supplement this process for G4H designers. The Learning Mechanics-Game Mechanics Model [19] is centered on mapping game mechanics to specific educational objectives when designing games for educational purposes. This is somewhat analogous to the process of mapping game mechanics to BCTs that we encourage in this study. These models and the research derived from them underscore the utility of mapping game mechanics to desired objectives for intervention development and evaluation.

We used the library of primary game mechanics developed by Järvinen [33], but other taxonomies of game mechanics may be useful for different Gs4H [32-37]. Game mechanics interact with one another and may have different effects at different times and different effects on different users; because of this inherent complexity, it is unlikely that a single, definitive catalog of game mechanics will emerge [32,70]. The conceptualization of game mechanics described by Järvinen [33] comports with the theoretical underpinnings and context of our G4H. Other taxonomies range from the conceptualization of game mechanics described by Schell [32] as consisting of the 7 essential elements of space, time, objects, actions, rules, skill, and chance to the extensive list of hundreds of game design patterns compiled by Bjork and Holopainen [71]. Designers should select from among existing taxonomies of game mechanics the one that provides the greatest utility for integrating the results of the preceding steps presented in this paper.

Evoking enjoyable participant experiences is likely critical to sustained adherence to Gs4H and their associated health-related behaviors but designing games that do so reliably is a challenge. Consumption of games, as opposed to consumption of books, movies, and so on, is inherently nonlinear, and this introduces some uncertainty in predicting player experience [31,32]. A key to honing the desired qualities of emergent gameplay is iterative development with frequent input from the priority population [15,31]. Semistructured qualitative interviews with participants who experience the G4H can be a useful tool for investigating to what degree participants are indeed undergoing the targeted experiences. Quantitative data may further help researchers divine participants' experiences. The Playful Experiences Questionnaire, for example, measures the categories promulgated in the Playful Experiences Framework [28,29]. Research with mixed methods designs that allow qualitative and quantitative data to build upon one another may lead to a deeper level of understanding.

We identified areas of needed research while conducting this project. First, more research is needed to elucidate how game experiences may affect SDT constructs. Greater clarity regarding if and how different experiences afforded by Gs4H affect critical psychosocial constructs may help game designers to develop more efficacious Gs4H. Second, more research that explicates how different G4H game mechanics engender specific player experiences is needed. Although this is likely to remain the purview of experts because of its inherent complexity, literature that helps to frame these links may help researchers to systematize iterative game development and communicate research proposals and study findings. Third, engagement is a key issue in securing a person's participation in, and thereby exposure to, a game. Thus, engagement is critical for G4H effectiveness. Engagement likely hinges on participants'

XSL•FO
RenderX

G4H-related experiences—both experiences stemming from the game and the health-promoting aspects of the G4H. More research is needed to characterize how specific G4H-related experiences correspond to effective engagement with the G4H [72,73].

Strengths and Limitations

The strengths of this study include adherence to recommended procedures for increasing transparency of scientific research (eg, the Template for Intervention Description and Replication Checklist), a systematic approach to game design, and the real-world application of the G4H development process to create a behavioral intervention. The limitations of this study include, first, that we did not conduct a formal systematic review of the literature. As G4H design exists at the confluence of several fields, this was not within the scope of this study. Although our study team has considerable expertise in G4H research, there may be additional relevant studies not covered in this paper. Furthermore, our study team did not include professional game designers or developers. We made extensive use of relevant scientific literature, but future research would benefit from the inclusion of individuals with expertise in these areas. The MECHA model is a useful starting point, but it would likely be improved by incorporating the unique insights that game designers and experts in human-computer interaction may provide. This may be especially useful as MECHA is applied to help design health-promoting video games.

Conclusions

G4H design combines the art of game design and the science of behavior change. In this paper, we have presented a process for systematically integrating these perspectives and illustrated its use in the design of a G4H centered on increasing physical activity in insufficiently active older adults. This systematic approach to G4H design may facilitate program planning, evaluation, and clearer communication of G4H interventions.

Acknowledgments

This study was supported in part by the Center for Energy Balance in Cancer Prevention and Survivorship, which is supported by the Duncan Family Institute for Cancer Prevention and Risk Assessment. This study was funded by an award from the National Institute on Aging (R01AG064092). This study was also supported by the Claude D Pepper Older Americans Independence Center (P30AG024832), The University of Texas Medical Branch at Galveston Clinical and Translational Science Award (TR001439), and the Sealy Center on Aging. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health. Additional salary support for EJL was provided by a Mentored Research Scholar Grant in Applied and Clinical Research (MRSG-14-165-01-CPPB) from the American Cancer Society and by a grant from the National Cancer Institute (R21CA218543).

Authors' Contributions

EJL and MCR conceived the project and developed the behavioral intervention with support from MCS, TB, DT, and KMBE. MCR wrote the manuscript with support from EJL, TB, DT, KMBE, and MCS.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mapping self-determination theory constructs, intervention functions, target participant experiences, behavior change techniques, and game mechanics.
[DOCX File , 19 KB - games_v9i4e29964_app1.docx]

References

1. Baranowski MT, Lieberman PD, Buday R, Peng W, Zimmerli L, Wiederhold B, et al. Videogame mechanics in games for health. Games Health J 2013 Aug;2(4):194-204. [doi: 10.1089/g4h.2013.0617] [Medline: 26192223]
2. Papastergiou M. Exploring the potential of computer and video games for health and physical education: a literature review. Comput Educ 2009 Nov;53(3):603-622. [doi: 10.1016/j.compedu.2009.04.001]
3. Lyons EJ, Tate DF, Ward DS, Ribisl KM, Bowling JM, Kalyanaraman S. Engagement, enjoyment, and energy expenditure during active video game play. Health Psychol 2014 Feb;33(2):174-181 [FREE Full text] [doi: 10.1037/a0031947] [Medline: 23527520]
4. Kappen DL, Mirza-Babaei P, Nacke LE. Older adults' physical activity and exergames: a systematic review. Int J Human Comput Interac 2018 Feb 23;35(2):140-167. [doi: 10.1080/10447318.2018.1441253]
5. Epstein DS, Zemski A, Enticott J, Barton C. Tabletop board game elements and gamification interventions for health behavior change: realist review and proposal of a game design framework. JMIR Serious Games 2021 Mar 31;9(1):e23302 [FREE Full text] [doi: 10.2196/23302] [Medline: 33787502]

XSL•FO
RenderX

6. Thompson D, Baranowski T, Buday R, Baranowski J, Thompson V, Jago R, et al. Serious video games for health how behavioral science guided the development of a serious video game. Simul Gaming 2010 Aug 01;41(4):587-606 [FREE Full text] [doi: 10.1177/1046878108328087] [Medline: 20711522]
7. Baranowski T, Thompson D, Buday R, Lu AS, Baranowski J. Design of video games for children's diet and physical activity behavior change. Int J Comput Sci Sport 2010;9(2):3-17 [FREE Full text] [Medline: 25364331]
8. Baranowski T, Baranowski J, Thompson D, Buday R. Behavioral science in video games for children's diet and physical activity change: key research needs. J Diabetes Sci Technol 2011 Mar 01;5(2):229-233 [FREE Full text] [doi: 10.1177/193229681100500204] [Medline: 21527086]
9. Baranowski T, Buday R, Thompson D, Lyons EJ, Lu AS, Baranowski J. Developing games for health behavior change: getting started. Games Health J 2013 Aug;2(4):183-190 [FREE Full text] [doi: 10.1089/g4h.2013.0048] [Medline: 24443708]
10. Bartholomew L, Parcel G, Kok G, Gottlieb N, Schaalma H, Markham C, et al. Planning Health Promotion Programs: An Intervention Mapping Approach. San Francisco: Jossey-Bass; 2006.
11. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
12. Michie S, Atkins L, Gainforth H. Changing behaviour to improve clinical practice and policy. In: Novos Desafios Novas Competências Contrib Atuais Psicol Braga Axioma-Publicações Fac Filos: Published online; 2016:41-60.
13. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
14. Nicholson S. A recipe for meaningful gamification. In: Gamification in Education and Business. Cham: Springer; 2015.
15. Deterding S. The lens of intrinsic skill atoms: a method for gameful design. Human Comput Interac 2015 May 15;30(3-4):294-335. [doi: 10.1080/07370024.2014.993471]
16. Verschueren S, Buffel C, Vander Stichele G. Developing theory-driven, evidence-based serious games for health: framework based on research community insights. JMIR Serious Games 2019 May 02;7(2):e11565 [FREE Full text] [doi: 10.2196/11565] [Medline: 31045496]
17. 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]
18. Deci EL, Ryan RM. Self-determination theory: a macrotheory of human motivation, development, and health. Can Psychol/Psychologie canadienne 2008;49(3):182-185. [doi: 10.1037/a0012801]
19. Arnab S, Lim T, Carvalho M, Bellotti F, Freitas S, Louchart S, et al. Mapping learning and game mechanics for serious games analysis. Br J Educ Technol 2015:391-411 [FREE Full text] [doi: 10.1111/bjet.12113]
20. Lu AS, Kharrazi H. A state-of-the-art systematic content analysis of games for health. Games Health J 2018 Feb;7(1):1-15 [FREE Full text] [doi: 10.1089/g4h.2017.0095] [Medline: 29293368]
21. 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]
22. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
23. Baranowski T. Games for health research—past, present, and future. Präv Gesundheitsf 2018 Jul 12;13(4):333-336. [doi: 10.1007/s11553-018-0657-y]
24. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. JAMA 2018 Nov 20;320(19):2020-2028. [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
25. Amireault S, Baier JM, Spencer JR. Physical activity preferences among older adults: a systematic review. J Aging Phys Act 2018 Oct 25:1-12. [doi: 10.1123/japa.2017-0234] [Medline: 29283793]
26. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012;9:78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
27. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
28. Korhonen H, Montola M, Arrasvuori J. Understanding playful user experience through digital games. In: Proceedings of the International Conference on Designing Pleasurable Products and Interfaces. 2009 Presented at: International Conference on Designing Pleasurable Products and Interfaces; Oct 13-16, 2009; Compiègne, France.
29. Boberg M, Karapanos E, Holopainen J, Lucero A. PLEX-Q: towards a playful experiences questionnaire. In: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play. 2015 Presented at: CHI PLAY '15: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play; Oct 5 - 7, 2015; London. [doi: 10.1145/2793107.2793124]

XSL•FO
RenderX

30. French DP, Olander EK, Chisholm A, Mc Sharry J. Which behaviour change techniques are most effective at increasing older adults' self-efficacy and physical activity behaviour? A systematic review. Ann Behav Med 2014 Oct;48(2):225-234. [doi: 10.1007/s12160-014-9593-z] [Medline: 24648017]
31. Hunicke R, LeBlanc M, Zubek R. MDA: a formal approach to game design and game research. In: Proceedings of the AAAI Workshop on Challenges in Game AI. 2004 Presented at: Proceedings of the AAAI Workshop on Challenges in Game AI; 2004; San Jose URL: https://www.scirp.org/(S(351jmbntvnsjt1aadkozje))/reference/referencespapers.aspx?referenceid=2048676
32. Schell J. The Art of Game Design: A Book of Lenses. Third edition. Boca Raton: Taylor & Francis Group; 2020.
33. Järvinen A. Games Without Frontiers: Methods for Game Studies and Design. Tampere University: Tampere University Press; 2008.
34. Adams E, Dormans J. Game Mechanics: Advanced Game Design. Berkeley: New Riders; 2012.
35. Hervas R, Ruiz-Carrasco D, Mondejar T, Bravo J. Gamification mechanics for behavioral change: a systematic review and proposed taxonomy. In: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2017 Presented at: Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare; Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare; Barcelona, Spain. [doi: 10.1145/3154862.3154939]
36. Toda A, Oliveira W, Klock A, Palomino P, Pimenta M, Gasparini I, et al. A taxonomy of game elements for gamification in educational contexts: proposal and evaluation. In: Proceedings of the IEEE 19th International Conference on Advanced Learning Technologies (ICALT). 2019 Presented at: IEEE 19th International Conference on Advanced Learning Technologies; Jul 15-18, 2019; Maceio, Brazil. [doi: 10.1109/icalt.2019.00028]
37. Cechetti N, Biduki D, De Marchi A. Gamification strategies for mobile device applications: a systematic review. In: Proceedings of the 12th Iberian Conference on Information Systems and Technologies (CISTI). 2017 Presented at: 12th Iberian Conference on Information Systems and Technologies; Jun 21-24, 2017; Lisbon, Portugal. [doi: 10.23919/cisti.2017.7975943]
38. Taylor A, Cable N, Faulkner G, Hillsdon M, Narici M, Van Der Bij A. Physical activity and older adults: a review of health benefits and the effectiveness of interventions. J Sports Sci 2004 Aug;22(8):703-725. [doi: 10.1080/02640410410001712421] [Medline: 15370483]
39. Jefferis BJ, Sartini C, Lee I, Choi M, Amuzu A, Gutierrez C, et al. Adherence to physical activity guidelines in older adults, using objectively measured physical activity in a population-based study. BMC Public Health 2014 Apr 19;14:382 [FREE Full text] [doi: 10.1186/1471-2458-14-382] [Medline: 24745369]
40. Evenson K, Buchner D, Morland K. Objective measurement of physical activity and sedentary behavior among US adults aged 60 years or older. Prev Chronic Dis 2012;9:E26 [FREE Full text] [Medline: 22172193]
41. Clarke CL, Sniehotta FF, Vadiveloo T, Argo IS, Donnan PT, McMurdo ME, et al. Factors associated with change in objectively measured physical activity in older people - data from the physical activity cohort Scotland study. BMC Geriatr 2017 Aug 14;17(1):180 [FREE Full text] [doi: 10.1186/s12877-017-0578-1] [Medline: 28806930]
42. Gretebeck KA, Sabatini LM, Black DR, Gretebeck RJ. Physical activity, functional ability, and obesity in older adults: a gender difference. J Gerontol Nurs 2017 Sep 01;43(9):38-46. [doi: 10.3928/00989134-20170406-03] [Medline: 28399320]
43. Aoyagi Y, Shephard RJ. Sex differences in relationships between habitual physical activity and health in the elderly: practical implications for epidemiologists based on pedometer/accelerometer data from the Nakanojo Study. Arch Gerontol Geriatr 2013;56(2):327-338. [doi: 10.1016/j.archger.2012.11.006] [Medline: 23312568]
44. Straight C, Brady A, Evans E. Sex-specific relationships of physical activity, body composition, and muscle quality with lower-extremity physical function in older men and women. Menopause 2015 Mar;22(3):297-303. [doi: 10.1097/GME.0000000000000313] [Medline: 25137244]
45. Brugnara L, Murillo S, Novials A, Rojo-Martínez G, Soriguer F, Goday A, et al. Low physical activity and its association with diabetes and other cardiovascular risk factors: a nationwide, population-based study. PLoS One 2016 Aug 17;11(8):e0160959 [FREE Full text] [doi: 10.1371/journal.pone.0160959] [Medline: 27532610]
46. Hodson L, Harnden K, Banerjee R, Real B, Marinou K, Karpe F, et al. Lower resting and total energy expenditure in postmenopausal compared with premenopausal women matched for abdominal obesity. J Nutr Sci 2014 Feb 13;3:e3 [FREE Full text] [doi: 10.1017/jns.2013.38] [Medline: 25191611]
47. Ward-Ritacco C, Adrian A, Johnson M, Rogers L, Evans E. Adiposity, physical activity, and muscle quality are independently related to physical function performance in middle-aged postmenopausal women. Menopause 2014 Oct;21(10):1114-1121. [doi: 10.1097/GME.0000000000000225] [Medline: 24618768]
48. Tudor-Locke C, Craig CL, Aoyagi Y, Bell RC, Croteau KA, De BI, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act 2011 Jul 28;8:80 [FREE Full text] [doi: 10.1186/1479-5868-8-80] [Medline: 21798044]
49. Przybylski AK, Rigby CS, Ryan RM. A motivational model of video game engagement. Rev General Psychol 2010 Jun 01;14(2):154-166. [doi: 10.1037/a0019440]
50. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. Motiv Emot 2006 Nov 29;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]

XSL•FO
RenderX

51. Caldwell AE, Masters KS, Peters JC, Bryan AD, Grigsby J, Hooker SA, et al. Harnessing centred identity transformation to reduce executive function burden for maintenance of health behaviour change: the Maintain IT model. Health Psychol Rev 2018 Sep;12(3):231-253 [FREE Full text] [doi: 10.1080/17437199.2018.1437551] [Medline: 29402182]
52. Dacey M, Baltzell A, Zaichkowsky L. Older adults' intrinsic and extrinsic motivation toward physical activity. Am J Health Behav 2008;32(6):570-582. [doi: 10.5555/ajhb.2008.32.6.570] [Medline: 18442337]
53. Janssen SL, Stube JE. Older adults' perceptions of physical activity: a qualitative study. Occup Ther Int 2014 Jun;21(2):53-62 [FREE Full text] [doi: 10.1002/oti.1361] [Medline: 24302685]
54. Maula A, LaFond N, Orton E, Iliffe S, Audsley S, Vedhara K, et al. Use it or lose it: a qualitative study of the maintenance of physical activity in older adults. BMC Geriatr 2019 Dec 12;19(1):349 [FREE Full text] [doi: 10.1186/s12877-019-1366-x] [Medline: 31830900]
55. Kosteli M, Williams SE, Cumming J. Investigating the psychosocial determinants of physical activity in older adults: a qualitative approach. Psychol Health 2016 Jun;31(6):730-749 [FREE Full text] [doi: 10.1080/08870446.2016.1143943] [Medline: 26964473]
56. McAuley E, Szabo A, Gothe N, Olson EA. Self-efficacy: implications for physical activity, function, and functional limitations in older adults. Am J Lifestyle Med 2011 Jul;5(4):10.1177/1559827610392704 [FREE Full text] [doi: 10.1177/1559827610392704] [Medline: 24353482]
57. Lindsay Smith G, Banting L, Eime R, O'Sullivan G, van Uffelen JG. The association between social support and physical activity in older adults: a systematic review. Int J Behav Nutr Phys Act 2017 Apr 27;14(1):56 [FREE Full text] [doi: 10.1186/s12966-017-0509-8] [Medline: 28449673]
58. Jancey JM, Clarke A, Howat P, Maycock B, Lee AH. Perceptions of physical activity by older adults: a qualitative study. Health Educ J 2009 Sep 29;68(3):196-206. [doi: 10.1177/0017896909339531]
59. Tech adoption climbs among older Americans. Pew Research Center. 2017. URL: https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2021-11-22]
60. Social media demographics to inform your brand's strategy in 2021. Sprout Social. 2021 Mar 9. URL: https://sproutsocial.com/insights/new-social-media-demographics/ [accessed 2021-11-22]
61. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: a randomized controlled pilot trial. JMIR Mhealth Uhealth 2017 Mar 06;5(3):e28 [FREE Full text] [doi: 10.2196/mhealth.6967] [Medline: 28264796]
62. Sinclair TJ, Grieve R. Facebook as a source of social connectedness in older adults. Comput Human Behav 2017 Jan;66:363-369. [doi: 10.1016/j.chb.2016.10.003]
63. 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]
64. Tudor-Locke C, Camhi S, Troiano R. A catalog of rules, variables, and definitions applied to accelerometer data in the National Health and Nutrition Examination Survey, 2003-2006. Prev Chronic Dis 2012;9:E113 [FREE Full text] [doi: 10.5888/pcd9.110332] [Medline: 22698174]
65. Rikli R, Jones J. Senior Fitness Test Manual. Champaign: Human Kinetics; 2013.
66. Ross RM, Murthy JN, Wollak ID, Jackson AS. The six minute walk test accurately estimates mean peak oxygen uptake. BMC Pulm Med 2010 May 26;10:31 [FREE Full text] [doi: 10.1186/1471-2466-10-31] [Medline: 20504351]
67. Markland D, Tobin V. A modification to the behavioural regulation in exercise questionnaire to include an assessment of amotivation. J Sport Exerc Psychol 2004;26(2):191-196. [doi: 10.1123/jsep.26.2.191]
68. Wilson P, Rodgers W, Loitz C, Scime G. “It's Who I Am … Really!’ The importance of integrated regulation in exercise contexts. J Appl Biobehav Res 2006;11(2):79-104. [doi: 10.1111/j.1751-9861.2006.tb00021.x]
69. Vlachopoulos SP. The basic psychological needs in exercise scale: measurement invariance over gender. Struct Equ Modeling 2008 Jan 11;15(1):114-135. [doi: 10.1080/10705510701758398]
70. Sicart M. Defining game mechanics. Game Stud Int J Comput Game Res 2008;8(2):1-14 [FREE Full text]
71. Björk S, Holopainen J. Games and design patterns. Game Des Read 2006:410-437.
72. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
73. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]

Abbreviations

BCT: behavior change technique
CHALLENGE: Challenges for Healthy Aging: Leveraging Limits for Engaging Networked
G4H: game for health

XSL•FO
RenderX

Gs4H: games for health
MECHA: Mechanics, Experiences, and Change
SDT: self-determination theory

Edited by N Zary; submitted 26.04.21; peer-reviewed by N Karimi, T Ratz; comments to author 02.07.21; revised version received 20.07.21; accepted 03.08.21; published 03.12.21.

Please cite as:
Robertson MC, Baranowski T, Thompson D, Basen-Engquist KM, Swartz MC, Lyons EJ
Using the Behaviour Change Wheel Program Planning Model to Design Games for Health: Development Study
JMIR Serious Games 2021;9(4):e29964
URL: https://games.jmir.org/2021/4/e29964
doi:10.2196/29964
PMID:34870604

©Michael C Robertson, Tom Baranowski, Debbe Thompson, Karen M Basen-Engquist, Maria Chang Swartz, Elizabeth J Lyons. Originally published in JMIR Serious Games (https://games.jmir.org), 03.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design of a Virtual Reality Interactive Training System for Public Health Emergency Preparedness for Major Emerging Infectious Diseases: Theory and Framework

Yue Luo[1*], MNS; Mei Li[1*], MNS; Jian Tang[2*], MNS; JianLan Ren[3*], MNS; Yu Zheng[4], MNS; XingLi Yu[2], MNS; LinRui Jiang[5], MNS; DingLin Fan[1], MNS; YanHua Chen[6], PhD

[1]School of Nursing, Southwest Medical University, Luzhou City, China
[2]Operating Room, The Affiliated Hospital of Southwest Medical University, Luzhou City, China
[3]Department of Anesthesiology, The Affiliated Hospital of Southwest Medical University, Luzhou City, China
[4]Department of Rheumatism and Immunology, The Affiliated Hospital of Southwest Medical University, Luzhou City, China
[5]Department of Cardiac Surgery, The Affiliated Hospital of Southwest Medical University, Luzhou City, China
[6]Department of Nursing, The Affiliated Hospital of Southwest Medical University, Luzhou City, China
[*]these authors contributed equally

Corresponding Author:
YanHua Chen, PhD
Department of Nursing
The Affiliated Hospital of Southwest Medical University
25, Taiping Road
Luzhou City, 646000
China
Phone: 86 18982765016
Email: Chen_yanhua25@163.com

Abstract

Background: Sufficient public health emergency preparedness (PHEP) is the key factor in effectively responding to and recovering from major emerging infectious diseases (MEIDs). However, in the face of MEIDs, PHEP is insufficient, so it is necessary to improve PHEP. The rapid development of virtual reality and human-computer interaction provides unprecedented opportunities for innovative educational methods.

Objective: This study designed a virtual reality interactive training system (VRITS) to provide an effective path for improving PHEP in the context of MEIDs so that the public can effectively respond to and recover from MEIDs.

Methods: This study used interactive narrative, situated learning and human-computer interaction theories as a theoretical framework to guide the design of the system. We used the literature research method and the Delphi method; consulted multidisciplinary experts, such as infectious diseases, disease control, psychology, and public health personnel, to determine the educational content framework; and set up an interdisciplinary team to construct an operating system framework for the VRITS.

Results: We named the VRITS "People's War Against Pandemic." The educational content framework includes 20 knowledge, emotion, and behavior skills in 5 aspects (cooperating with prevention and control work, improving emergency response ability, guaranteeing supplies and equipment, preparing economic resources, and maintaining physical and mental health). The operating system framework includes virtual interactive training, knowledge corner, intelligent evaluation, and community forum modules, and the core module is the virtual interactive training module. In this module, users control virtual characters to move in various scenes, and then identify and analyze the controllability and harmfulness of the evolving pandemic and select the correct prevention and control strategy to avoid infecting themselves and others.

Conclusions: The development and sharing of the multidisciplinary theoretical framework adopted by People's War Against Pandemic can help us clarify the design ideas and assumptions of the VRITS; predict training results; understand the ability of training to change emergency knowledge, emergency emotion, and behavioral responses to MEIDs; and promote the development of more effective training systems based on virtual reality.

(JMIR Serious Games 2021;9(4):e29956) doi:10.2196/29956

XSL•FO
RenderX

KEYWORDS

virtual reality; major emerging infectious diseases; emergency preparedness; interactive narrative theory; situated learning theory; human-computer interaction theory; emergency simulation; public health; training; VR; epidemic; medical education; education

Introduction

Background

Since the 21st century, major emerging infectious diseases (MEIDs) have had a disastrous impact on humanity. In particular, the worldwide COVID-19 pandemic has infected more than 100 million people in more than 200 countries and regions, and the number of deaths has reached approximately 2.5 million [1]. It has led to the collapse of medical systems, the outbreak of economic crises, and social disorder in many countries [2]. It can be seen that MEIDs are highly destructive, unexpected, complex, and uncertain, and it is difficult for the public to effectively deal with them. This is related to poor public health emergency preparedness (PHEP), which is manifested in the lack of emergency knowledge and skills and the existence of negative emotional and behavioral responses [3-5]. During the COVID-19 pandemic, the public exposed problems such as inadequate scientific literacy, weak ability to select and judge information, and failure to actively cooperate with the pandemic prevention and control work [6-8]. Therefore, the public needs training to increase emergency preparedness and systematic theoretical training to acquire practical skills to prepare for potential risks. Compared with general training, systematic theoretical training enables training to be carried out in a predetermined framework and reflects the requirements in design elements to meet the needs of users [9].

Nelson et al [10] defined PHEP as "the capability of the public health and health care systems, communities, and individuals to prevent, protect against, quickly respond to, and recover from health emergencies, particularly those whose scale, timing, or unpredictability threatens to overwhelm routine capabilities." With the increasing number of MEIDs, governments and public health departments are paying increasing attention to the study of emergency preparedness. Oppenheim et al [11] constructed a global national-level emergency preparedness index framework from the following 5 dimensions: public health infrastructure, physical infrastructure, institutional capacity, economic resources, and public health communication. In 2017, the European Center for Disease Prevention and Control developed a competency model for European Union public health professionals, including the ability of discovery and assessment, policy making, adaptation and implementation, coordination and communication, and emergency risk communication with the public [12]. However, these are studies conducted from a state or institutional level rather than from the public level. This study defines PHEP for MEIDs as the ability of the public to effectively respond to and recover from MEIDs, including the preparedness of emergency skills, legal compliance, economic estimation, avoiding secondary disasters, and physical and mental health.

Traditional emergency preparedness training mainly includes books, networks, classroom learning, and campus lectures [13]. In addition, emergency drills are usually an effective means to improve PHEP [14]. There are various forms of drills, including tabletop exercises, functional exercises, and full-scale exercises. A 10-week tabletop drill of hospital disaster preparedness from 9 participating hospitals showed that their ability of emergency response and treatment (disaster preparedness of equipment, risk communications, positive responses, and surge capacity) improved [15]. The systematic review showed that current drills are mostly aimed at public health professionals, and then the drills are extended to nonpublic health professionals, such as medical students, dentists, and dental health workers with special functional needs to serve the community [16]. Emergency drills for MEIDs need a lot of workforce and material resources. In fact, it is difficult for the public to get an opportunity to participate in such emergency drills, so they may face the challenge of implementation.

With the rapid development of mobile internet and virtual reality (VR) technology, the superiority of virtual interactive training resources is attracting increasing attention, which has a positive impact on medicine and health, education and teaching, engineering technology, etc [17-19]. Virtual interaction adopts the mode of the combination of VR and human-computer interaction and uses the computer to generate a 3D space composed of graphics so that users can visually feel immersed in the virtual environment and realize the information exchange process between human and computer through human-computer interaction. Virtual Police (ViPOL) [20] uses a multidisciplinary background to put theory into practice and develop a virtual training system. However, most of these training systems only describe theoretical knowledge but rarely explain their theoretical basis. This restricts the observation angle, the way of thinking, and the ability to understand effective mechanisms, which are important for the design of a comprehensive VRITS.

Objectives

To address the public's need to participate in emergency drills, we combined emergency preparedness training with virtual interaction to develop a VRITS for PHEP for MEIDs. Here, we introduced the educational content, theoretical, and operating system frameworks of the VRITS for emergency preparedness for MEIDs. The system is named People's War Against Pandemic, aiming to provide the public with information on emergency knowledge and skills, emergency emotion, and behavioral responses to MEIDs in order to effectively respond to and recover from the pandemic.

Methods

Educational Content Framework

Through the method of literature review, we initially constructed an educational content framework for PHEP for MEIDs based on the global national-level emergency preparedness index framework [8], the 4-level response for public emergency and citizen health emergency literacy in China. Furthermore, the framework includes 6 first-level indices and 21 second-level

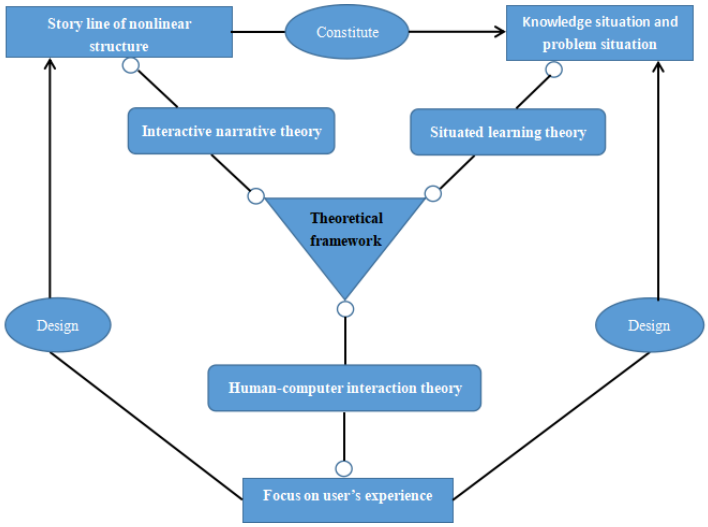
XSL•FO
RenderX

indices. The 6 first-level indices involve cooperating with prevention and control work, improving emergency response ability, perfecting basic equipment, ensuring personal safety, preparing economic resources, and managing self-emotion. Then, we consulted 12 experts through the Delphi method to revise the framework, including 3 public health experts, 3 infectious diseases experts, 4 disease control experts, and 2 psychology experts. These experts had high theoretical knowledge and rich practical experience and carried authority and representativeness. In addition, these experts had a bachelor's degree or above, an intermediate title or above, and 8 years of working experience or more. We sent consultation questionnaires to the qualified experts by email. Index screening was mainly based on the importance of each index evaluated by the experts and their specific suggestions, and the indices were screened by the arithmetic mean, the coefficient of variation (CV) of each index, and the Kendall coefficient of concordance (KCC). Importance was assigned on a scale of 1-5, ranging from very unimportant to very important. The retention criteria of the indices were that the arithmetic mean was greater than 3.50, the CV was less than 0.25, and the KCC of the second round was greater than that of the first round.

Theoretical Framework

People's War Against Pandemic constructed the operating system using interactive narrative, situated learning, and human-computer interaction theories as a theoretical framework. Interactive narrative theory provides a reference for the design of a story line in the multidisciplinary background of a VRITS [21]. Situated learning theory provides a theoretical basis for the construction of an immersive, realistic, social, and practical environment [22]. Human-computer interaction theory centers on the user's experience, which provides a theoretical basis for designing a user-friendly human-computer interaction interface [23]. In the design process of the VRITS, we integrated the 3 theories into a theoretical framework of the system (Figure 1). The story line that was designed with a nonlinear structure constituted the learning situation, and the human-computer interaction interface between the story line and the learning situation was designed with the users' experience as the center.

Figure 1. The theoretical framework.



Interactive Narrative Theory: Users Decide the Development of the Plot

Interactive narrative theory is a branch derived from narrative theory. It is a new narrative method that adds interaction on the basis of narrative. The characteristic of interactive narrative theory is that the development of the plot is decided by the users at the key nodes of the plot, and the small plots are connected to form a complete story in the process of interaction [24]. Compared with the traditional narrative mode of a single linear structure, interactive narrative is a nonlinear structure. Although the single linear narrative structure has branches, it eventually forms a 1D linear structure after simplification, which leads to the same result after each training session. Repetitive training can easily make the users tired, which is not conducive to effective absorption of knowledge. In the nonlinear narrative structure, the user is the controller of the whole process. At each interaction point, the user can randomly select an option, and the story and feedback change accordingly [25]. Interactive narrative pays more attention to the process of exploration. Even if the users fail, they can start again.

Based on the interactive narrative theory of a nonlinear structure, the plot of People's War Against Pandemic was designed. The plot is about a MEID (COVID-19, cholera, Ebola, plague, or unknown-cause infection) outbreaking in a specific place. Users need to identify and analyze the controllability and harmfulness of the MEID, select the correct prevention and control strategies to control the virtual characters to move in various scenes (residential areas, medical institutions, pharmacies, police stations, banks, schools), and choose the correct behaviors to avoid infecting themselves and others. In this process, the users' knowledge, emotions, and behavioral responses to prevention and control of the MEID are trained. After obtaining the users' instruction information, the system creates multimodal interactive data to activate preset character scenes and synchronously generates matching dynamic pictures to form a complete training story.

In the whole training process, using the first-person perspective can increase the immersion of the users and trigger interaction nodes more conveniently. There are 5 main story lines for 5 kinds of infectious diseases, and each main story line has more

XSL•FO
RenderX

than 5 interactive nodes. The first node is the selection of the type of infectious disease. Since the source of infection, transmission route, susceptible population, and harmfulness of each infectious disease are different, the choice of each node after the first node is different. Therefore, the choice of the first node is of vital importance. Then it leads to different plots with 4 interactive nodes of different scenes, prevention and control strategies, response measures, and personal situations. The users can improve their emergency preparedness in the process of making continuous choices. This highlights the causal relationship between decision making and results in the narrative process.

Situated Learning Theory: Learning in Real Situations

Situated learning theory holds that learning is not only an individual information acquisition process but also a social and practical participation process [26]. In the beginning, when situated learning theory was proposed, learners played roles in the real world. With the development of information technology, the environment of situated learning expanded to an online situation, simulation laboratories, and other virtual scenes [27], among which the VR interactive system is more immersive and realistic. In the emergency preparedness training in MEIDs, in view of the explosive and highly infectious characteristics of infectious diseases, users cannot learn in the real situation. Therefore, we designed a virtual learning situation based on the practical and social characteristics of situated learning, combined with the key nodes of interactive narrative stories.

The design of a virtual learning situation includes the design of a knowledge situation and a problem situation. A knowledge situation refers to the background and process situation in which knowledge occurs, which can help users construct the meaning of knowledge. A problem situation [28] emphasizes that learning should be placed in meaningful and complex situations so that users can take the initiative to learn and improve the ability to solve problems during the process of solving problems. We designed the knowledge and problem situations based on the nonlinear structured story line and the educational content for emergency preparedness. In total, 5 main story lines were designed with 5 complete learning situations, each with 5 branches of learning situations and 20 specific knowledge and problem situations.

Human-Computer Interaction Theory: Focus on Users' Experience

Human-computer interaction refers to the process of information exchange between human and computer in a specific way with a specific language to achieve a specific function [29]. In the process of human-computer interaction, usability becomes important to accomplish users' tasks efficiently, clearly, and safely. The usability criteria include a low learning cost, fast task execution, a low error rate, high satisfaction, and a high return visit rate [30]. Therefore, the design of human-computer interaction needs to focus on the users' experience by taping their needs and analyzing their characteristics to make them have a good sense of experience.

To design an interactive interface that allows users to have a good sense of experience, we set up an interdisciplinary team of 10 experts from the fields of infectious diseases, education, psychology, and computers to repeatedly modify the preliminary design plan and interactive interface script and finally determine the design plan and script. The design plan is the function and realization path of each module of the operating system framework. The content of the script includes the interaction points involved in each interactive interface, all of which constitute the learning situation and story line. The preliminary design plan and script were modified and reviewed by experts from the fields of infectious diseases, education, and psychology and then handed to experts from the computer field to evaluate the implementation path of the interactive interface and preliminarily design several interactive interfaces. After the evaluation, the modification opinions were returned. Experts from the fields of infectious diseases, education, and psychology modified them. Experts from the computer field evaluated and again modified them until the final version was determined. In addition, before designing the VRITS, we conducted an investigation of the needs of more than 4000 Chinese students and found that in terms of training content, Chinese students have low awareness of infectious diseases and prevention and control skills [31]. Therefore, we paid more attention to the interface design of infectious disease knowledge and prevention and control skills.

Results

Educational Content Framework for PHEP for MEIDs

In the first round of expert consultation, the arithmetic mean of each first-level index was 4.33-4.92 and the CV was 0.06-0.15. In addition, the arithmetic mean of each second-level index was 4.08-4.92 and the CV was 0.06-0.34. According to the experts' specific suggestions, the second-level indices assigned to "ensure personal safety" were more suitable to be assigned to other first-level indices, and the first-level index "ensure personal safety" was deleted. Therefore, the second-level indices in this deleted index were assigned to "cooperate with prevention and control work" and "improve emergency response ability." In addition, some changes were made to other indices according to the experts' suggestions. Thus, 5 first-level indices and 20 second-level indices were formed, and a second round of expert consultation was conducted.

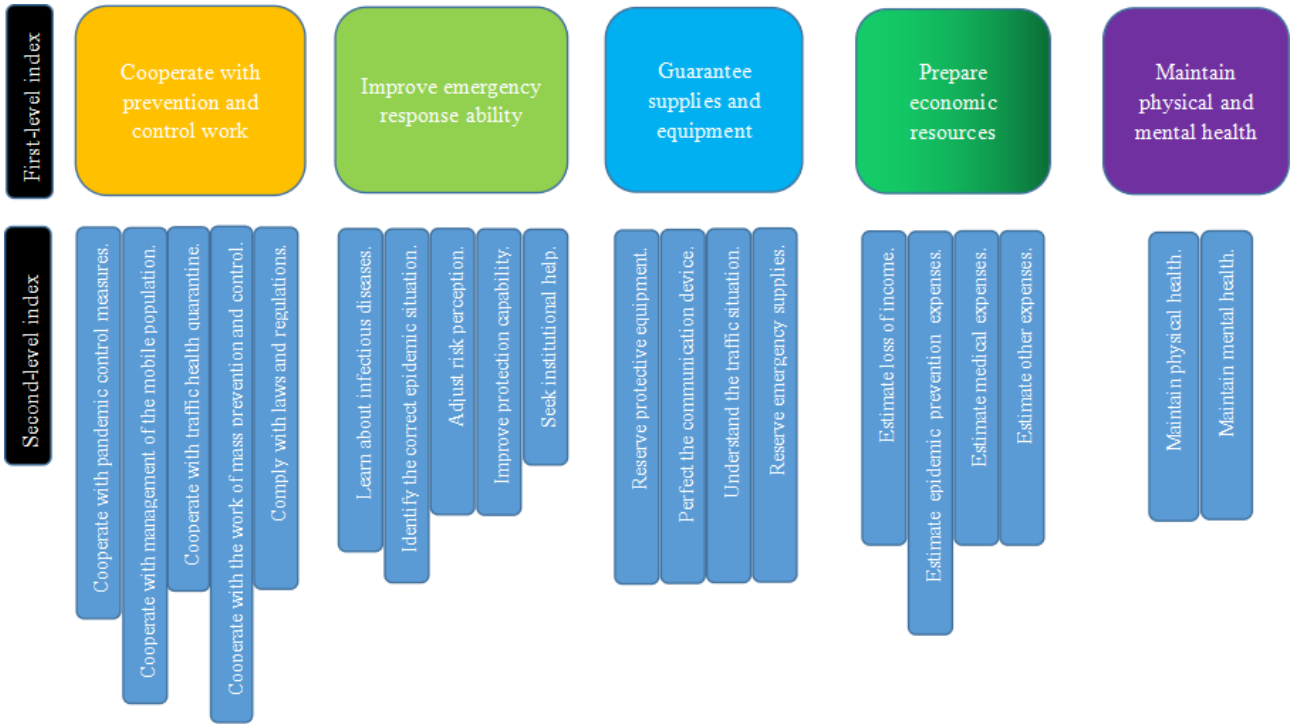
In the second round of consultation, the arithmetic mean of each first-level index was 4.42-4.92 and the CV was 0.06-0.12. In addition, the arithmetic mean of each second-level index was 3.67-4.92 and the CV was 0.06-0.22. According to the results of the second round of consultation, the scores of each index met the index retention criteria, and the KCC of the second round was greater than that of the first round (Table 1). The experts did not put forward specific opinions, so the index framework was not modified. Finally, the educational content framework for PHEP for MEIDs included 5 first-level indices and 20 second-level indices. The 5 first-level indices were "cooperating with prevention and control work," "improving emergency response ability," "guaranteeing supplies and equipment," "preparing economic resources," and "maintaining physical and mental health" (Figure 2).

XSL•FO
RenderX

Table 1. Kendall coefficient of concordance (KCC) for the 2 rounds of consultation.

Round	Entries, n	W	χ^2	*P* value
First round				
First-level index	6	0.258	15.474	.01
Second-level index	21	0.269	64.448	<.001
All indices	27	0.245	76.385	<.001
Second round				
First-level index	5	0.270	12.952	.01
Second-level index	20	0.358	81.626	<.001
All indices	25	0.350	100.713	<.001

Figure 2. Educational content framework for PHEP under MEIDs. MEID: major emerging infectious disease; PHEP: public health emergency preparedness.

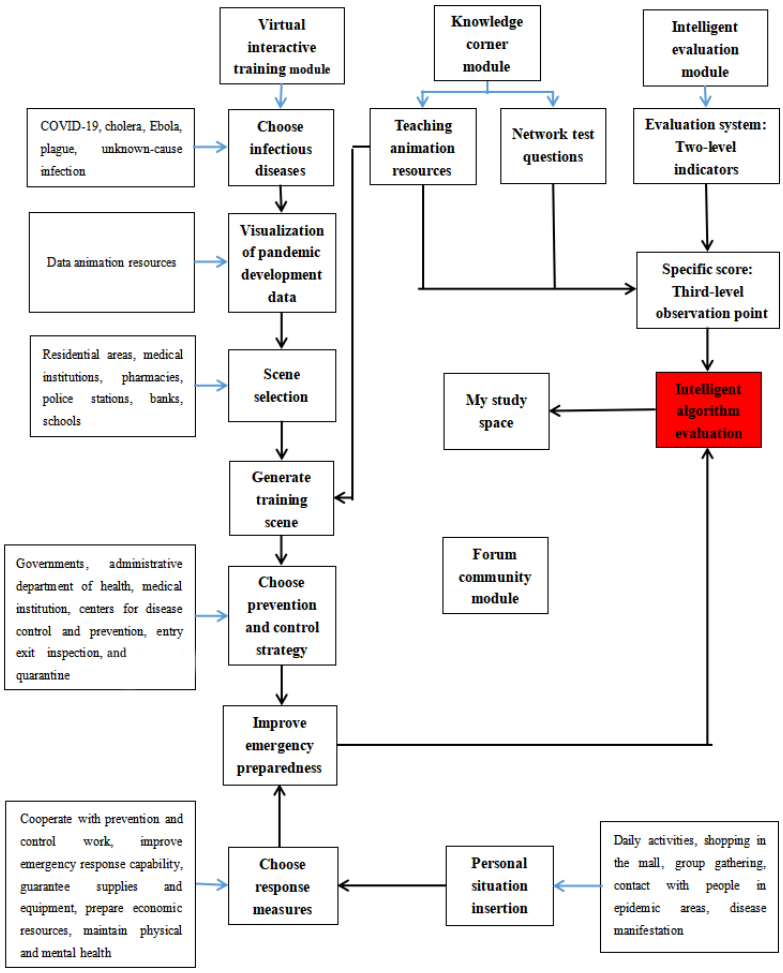


Operating System Framework

The operating system framework of the VRITS consists of 4 modules: virtual interactive training, knowledge corner, intelligent evaluation, and forum community. Figure 3 shows the overall design framework of the operating system. The development software we used was Unity 3d, the development tool was Microsoft Visual Studio, and the modeling tool was Autodesk 3ds Max. Users can access the system through a link on the web pages on computers, mobile phones, and tablets, and they can get it for free after registration.

XSL•FO
RenderX

Figure 3. Operating system framework of the VRITS. VRITS: virtual reality interactive training system.



Virtual Interactive Training Module

The characters in the virtual interactive training module are generated by computer (Figure 4). The users control the virtual characters to select specific prevention and control measures to complete the training tasks according to the visualized information about the pandemic development data and tasks prompted in the interactive interface. The virtual interactive training module is the core module of People's War Against Pandemic, which contains 5 situations. The users randomly select a scene to enter, and multiple choices are set in the interactive nodes of each scene, including the selection of training routes, of answers to knowledge points, and of props in the props column. Every choice can get dynamic feedback. In the selection process, if a choice is wrong, the interactive interface prompts the users with an error and the score is deducted. The interactive text box of the interactive interface of each node prompts the users about the development of the pandemic, and a personal situation animation is inserted in the process of pandemic prevention, such as daily activities, shopping, group gathering, and contacting infected people in the pandemic area. If the users have corresponding clinical symptoms in their personal situation, they need to judge whether they are infected and then take corresponding countermeasures.

In the whole virtual interactive situation, if the virus is rampant and the pandemic spreads, a dialog box appears in the interactive interface to prompt the users to improve the level of prevention and control. Individuals should train their decision-making skills as well as their 20 knowledge and behavior skills in 5 aspects (cooperating with prevention and control work, improving emergency response ability, guaranteeing supplies and equipment, preparing economic resources, and maintaining physical and mental health). The success or failure of pandemic prevention and control is prompted by corresponding animation. If the user makes the right choices, they are not be infected and the interactive interface prompts "Training successful. Please check the score." If the training fails, a warning appears on the interactive interface to prompt the virtual character to be infected with the disease (Figure 4). More seriously, it promotes the spread of the virus, the number of infected people rises sharply, and the pandemic cannot be controlled. Then the animation of secondary disasters corresponding to the diseases (crowd gathering, panic buying of supplies, shortage of medical resources) appears in the interactive interface as warning education.

XSL•FO
RenderX

Figure 4. Interface of the VRITS. There are four interfaces in the picture. Interface 1 is the login interface; Interface 2 is the image of the virtual character and the entrance of the four modules; Interface 3 is the entrance of the different scenes; and Interface 4 is a warning interface for the virtual character infected with the disease after the training fails. VRITS: virtual reality interactive training system.

Knowledge Corner Module

The knowledge corner module contains all kinds of knowledge related to MEIDs (including pathogeny, clinical manifestations, diagnosis, prevention and treatment). This knowledge is conveyed in the form of words, pictures, and videos, which is the supplement and integrity of the educational content of the virtual interactive training module. The administrator uploads the latest knowledge of infectious diseases (including the protection skills of infectious diseases, laws and regulations related to MEIDs, and secondary disasters caused by the pandemic) from the background of the system, and each knowledge point matches the network test questions and answers, which is convenient for users to test and then correct the answers.

Intelligent Evaluation Module

The intelligent evaluation module takes the indices at all levels of the educational content framework for emergency preparedness as specific scoring points. Every behavior of the users in the training process is recorded and intelligently scored. The training starts out on a 100-point scale, with points deducted in real time if the users make an incorrect choice. At the end of the training, the users can view the total score and the score for each first-level index. Automatic interactive text analysis technology, such as participation analysis, social network analysis, and content analysis, is used to obtain the users' participation in learning, social network, concerned learning content, and online behavior information to analyze their characteristics and to push their personalized training resources. The module can also provide data collection, integration, classification, and application functions and offer objective and real-time big data for the VR interactive training in MEIDs.

Forum Community Module

In the forum community module, users can enter the forum community to participate in questions, discussions, and communications. They can publish topic posts related to the pandemic, help posts for difficulties encountered in simulation training, experience posts for success in fighting the pandemic, and announcement posts issued by administrators. Using the interactive, direct, and group characteristics of the forum, we can realize the interaction between user and user, user and interactive interface, and user and teaching content. This kind of real-time communication makes it easy to find and solve problems in time.

Discussion

Principal Findings

We drew knowledge from multiple disciplines, such as education, medicine, psychology, public health, and computers, which provided reference information for the theoretical framework design of People's War Against Pandemic. Then we put forward the theoretical basis of multidisciplinary integration of interactive narrative, situated learning, and human-computer interaction and designed a VRITS. A clear theoretical basis can help us clarify the design ideas and assumptions of the VRITS; predict training results; understand the way of training to change emergency knowledge, emergency emotion, and behavioral responses to MEIDs; and promote the development of a more effective VRITS.

Whether interactive narrative theory plays a core role or a peripheral role in a VRITS usually depends on the teaching objectives of virtual interaction and the consequent changes in knowledge, emotion, and behavior. In the prevention and treatment of some chronic diseases and mental diseases based on virtual interactive games, interactive narrative often plays a

XSL•FO
RenderX

secondary incentive role [32]. The changes in the symptoms of chronic diseases and mental state are usually rooted in daily life. In People's War Against Pandemic, the users' knowledge and skills, emotional adjustment, and behavioral responses of emergency preparedness for MEIDs run through the main line of the whole interactive narrative, making it the core position in People's War Against Pandemic. Therefore, users transform the knowledge and skills learned from virtual interaction into specific attitudes or behavioral responses in life, which is closely related to the role of the interactive narrative. In addition, how to deal with the relationship between story and interaction in interactive narrative is also particularly important [24]. In the process of virtual interactive training, there are many branches in each interactive node, and the choice of each branch has an impact on the final result. Eventually, each branch of all interactive nodes is arranged and combined into different stories. Each interaction determines the direction of the story, and the story is composed of every interaction.

Situated learning theory also plays an important role in a VRITS. The purpose of a virtual learning situation is to create a social and participatory experience for the users and realize the functions that cannot be completed in real life [33]. It is the cornerstone of the system. In the past few years, more fields have applied situated learning theory to design virtual learning situations. "Tawaf" [34] emphasizes the positive role of situated learning in improving learners' interest in learning by solving problems in a 3D VR environment with virtual characters. A study on a virtual learning environment for disaster risk reduction explored the factors that affect the sense of reality of a learning situation in a virtual environment [35]. Among them, the positive factors to pay attention to are socialization and learning from mistakes, which give learners a sense of reality in face-to-face interaction and to safely fail in a virtual environment. In brief, the use of a virtual learning situation can enhance the reality of the scene, the immersion of training, and the sense of experience and stimulate the motivation and efficiency of learning. In addition, situated learning and interactive narrative are integrated. The learning situation is composed of story lines with a nonlinear structure, and all stories take place in the learning situation, which is the carrier of the story lines.

If interactive narrative theory is the core of a VRITS and situated learning theory is the cornerstone of the VRITS, then human-computer interaction theory plays the role of a bridge of information exchange between the system and the users. According to the users' cognitive ability and information acquisition ability, the human-computer interaction interface is designed centering on the users' experience and the information communication between human and computer is finally realized [36]. Eventually, based on human-computer interaction theory, this study designed the interactive interface of a learning situation and story line, which organically integrated the 3 theories to make the information communication smooth and efficient.

The PHEP education for MEIDs mainly includes popular science knowledge, protection skills, and cooperation in pandemic prevention and control [37]. In addition, emergency preparedness, such as social cooperation, legal compliance, economic prediction, social stability, and avoiding of secondary disasters, are also important. Therefore, this study constructed an educational content framework of emergency preparedness to train users on 20 items of emergency knowledge and skills, emotion, and behavioral responses in 5 aspects. The findings will enrich the theoretical system of emergency preparedness education for MEIDs and provide the public with a panoramic understanding of the response to MEIDs so that their comprehensive preparation can be promoted and ultimately an effective response can be achieved.

As a new training system, People's War Against Pandemic not only has the characteristics of a 3D stereoscopic combination of sound, text, image, and interaction but also uses a sense of reality of VR scenes to fully arouse the interest of users and satisfy their curiosity and thirst for knowledge. Through People's War Against Pandemic simulation training, users can keep in mind the correct process and knowledge points of prevention and control of MEIDs in constantly wrong choices and effectively improve their ability to comprehensively apply knowledge and skills to improve emergency preparedness and reduce disease infection. At the same time, this training allows users to place themselves in a virtual situation from a first-person perspective, mobilizing the users' subjective initiative and stimulating their learning motivation through role-playing and a realistic virtual environment.

As an innovative way of emergency training for infectious diseases, People's War Against Pandemic is an attempt to develop emergency training resources for MEIDs. Its design of educational content is more comprehensive, covering many core parts of knowledge and skills related to MEIDs. In addition, the system can judge whether the prevention and control strategies for infectious diseases selected by users are correct behaviors, and timely feedback and adjustment of problems in the training process will play a substantial positive role in improving the learning effect [38]. Based on this, People's War Against Pandemic has set up a real-time feedback function. The intelligent evaluation module records and evaluates each option of each user and presents the results in the interactive interface in the form of a score, which effectively realizes the real-time feedback of information. Offline emergency drills often face implementation challenges due to the cost of workforce, material, and financial resources. The online resources of People's War against Pandemic can be promoted to the public through the internet to solve the bottleneck problem of offline emergency drills. The system can be released by the municipal Center for Disease Control and Prevention and tried out among community residents. In the process of using the system, the residents' use of experience, satisfaction, opinions, and suggestions can be collected to modify and improve the system, and finally it can be promoted and applied nationwide.

Limitations

There are some limitations in this study. The immersion of the VR interactive system designed in this study may make some young or low-self-control users indulge in it. However, this problem may be solved by limiting the length of training [39]. In addition, it will take some time to popularize the VR interactive devices. Some users may experience the effect of

XSL•FO
RenderX

the VR interactive system on mobile software, but they cannot experience a more realistic sense of operation with VR devices. Moreover, users cannot customize roles for the time being. At present, the system roles are unified as citizens, and there is no division by age, gender, and occupation. Therefore, it may have some impact on users' immersion and usability. This problem can be solved when the system is upgraded and improved in the future. Most importantly, we did not conduct a randomized controlled trial on the effectiveness of VR interactive systems. Therefore, in the future, we need to carry out empirical research to verify the effectiveness of virtual interaction systems.

Conclusions

A VRITS is a product of a highly integrated multidisciplinary background. The application of virtual interaction to the emergency preparedness training in MEIDs may solve the bottleneck problem in the current emergency preparedness training in MEIDs. Based on a solid theoretical foundation, rich teaching resources, a virtual learning situation, a complete story line, and an interactive function, the system realizes the visualization of emergency preparedness training for MEIDs, which may stimulate the learning motivation of users, increase their sense of experience and immersion, enhance their learning participation and learning efficiency, and improve the training effect. In future practice, to improve the level of PHEP for MEIDs, it is necessary to further perfect the VRITS and promote its use for emergency preparedness education of MEIDs.

Acknowledgments

This research was supported by the Sichuan Science and Technology Program (21KPZP0280) and the Sichuan Hospital Management and Development Research Center, Southwest Medical University (SCYG2021-25). The authors would like to thank the experts for their guidance.

Authors' Contributions

The concept was developed by YL; the methodology by JT, YZ, LRJ, and DLF; and the design by YL, ML, JLR, YZ, JT, and YHC. The original draft preparation (writing) was performed by YL, ML, JLR, and JT; review and editing (writing) by YL, ML, JLR, JT, XLY, and YHC; project administration by YL and YHC; and funding acquisition by YHC.

Conflicts of Interest

None declared.

References

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. URL: https://covid19.who.int/ [accessed 2020-03-06]
2. Mofijur M, Fattah IMR, Alam MA, Islam ABMS, Ong HC, Rahman SMA, et al. Impact of COVID-19 on the social, economic, environmental and energy domains: lessons learnt from a global pandemic. Sustain Prod Consum 2021 Apr;26:343-359 [FREE Full text] [doi: 10.1016/j.spc.2020.10.016] [Medline: 33072833]
3. Pedrosa AL, Bitencourt L, Fróes ACF, Cazumbá MLB, Campos RGB, de Brito SBCS, et al. Emotional, behavioral, and psychological impact of the COVID-19 pandemic. Front Psychol 2020;11:566212 [FREE Full text] [doi: 10.3389/fpsyg.2020.566212] [Medline: 33117234]
4. Gallè F, Sabella EA, Da Molin G, De Giglio O, Caggiano G, Di Onofrio V, et al. Understanding knowledge and behaviors related to CoViD-19 epidemic in Italian undergraduate students: the EPICO Study. Int J Environ Res Public Health 2020 May 16;17(10):3481 [FREE Full text] [doi: 10.3390/ijerph17103481] [Medline: 32429432]
5. Torales J, O'Higgins M, Castaldelli-Maia JM, Ventriglio A. The outbreak of COVID-19 coronavirus and its impact on global mental health. Int J Soc Psychiatry 2020 Jun;66(4):317-320. [doi: 10.1177/0020764020915212] [Medline: 32233719]
6. Research Group of Southwest University, Sun C. Study on the impact of COVID-19 epidemic on young college students-Based on the investigation of 45 college students in China. China Youth Study 2020(4):43-48+12. [doi: 10.19633/j.cnki.11-2579/d.2020.0052]
7. Han Y, Jiang B, Guo R. Factors affecting public adoption of COVID-19 prevention and treatment information during an infodemic: cross-sectional survey study. J Med Internet Res 2021 Mar 11;23(3):e23097 [FREE Full text] [doi: 10.2196/23097] [Medline: 33600348]
8. Jin X, Li J, Song W, Zhao T. The impact of COVID-19 and public health emergencies on consumer purchase of scarce products in China. Front Public Health 2020;8:617166 [FREE Full text] [doi: 10.3389/fpubh.2020.617166] [Medline: 33344410]
9. Gralla E, Goentzel J, Chomilier B. Case study of a humanitarian logistics simulation exercise and insights for training design. J Humanit Logist Supply Chain Manag 2015;5(1):113-138 [FREE Full text] [doi: 10.1108/jhlscm-01-2014-0001]
10. Nelson C, Lurie N, Wasserman J, Zakowski S. Conceptualizing and defining public health emergency preparedness. Am J Public Health 2007 Apr;97 Suppl 1(Suppl 1):S9-11. [doi: 10.2105/AJPH.2007.114496] [Medline: 17413078]

XSL•FO
RenderX

11. Oppenheim B, Gallivan M, Madhav NK, Brown N, Serhiyenko V, Wolfe ND, et al. Assessing global preparedness for the next pandemic: development and application of an Epidemic Preparedness Index. BMJ Glob Health 2019;4(1):e001157 [FREE Full text] [doi: 10.1136/bmjgh-2018-001157] [Medline: 30775006]
12. Stoto MA, Savoia E, Nelson C, Piltch-Loeb R, Guicciardi S, Takacs J, et al. Development of a public health emergency preparedness competency model for European Union countries. Euro Surveill 2018 Dec;23(49):1700631 [FREE Full text] [doi: 10.2807/1560-7917.ES.2018.23.48.1700631] [Medline: 30621822]
13. Valesky W, Silverberg M, Gillett B, Roblin P, Adelaine J, Wallis LA, et al. Assessment of hospital disaster preparedness for the 2010 FIFA World Cup using an internet-based, long-distance tabletop drill. Prehosp Disaster Med 2011 Jun;26(3):192-195. [doi: 10.1017/S1049023X11006443] [Medline: 22107770]
14. Skryabina E, Reedy G, Amlôt R, Jaye P, Riley P. What is the value of health emergency preparedness exercises? A scoping review study. Int J Disaster Risk Reduct 2017 Mar;21:274-283 [FREE Full text] [doi: 10.1016/j.ijdrr.2016.12.010]
15. Ali M, Williams MD. No-notice mystery patient drills to assess emergency preparedness for infectious diseases at community health centers in New York City, 2015-2016. J Community Health 2019 Apr;44(2):387-394 [FREE Full text] [doi: 10.1007/s10900-018-00595-5] [Medline: 30604224]
16. Savoia E, Lin L, Bernard D, Klein N, James LP, Guicciardi S. Public health system research in public health emergency preparedness in the United States (2009-2015): actionable knowledge base. Am J Public Health 2017 Sep;107(S2):e1-e6. [doi: 10.2105/AJPH.2017.304051] [Medline: 28892437]
17. Riva G, Serino S. Virtual reality in the assessment, understanding and treatment of mental health disorders. J Clin Med 2020 Oct 26;9(11):3434 [FREE Full text] [doi: 10.3390/jcm9113434] [Medline: 33114623]
18. Wilcha RJ. Effectiveness of virtual medical teaching during the COVID-19 crisis: systematic review. JMIR Med Educ 2020 Nov 18;6(2):e20963 [FREE Full text] [doi: 10.2196/20963] [Medline: 33106227]
19. Song J, Shi Z, Wang L, Wang H. Improved virtual gyroscope technology based on the ARMA model. Micromachines (Basel) 2018 Jul 11;9(7):348 [FREE Full text] [doi: 10.3390/mi9070348] [Medline: 30424281]
20. Moskaliuk J, Bertram J, Cress U. Training in virtual environments: putting theory into practice. Ergonomics 2013;56(2):195-204. [doi: 10.1080/00140139.2012.745623] [Medline: 23231585]
21. Ostrin G, Frey J, Cauchard JR. Interactive narrative in virtual reality. 2018 Presented at: Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia; Nov 25-28, 2018; Cairo p. 436-467. [doi: 10.1145/3282894.3289740]
22. Huang H, Liaw S. Applying situated learning in a virtual reality system to enhance learning motivation. Int J Inf Educ Technol 2011;1(4):298-302 [FREE Full text] [doi: 10.7763/ijiet.2011.v1.48]
23. Shneiderman B, Plaisant C, Cohen MS, Jacobs S. Designing the user interface strategies for effective human-computer interaction. Hoboken, NJ: Pearson; Mar 1987:6.
24. van Lent M. Interactive narrative. IEEE Comput Graph Appl 2006;26(3):20-21 [FREE Full text] [doi: 10.1109/mcg.2006.59] [Medline: 16711213]
25. Riedl M, Bulitko V. Interactive narrative: an intelligent systems approach. AIMag 2012 Dec 06;34(1):67-77 [FREE Full text] [doi: 10.1609/aimag.v34i1.2449]
26. Lave J, Wenger E. Learning in doing: social, cognitive and computational perspectives. In: Situated Learning: Legitimate Peripheral Participation. Cambridge: Cambridge University Press; 1991.
27. Artemeva N, Rachul C, O'Brien B, Varpio L. Situated learning in medical education. Acad Med 2017 Jan;92(1):134 [FREE Full text] [doi: 10.1097/ACM.0000000000001495] [Medline: 28030424]
28. Matlala S. Educators' perceptions and views of problem-based learning through simulation. Curationis 2021 Mar 10;44(1):e1-e7 [FREE Full text] [doi: 10.4102/curationis.v44i1.2094] [Medline: 33764129]
29. Hewett TT, Baecker R, Card S, Carey T, Gasen J, Mantei M, et al. ACM SIGCHI curricula for human-computer interaction. In: Tech Rep. New York: Association for Computing Machinery; 1992:173.
30. Olson GM, Olson JS. Human-computer interaction: psychological aspects of the human use of computing. Annu Rev Psychol 2003;54:491-516. [doi: 10.1146/annurev.psych.54.101601.145044] [Medline: 12209025]
31. Yu X, Li A, Tang J, Xie H, Zhu J, Xiong W, et al. Risk perception, emergency preparedness and emergency literacy among medical students under COVID-19 epidemic. Chin J School Health 2020;41(1):1646-1649. [doi: 10.16835/j.cnki.1000-9817.2020.11.013]
32. Cavazza M, Charles F. Towards interactive narrative medicine. Stud Health Technol Inform 2013;184:59-65. [Medline: 23400131]
33. Dawley L, Dede C. Situated learning in virtual worlds and immersive simulations. In: Handbook of Research on Educational Communications and Technology. New York: Springer; 2014:723-734.
34. Yasin AM, Darleena Z, Isa MAM. Avatar implementation in virtual reality environment using situated learning for “Tawaf”. Procedia Soc Behav Sci 2012 Dec;67:73-80 [FREE Full text] [doi: 10.1016/j.sbspro.2012.11.308]
35. Caroca J, Bruno M, Aldunate R. Situated learning based on virtual environment for improving disaster risk reduction. J E-Learn Knowl Soc 2016;12(4):81-92 [FREE Full text]
36. Good A, Omisade O. Linking activity theory with user centred design: a human computer interaction framework for the design and evaluation of mHealth interventions. Stud Health Technol Inform 2019 Jul 30;263:49-63. [doi: 10.3233/SHTI190110] [Medline: 31411152]

XSL•FO
RenderX

37. Yang S, Cho SI. Middle East respiratory syndrome risk perception among students at a university in South Korea, 2015. Am J Infect Control 2017 Jun 01;45(6):e53-e60 [FREE Full text] [doi: 10.1016/j.ajic.2017.02.013] [Medline: 28385465]
38. Van Ginkel S, Ruiz D, Mononen A, Karaman C, Keijzer A, Sitthiworachart J. The impact of computer-mediated immediate feedback on developing oral presentation skills: an exploratory study in virtual reality. J Comput Assist Learn 2020 Feb 13;36(3):412-422. [doi: 10.1111/jcal.12424]
39. Coyne L, Merritt TA, Parmentier BL, Sharpton RA, Takemoto JK. The past, present, and future of virtual reality in pharmacy education. Am J Pharm Educ 2019 Apr;83(3):7456 [FREE Full text] [doi: 10.5688/ajpe7456] [Medline: 31065173]

Abbreviations

CV: coefficient of variation
KCC: Kendall coefficient of concordance
MEID: major emerging infectious disease
PHEP: public health emergency preparedness
VR: virtual reality
VRITS: virtual reality interactive training system

Edited by N Zary; submitted 26.04.21; peer-reviewed by J Thompson, H Oyama; comments to author 28.06.21; revised version received 14.08.21; accepted 09.09.21; published 14.12.21.

Please cite as:
Luo Y, Li M, Tang J, Ren J, Zheng Y, Yu X, Jiang L, Fan D, Chen Y
Design of a Virtual Reality Interactive Training System for Public Health Emergency Preparedness for Major Emerging Infectious Diseases: Theory and Framework
JMIR Serious Games 2021;9(4):e29956
URL: https://games.jmir.org/2021/4/e29956
doi:10.2196/29956
PMID:34904951

©Yue Luo, Mei Li, Jian Tang, JianLan Ren, Yu Zheng, XingLi Yu, LinRui Jiang, DingLin Fan, YanHua Chen. Originally published in JMIR Serious Games (https://games.jmir.org), 14.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Awareness, Prevention, Detection, and Therapy Applications for Depression and Anxiety in Serious Games for Children and Adolescents: Systematic Review

Kim Martinez[1], MSc; Maria Isabel Menéndez-Menéndez[1], PhD; Andres Bustillo[2], PhD

[1]Department of History, Geography and Communication, University of Burgos, Burgos, Spain

[2]Department of Computer Engineering, University of Burgos, Burgos, Spain

Corresponding Author:
Kim Martinez, MSc
Department of History, Geography and Communication
University of Burgos
Don Juan de Austria
Burgos, 09001
Spain
Phone: 34 947 49 91 12
Email: kmartinez@ubu.es

Abstract

Background: Depression and anxiety in children and adolescents are major health problems worldwide. In recent years, serious games research has advanced in the development of tools to address these mental health conditions. However, there has not been an extensive analysis of these games, their tendencies, and capacities.

Objective: This review aims to gather the most current serious games, published from 2015 to 2020, with a new approach focusing on their applications: awareness, prevention, detection, and therapy. The purpose is also to analyze the implementation, development, and evaluation of these tools to obtain trends, strengths, and weaknesses for future research lines.

Methods: The identification of the serious games through a literature search was conducted on the databases PubMed, Scopus, Wiley, Taylor and Francis, Springer, PsycINFO, PsycArticles, Web of Science, and Science Direct. The identified records were screened to include only the manuscripts meeting these criteria: a serious game for PC, smartphone, or virtual reality; developed by research teams; targeting only depression or anxiety or both; aiming specifically at children or adolescents.

Results: A total of 34 studies have been found that developed serious games for PC, smartphone, and virtual reality devices and tested them in children and adolescents. Most of the games address both conditions and are applied in prevention and therapy. Nevertheless, there is a trend that anxiety is targeted more in childhood and depression targeted more in adolescence. Regarding design, the game genres arcade minigames, adventure worlds, and social simulations are used, in this order. For implementation, these serious games usually require sessions of 1 hour and are most often played using a PC. Moreover, the common evaluation tools are normalized questionnaires that measure acquisition of skills or reduction of symptoms. Most studies collect and compare these data before and after the participants play.

Conclusions: The results show that more awareness and detection games are needed, as well as games that mix the awareness, prevention, detection, and therapy applications. In addition, games for depression and anxiety should equally target all age ranges. For future research, the development and evaluation of serious games should be standardized, so the implementation of serious games as tools would advance. The games should always offer support while playing, in addition to collecting data on participant behavior during the game to better analyze their learning. Furthermore, there is an open line regarding the use of virtual reality for these games due to the capabilities offered by this technology.

(JMIR Serious Games 2021;9(4):e30482) doi:10.2196/30482

KEYWORDS

serious games; depression; anxiety; children; adolescents; virtual reality; mental health; detection; awareness; prevention; therapy

XSL•FO
RenderX

Introduction

The recent worldwide increase of mental health conditions is a serious health problem. According to the World Health Organization [1], there was a 13% rise in mental health issues from 2007 to 2017. People affected by mental illnesses experience a severe deterioration in their relationships with family, friends, and the community and in daily task performance. The World Health Organization estimates the most common mental health conditions, depression and anxiety, cost the global economy US $1 trillion each year [2]. These issues develop from a very young age; in fact, around 20% of the world's children and adolescents have a mental health condition. In addition to personal and financial costs, these problems can head to suicide, which is the second leading cause of death among people aged 15 to 29 years [1]. Therefore, research must address this situation by developing resources for mental health education and therapy for the very young.

To address this social demand, research on serious games regarding mental health has emerged in recent decades. Serious games use the ludic medium of video games to achieve learning goals [3]. Their utility resides in creating immersive and entertaining environments built on learning theories and methodologies based on empirical research to maximize education outcomes [4]. Serious games applied to mental health adapt cognitive behavioral therapy (CBT) to the experience allowing exploratory learning and behavior practice [5]. These tools are especially suitable for children and adolescents, since they are familiar with technology and are attracted to graphics and playability [6]. Furthermore, the development and democratization of technology allows the creation of high-quality games for the most common user devices, PCs and smartphones, and the most novel, such as virtual reality (VR) headsets [7].

Serious games have been proposed since the late 1990s for depression therapy, as the review of Li and Foo [8] outlined. Although these first games targeted varied audiences, the vast majority were applied to children and adolescents [9]. However, these first solutions were very rare, and the first reviews on these topics covered a diverse mix of disorders and neurobiological conditions like anxiety disorder, autism spectrum disorder, and attention deficit hyperactivity disorder [10-14]. More recently, serious games that facilitate exposure therapy for phobias use virtual reality to maximize their impact [15].

The standardization of serious games as reliable solutions for anxiety and depression has taken place in the last 5 years, making possible new reviews focused on serious games developed for depression care [16] and anxiety treatment [17,18]. In addition, Villani et al [19] have studied serious games for emotional regulation, an important approach to prevent depression and anxiety.

Gamified technologies for improving mental health have also been considered [20], although these are not the same as serious games. A gamified experience has gaming elements like rewards, characters, and rules but does not focus on playfulness and fun [21]. Finally, the latest reviews document the different psychiatric applications of serious games [22] and the importance of their application in children and adolescents [7,23]. Table 1 shows the main characteristics of these reviews for anxiety and depression—number of games studied, article publication ranges, disorders considered, and game targets—and their main conclusions.

The analysis provided in Table 1 reveals some limitations of the reviews. The first is the low number of studied games, ranging from 5 to 34 but mostly mixing serious games with commercial games or gamified applications. Furthermore, even being recent studies, there is not a novelty inclusion criteria, so they consider articles published as far back as 1989. Second, regarding subject and target, none of these papers focuses on all the applications for depression and anxiety a game could have, from awareness to therapy. Finally, the conclusions of these reviews point out different game design and evaluation issues, but these may be influenced by article limitations. There is a research gap concerning the latest games developed for all possible applications and their specific targets.

This review aims to study serious games applied to depression and anxiety targeting children and adolescents published from 2015 to 2020 and determine if conclusions obtained by the previous studies apply in the same way to new serious games or if this field is evolving. The paper looked for 4 possible applications in mental health: (1) awareness: learning to recognize mental health conditions and erase the stigma associated with help-seeking behavior [24]; (2) prevention: applying emotional regulation strategies in the day-to-day to improve psychological well-being and relationship satisfaction and to avoid mental health issues [25]; (3) detection: finding through game data any mental health conditions or emotional disorders [26]; and (4) therapy: learning CBT techniques to help improve mental well-being and depression or anxiety symptoms for those already diagnosed [27].

XSL•FO
RenderX

Table 1. Comparison of published reviews of serious games for mental health.

Review	Number of studies	Years of publication	Subject and target	Conclusions
Dias et al [16]	28 serious games and gamified apps	2007-2016	Depression care for every possible target	The primary aim of researchers is improving treatment engagement, but there is a lack of effectiveness evaluation.
Barnes and Prescott [17]	5 serious games	2011-2016	Therapy for adolescents with anxiety disorders	The research is limited, but the findings suggest games have the potential to reduce anxiety levels in adolescents.
Villani et al [19]	23 serious and commercial games	2007-2017	Video game effects on emotional regulation and mental health well-being for every possible target	An initial guideline to design serious games for emotional or social abilities and the need for building intervention protocols around commercial games.
Vajawat et al [22]	29 serious and commercial games	2008-2020	ADHD[a], autistic spectrum disorders, eating disorders, posttraumatic stress, impulse control disorders, depression, schizophrenia, dementia, and healthy aging for every possible target	Serious games have great potential but need to explore more applications and targets; review includes the need for standardization of guidelines, more comparison between studies, and the incorporation of VR[b] and artificial intelligence.
David et al [23]	34 serious games	1989-2014	Mental health promotion and health behavioral change for children and adolescents	Serious games are not ready for dissemination as a stand-alone treatment/prevention strategy but have the potential to serve as valuable clinical tools.
Zayeni et al [7]	22 serious and commercial games	2012-2019	Therapeutic and preventive video games for children and adolescents	Video games can be an effective tool for psychotherapy but there is a lack of effectiveness evaluation.

[a]ADHD: attention deficit hyperactivity disorder.

[b]VR: virtual reality.

Methods

The serious games selection process was carried out in 4 phases as summarized in Figure 1: identification, screening, eligibility, and inclusion. The identification through literature search was conducted on the following databases: PubMed, Scopus, Wiley, Taylor and Francis, Springer, PsycINFO, PsycArticles, Web of Science, and Science Direct. The years of publication were limited from 2015 to 2020 as previous reviews had studied serious games with similar purposes prior to these dates. The search terms used were "(serious game OR video game OR applied game OR computer game OR mobile game OR online game OR gaming) AND (children OR adolescent OR childhood OR adolescence) AND (cognitive behavioral therapy OR cognitive training OR anxiety treatment OR anxiety disorder OR mental health OR depression OR stigma OR helping behavior OR meditation)." Bibliographical references of the reviews included in Table 1 were also examined and added to this first selection of studies.

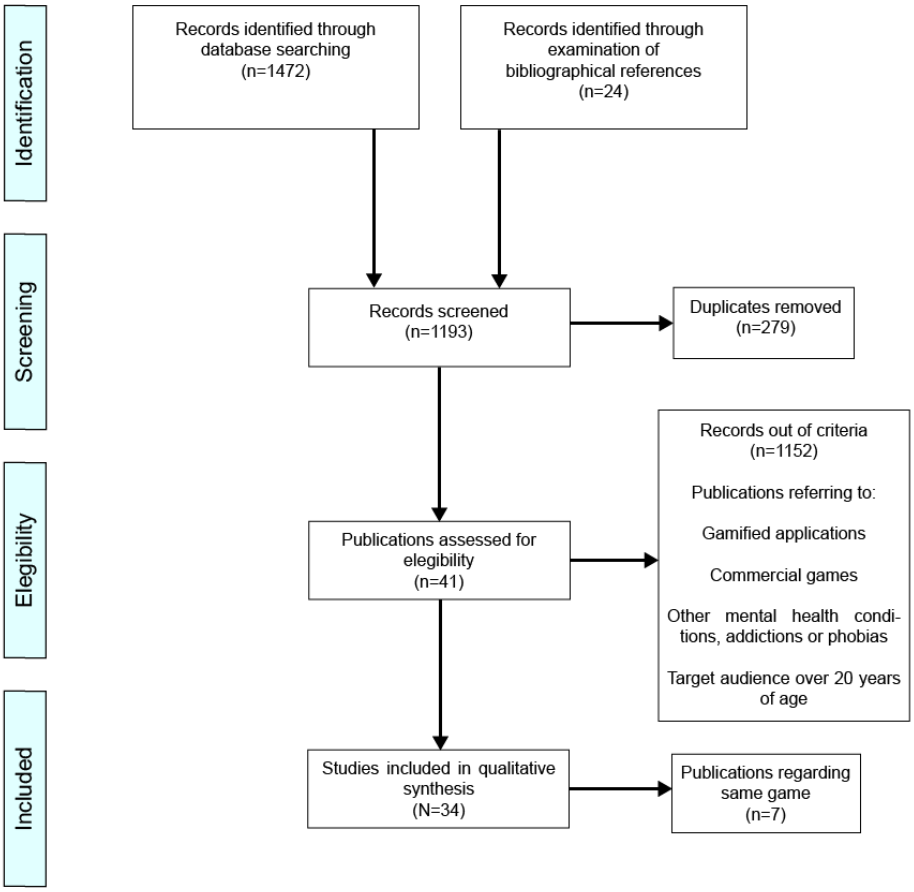
The total of 1966 records were identified through this search, and the bibliographical examination added another 24 studies. The following step, as Figure 1 shows, was removing all duplicates, which excluded 381 articles. The remaining 1609 studies were screened to find those studies that considered serious games for depression and anxiety and possible applications to awareness, prevention, detection, and therapy.

Inclusion criteria include games for PC, smartphone, or VR addressing depression or anxiety developed by research teams for children and adolescents. Exclusion criteria include gamified apps; existing commercial games; games for those aged 20 years and older; and games dedicated to other mental health conditions, addictions, or phobias. These criteria were met by only 41 papers.

In cases where the same research team published multiple articles on the same game with the same application, only the most recent papers were selected as these papers were the most developed by the researchers as they had more time for the design and testing and included the highest number of participants. This excluded 7 papers, so 34 studies were included in the qualitative analysis and listed with characteristics in Multimedia Appendix 1 [18,25,28-59].

XSL•FO
RenderX

Figure 1. Flow diagram of the selection process.



Results

For each of the 34 studies, the following information was extracted: game name, mental health condition addressed, application, target age, mental health purpose, year of publication, game genre, tools used for evaluation, type of game performance, devices used, time required, groups of participants, results, and references. From these data, the results were interrelated to obtain the use of serious games and trends in their development.

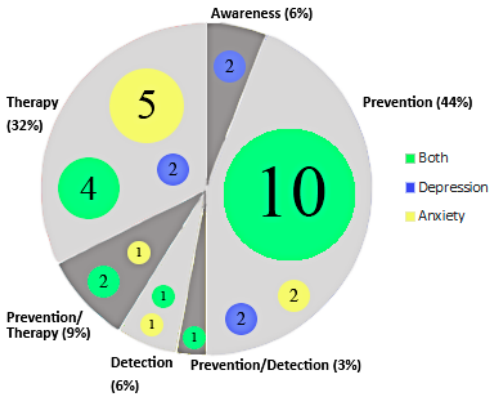
Serious Games Applications

Each game was classified as using 1 of the 4 possible applications contemplated in this review: awareness, prevention, detection, and therapy. The application was also put in relation with the mental health condition to which the game applies, as Figure 2 shows. The majority (30/34, 88%) of serious games serve 1 application. Only 12% (4/34) of games were hybrids mixing 2 applications. The most common (15/34, 44%) application was prevention, followed by therapy (11/34, 32%). On the other hand, awareness and detection were infrequently used, with 2 games each (2/34, 6%). Finally, 9% (3/34) of games used prevention and therapy and 3% (1/34) used prevention and detection.

Figure 2 also shows the relationship between type of application and type of disorder. Of the 34 papers, 26% (9/34) focused on anxiety, 18% (6/34) focused on depression, and 56% (19/34) addressed both issues. Most (10/14, 71%) of the serious games that apply to both mental health conditions were used for prevention. These were games that teach emotional regulation and conflict resolution, skills necessary for a positive mental well-being. It was also common (4/11, 36%) that gaming therapy was equally effective for depression and anxiety since learning CBT was applied for both conditions. Games that exclusively treated depression were equally divided into prevention (2/6, 33%), therapy (2/6, 33%), and awareness (2/6, 33%), which only has cases of this condition. Of the anxiety games, most were used for therapy (5/9, 56%) with the rest for prevention (2/9, 22%), detection (1/9, 11%), and prevention/therapy (1/9, 11%). It stands out that games that treat and prevent anxiety (7/34, 21%) were more numerous than those for depression (4/34, 12%).

XSL•FO
RenderX

Figure 2. Application of games by mental health condition.



Target Age

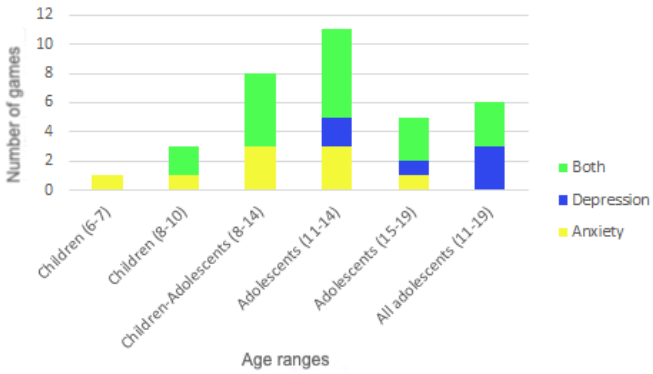
The age range for which these serious games are targeted was 6 to 19 years, but there was uneven distribution (Multimedia Appendix 1). Only 3% (1/34) targeted children aged 6 to 7 years, but 21% (7/34) targeted children aged 8 years. Adolescence was the preferred target group of the reviewed papers, especially ages 11 to 14 years (11/34, 32%). The number of games for adolescents aged 18 to 19 years (4/34, 12%) decreased but was still more numerous than for the youngest kids.

Games were targeted to children according to stages of emotional development [60]: (1) young children aged 6 to 7 years—the earliest stage of emotional immaturity; (2) children aged 8 to 10 years—the first phase of development of empathy and self-recognition; (3) children and young adolescents aged 8 to 14 years—the age range that covers the passage from childhood to adolescence and produces more emotional issues; (4) young adolescents aged 11 to 14 years—the second phase of self-recognition based on comparison with the group, decisive in self-esteem; (5) older adolescents aged 15 to 19 years—self-consolidation stage for the achievement of adult maturity; and (6) all adolescents aged 11 to 19 years—the entire adolescence that includes all self-consolidation and emotional regulation learning.

Age ranges must be considered when addressing mental health conditions (distribution shown in Figure 3). Anxiety was the only condition treated in children aged 6 to 7 years but was mainly seen in children aged 8 to 14 years. Serious games for depression were only targeted at children aged 11 years and older, and half (3/34, 9%) of these games span the entire range of adolescence. Papers dedicated to both conditions applied to all ages (19/34, 56%) and especially children aged 8 to 14 years (11/34, 32%). The trend that Figure 3 shows is that serious games for anxiety were centered in childhood (5/9, 56%) and games for depression in adolescence (6/6, 100%).

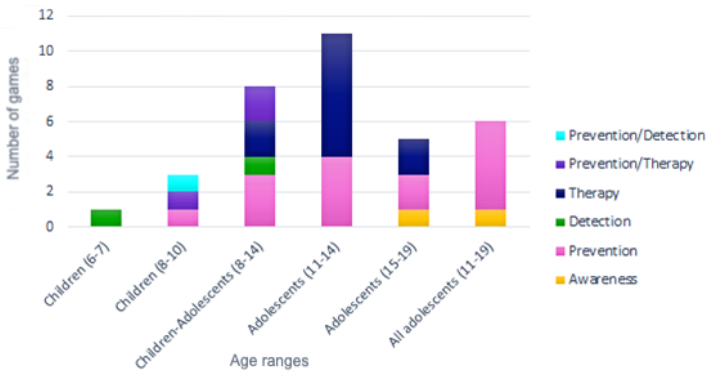
The other feature to study regarding age range was game application, as can be seen in Figure 4. Games for detection (2/34, 6%) were limited to childhood, those aged 6 to 7 years, and for children aged 8 to 14 years. In addition, there was a game using detection and prevention for children aged 8 to 10 years (1/34, 3%). On the other hand, awareness games (2/34, 6%) were only targeted to adolescents. The prevention application focused on all age ranges, with its highest percentage in adolescence (11/15, 73%). Therapy also targeted all children, with its highest use on children aged 11 to 14 years (7/11, 64%). Finally, games that mix prevention and therapy were used only in children aged 8 to 14 years (2/34, 6%).

Figure 3. Mental health condition by age range.



XSL•FO
RenderX

Figure 4. Game application by age range.

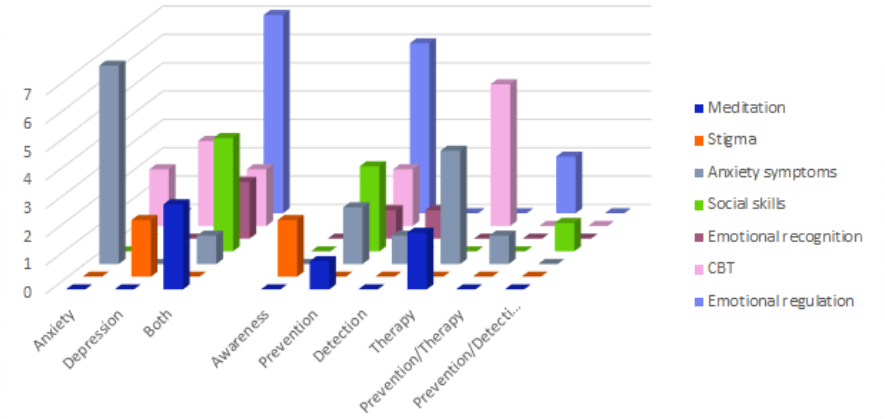


Mental Health Purpose

Mental health purpose is a characteristic that has been obtained from the qualitative analysis of these serious games. The educational mental health issues that these 34 games address have been classified into 7 categories: training of emotional regulation, emotional recognition, and social skills; reduction of anxiety symptoms and stigma against mental health conditions; CBT; and meditation learning. Concerning mental health conditions, Figure 5 relates the purpose of the game to the issue addressed and highlights how training of emotional and social skills and learning of meditation, necessary skills for positive mental well-being, are dedicated to both anxiety and depression. Logically, games that aimed to reduce anxiety symptoms (8/34, 24%) were targeted at anxiety, except for one that targeted both. However, games aiming to reduce stigma (2/34, 6%) only focused on depression. Finally, games that apply CBT techniques (7/34, 21%) were almost evenly distributed between anxiety and depression.

Comparing mental health purpose with game application, as shown in Figure 5, it is possible to discern the trend that training emotional and social skills was used mainly (10/14, 71%) for prevention. On the other hand, therapy was divided between the learning of CBT and meditation (7/11, 64%) and games aiming to reduce anxiety symptoms (4/11, 36%). These mental health purposes also worked for prevention (5/18, 28%), but to a lesser extent. On the other hand, awareness was completely related to stigma reduction (2/2, 100%). Moreover, the use of detection was divided into emotional recognition (1/3, 33%), social skills (1/3, 33%), and anxiety symptoms (1/3, 33%). In conclusion, the most addressed mental health issues were emotional regulation training (8/34, 24%), anxiety symptoms reduction (8/34, 24%), and CBT application (7/34, 21%).

Figure 5. Mental health purpose by its condition and game application. CBT: cognitive behavioral therapy.



Game Genre

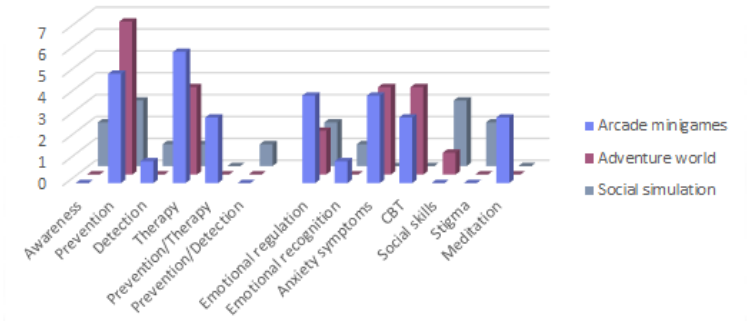
Qualitative analysis of the papers allowed them to be organized into 3 game genres: (1) arcade minigames that have short durability and simple playability with various objectives; (2) social simulations of day-to-day environments and issues to learn optimal resolutions and necessary skills; (3) adventure world in which the player takes on the role of an avatar and must interact with missions, characters, and items. This playability characteristic has been put in relation with the game application in Figure 6 to show how these are adapted to serious games. Arcade minigames were used in 44% (15/34) of games, followed by adventure worlds in 32% (11/34) and social simulations in 24% (8/34). Regarding game application, the main ones, prevention and therapy, were equally distributed between arcade minigames (15% [5/34] and 18% [6/34], respectively), social simulations (9% [3/34] and 3% [1/34], respectively), and adventure worlds (21% [7/34] and 12% [4/34], respectively). However, it should be noted that for awareness, only social simulation (2/2, 100%) was used, and adventure world is not a genre applied for detection.

Figure 6 also interrelates the game genre with mental health purpose to continue clarifying the application of gameplay to mental health learning. It is revealed that emotional regulation (8/34, 24%) was the only one that used all game genres.

XSL•FO
RenderX

Emotional recognition was addressed with genres arcade minigames (1/2, 50%) and social simulation (1/2, 50%) but not adventure world. Social simulations (8/34, 24%) were used to decrease depression stigma but not anxiety symptoms. Furthermore, CBT learning used the genres arcade minigames (3/7, 43%) and adventure world (4/7, 57%), but meditation only used arcade games (3/34, 9%). Finally, social skills training used the genres social simulations (3/4, 75%) and adventure world (1/4, 25%) .

Figure 6. Game genre by game application and mental health purpose. CBT: cognitive behavioral therapy.



Evaluation Tools

Assessing the evaluation of the articles result in 3 important issues: objectives of the paper, methods of data collection, and comparison of data. No more variables in common could be found due to the mix of applications and mental health purposes than these serious games address. The qualitative analysis of the papers resulted in 4 possible goals: increase emotional skills, reduce mental health condition symptoms, evaluate game usefulness, and assess game design. The survey analysis shows that 41% (14/34) of games measured skills, 32% (11/34) tested symptoms, 18% (6/34) surveyed the usefulness of the game, and 9% (3/34) evaluated the experience design. Figure 7 interrelates these objectives with the game application to determine the actual focus of this field research. The graphic reveals that most prevention games measured skills (10/15, 67%) and symptoms (4/15, 27%). Detection was also found in these 2 objectives (1/2 [50%] and 1/2 [50%], respectively), while therapy games measured symptoms (6/11, 54%) or verified that it is an effective tool (3/11, 27%). Awareness games aimed to prove that these are useful tools (2/2, 100%) as well. Finally, the studies that were still in the process of assessing game design corresponded to therapy application (3/3, 100%).

Regarding methods of data collection, the qualitative analysis identified 4 tools used in the articles: normalized questionnaires, self-made questionnaires, game data, and interviews. The most common were normalized questionnaires; 74% (25/34) of papers used them to assess changes after the serious game was played. These questionnaires are proven tools in psychology and education that in each case measured symptoms or emotional capacities as appropriate. Furthermore, 24% (8/34) used self-made questionnaires to test their developed experience; 38% (13/34) analyzed game data, and 12% (4/34) conducted interviews with the participants. Figure 8 shows how the data collection tools were distributed according to publication objectives, so the trend to measure the games was revealed. Questionnaires were the most commonly used tools, and this choice stands out to measure skills and symptoms. It is also remarkable within these objectives that normalized questionnaires (25/34, 74%) were used more often than self-made ones (8/34, 24%), except for surveying usefulness (both 3/6, 50%) and checking game design (both 1/3, 33%) when they were chosen equally. The use of game data was highlighted to measure skills (5/14, 36%) and symptoms (6/11, 54%) but was only used once each for useful tool (1/6, 17%) and design assessment (1/3, 33%). Interviews were the least common method, used twice for symptoms (2/11, 18%) and once each for useful tool (1/6, 17%) and design (1/3, 33%).

Finally, the moment in which the performance data were collected was analyzed: before playing (pre), while playing through the game data (during), and after playing (post). These 3 moments of data collection are combined in the different papers, so there are 5 options: post, during/post, pre/post, pre/during, and pre/during/post. A total of 12% (4/34) of studies performed only postcollection, 6% (2/34) performed during/post comparison, and 9% (3/34) performed the pre/during type. In conclusion, most of the games compared data before and after the serious game. The most common combination (17/34, 50%) was the pre/post evaluation, and 23% (8/34) of studies performed all possible measurements: pre/during/post. Figure 9 shows the relationship of the paper objective with the type of data comparison to get the trend of every measure collection. To measure skills, the most common (10/14, 71%) was a pre-post evaluation, although there were also a few cases of pre/during (2/14, 14%) and pre/during/post comparison (2/14, 14%). The pre/during/post comparison (6/11, 54%) was the most used for symptom measuring, followed by pre/post (4/11, 36%). For the useful tool and design testing, the types of comparisons chosen were similar, but there is no case of pre/during/post evaluation.

XSL•FO
RenderX

Figure 7. Paper objective by game application.

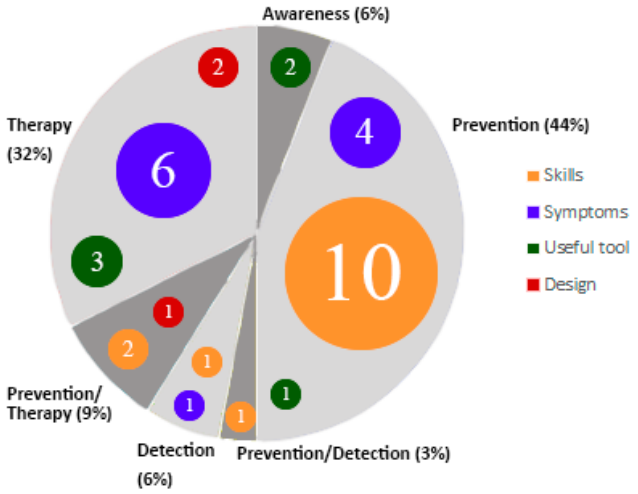


Figure 8. Paper objective by type of evaluation.

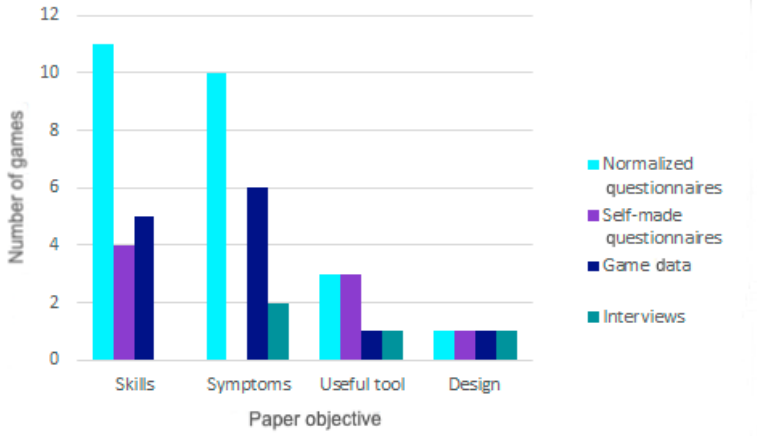
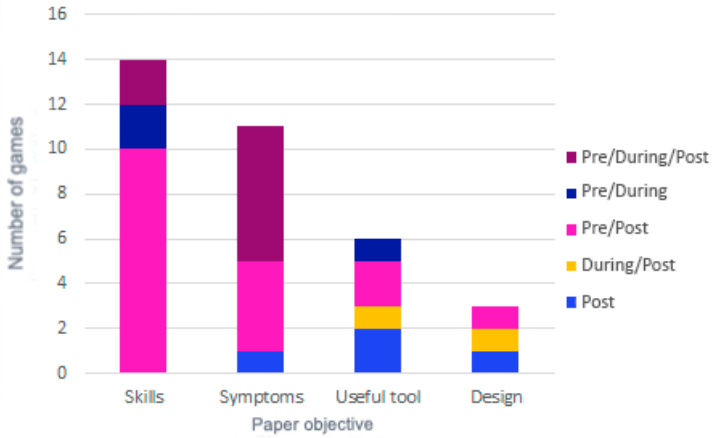


Figure 9. Paper objective by type of data comparison.



Type of Game Performance

All papers in this review proved the serious games in different ways in their tests so the participants could achieve the set goal. Qualitative analysis discovered 3 parameters that must be taken into account in the performance data: game was used as the only tool to learn so the participants achieve the mental health purpose, participants received support during the game through the gameplay or research assistants, and participants and research assistants performed a debriefing after playing in which participants and educators or psychologists discussed what was played to draw conclusions. The analysis revealed that most of the serious games were performed as the only tool, support was not offered, and debriefing was not performed. Figure 10 shows the distribution of these performance parameters by game application to determine implementation differences depending on the goal of the game. It is important to note that only therapy games (6/8, 75%) used the game as a back tool. Regarding support or lack of it, all applications were proportional in both

XSL•FO
RenderX

cases. However, games that did not perform a debriefing were more common in therapy applications (12/14, 86%).

The performance of serious games within the different age ranges is shown in Figure 11 to determine if the games were planned differently depending on the target. The use of the game as an only tool was proportional to the number of games in each range. Support during the game was offered to children (6/12, 50%) more often than to adolescents (5/22, 23%). Debriefing was not offered after any game in children aged 10 years and younger. Debriefing was offered after some games for adolescents aged 11 to 14 years (5/19, 26%) and was much more common (6/11, 54%) for older adolescents.

Figure 10. Game performance by game application.

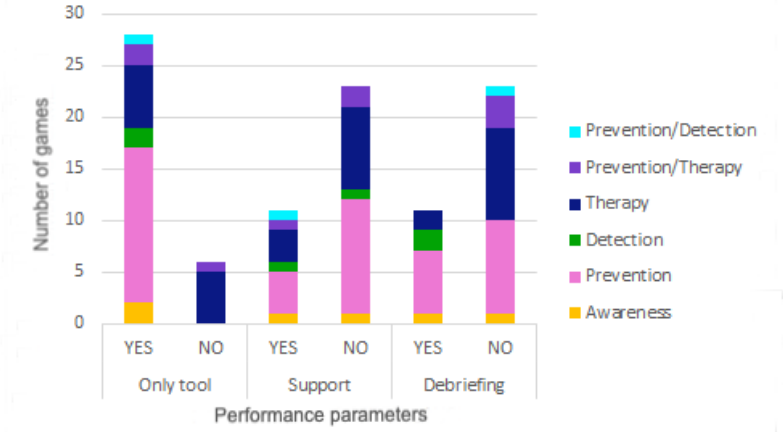
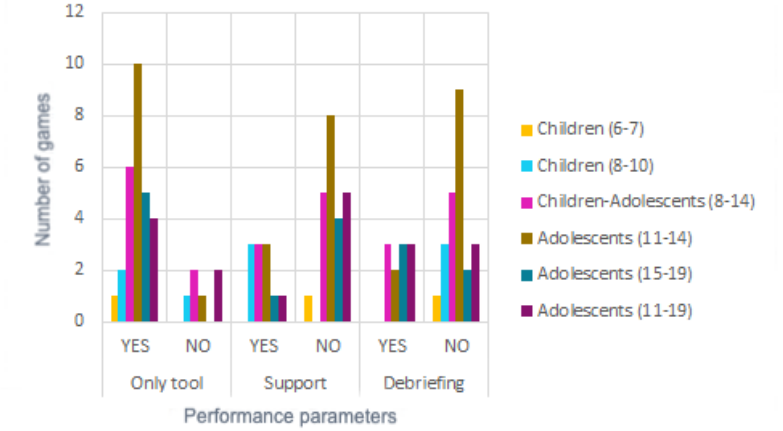


Figure 11. Game performance by age range.



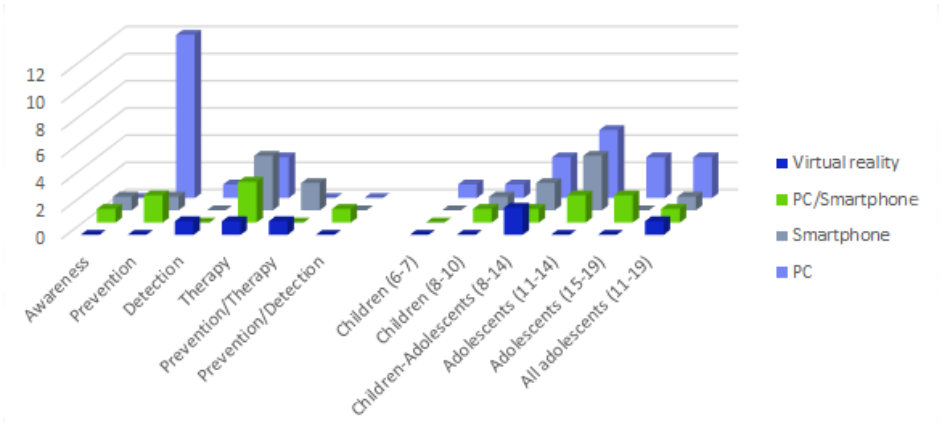
Game Devices

Games were developed for PC (16/34, 47%), smartphone (8/34, 23%), PC and smartphone (7/34, 21%), and VR (3/34, 9%). This technical aspect has been related to game application in Figure 12 to reveal potential development trends based on the objective. The most notable trends are that the majority (12/16, 75%) of PC games were dedicated to prevention and most (4/8, 50%) smartphone games to therapy. On the other hand, awareness (2/2, 100%) was mainly used on smartphones (1/2, 50%) and detection (1/2, 50%) on PC and VR.

Figure 12 shows the choice of device according to the targeted age range to discern any possible difference by target maturity. The result is that serious games for children aged 10 years and younger were developed for PC and smartphone equally. For the age range 8 to 14 years, development for PC was slightly preferred (4/8, 50%); in addition, most VR games (2/3, 66%) were targeted to this group. Games were developed for PC more often (16/22, 73%) for adolescents; in fact, no games for those aged 15 to 19 years were developed exclusively for smartphones.

XSL•FO
RenderX

Figure 12. Device used by game application and age range.



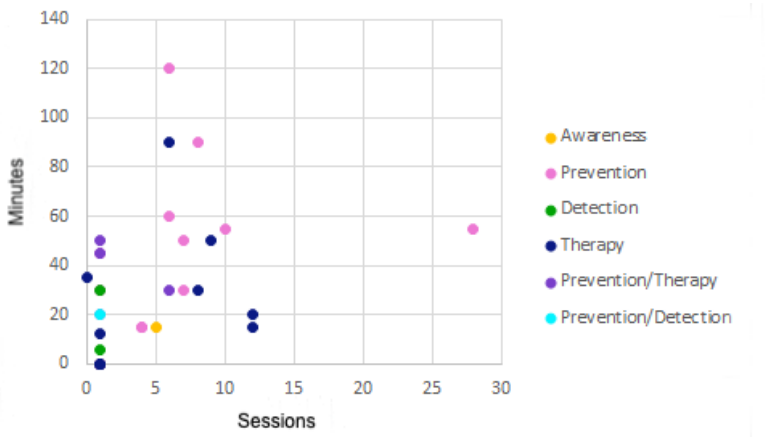
Game Length

The game tests were performed in different ways with respect to time: 26% (9/34) of games used a single session, 50% (17/34) were played in several sessions, and 24% (8/34) of studies did not indicate the playing time. Among the games that used several sessions, the average was 8.1 (range 4-20). The largest number of sessions was a singularity, since the rest of the games were within the range of 4 to 12 sessions. Regarding the session duration, in which all available data has been included, the average was 39.7 minutes. The shortest test was 5.5 minutes and the longest 120 minutes. Data do not always refer to strict playing time as there were tests in which a debriefing was performed and that process was included in this time.

The distribution of number of sessions and minutes by game application is shown in Figure 13 to determine possible trends of performance and objective. Any study that did not specify the duration time or number of sessions has been given a value of 0 on its corresponding axis. There is a trend whereby the games with the highest number of sessions and the longest length correspond to prevention, followed by therapy. On the other hand, all detection games included 1 game session of less than 40 minutes. Finally, of the 2 awareness games, one does not give time data and the other did 5 sessions of 15 minutes.

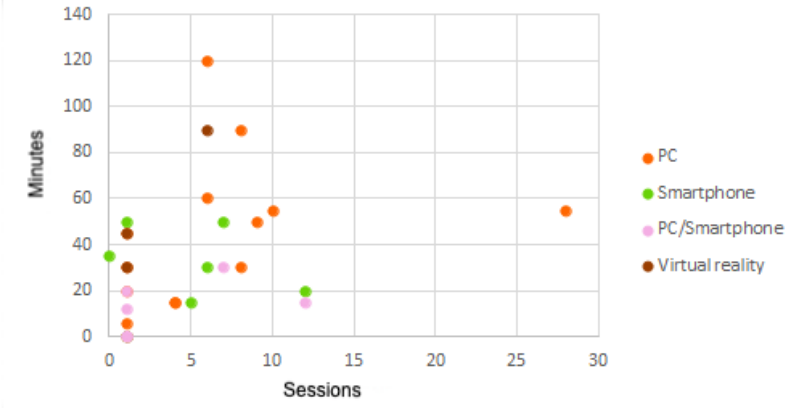
The time spent with the device used has been interrelated in Figure 14 to show the possible relationships of time played with technical development. Games that required the highest number of sessions and most length were most often PC games. All games for smartphone or smartphone and PC took less than an hour, although the number of sessions was highly variable. Most VR games (2/3, 66%) used only 1 session of less than 50 minutes.

Figure 13. Required time by game application.



XSL•FO
RenderX

Figure 14. Required time by device used.

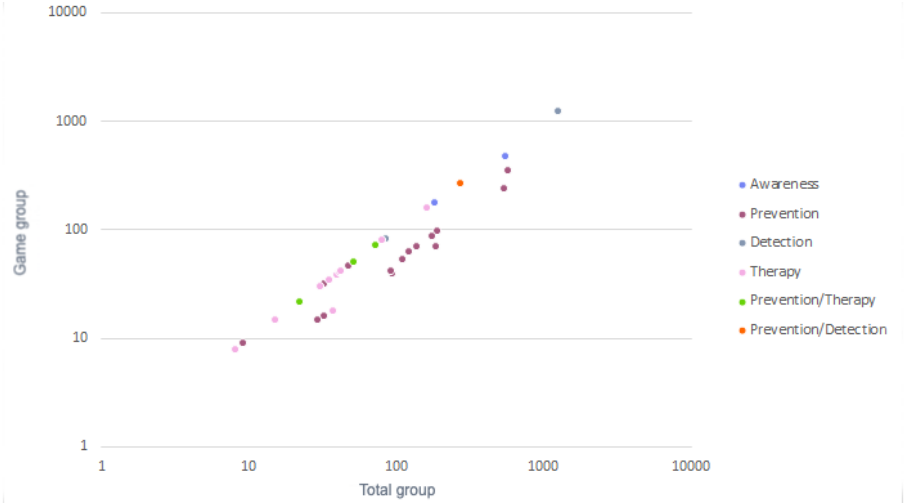


Groups of Participants

All games in this review were tested with groups of participants ranging from 8 to 1236 people. The last case was an exception because the game with next highest number had 574 children. The vast majority (29/34, 85%) had fewer than 200 participants, and the average was 156.79. Of the games, 38% (13/34) included a control group while 62% (21/34) evaluated only the group that played the game. The relationship between the total group of participants, the game group, and game application can be seen in Figure 15. The objective was to find a trend between the aim of the games and the number of participants in their tests, in addition to the division between game and control groups.

There was a big difference between therapy groups, which averaged 45.54 participants, and prevention groups, which averaged 157.53 participants. Detection and awareness applications were also tested with large groups. Furthermore, the representation of so many points on the diagonal between both axes shows that most games (20/34, 59%) were tested with the total number of participants. On the contrary, there was a trend in which prevention games (12/15, 80%) usually used control groups.

Figure 15. Total groups and game groups by game application.



Discussion

Present Trends and Future Lines

Implementation of Serious Games

Most games identified treat both anxiety and depression. In addition, there were more games with prevention and therapy applications. This trend is logical since depression and anxiety are closely related and they can lead to each other. Likewise, many therapeutic and emotional management techniques are equally valid for both conditions [61]. Taking a closer look at the results, more games address prevention than therapy. It is true that the main focus of education and psychology should be to avoid suffering mental health issues. Furthermore, these tools can be applied to many participants in school or high school classes. On the other hand, the application of therapy games is smaller since these are logically reserved for children and adolescents already diagnosed.

Continuing with the applications, results reveal that more awareness and detection games are needed as there are only 2 each. It is also noted that, so far, there are only awareness games about depression, probably because these recent games have followed an emerging line of research that fosters emotional awareness to prevent adolescence depression [62]. Ideally, both applications should be mixed since the lack of emotional awareness is equally associated with symptoms of depression and anxiety at these ages [60]. Another option would be to add detection to the development of prevention games. When emotional regulation techniques are being taught, game data

XSL•FO
RenderX

can be collected to detect any possible behavior problem or lack of emotional recognition. Likewise, adding awareness to prevention games teaches maintaining a positive well-being while recognizing its importance and erasing mental health stigma.

Another relevant interrelation found in the games implementation is age and mental health condition. Most of these games are dedicated to adolescence but should be applied earlier as children begin to develop depression and anxiety from the age of 3 years [63]. There is also a trend that games for depression are exclusively addressed to adolescence and games for anxiety exclusively to childhood. Once again, this bias may follow the current line of research that focuses depression on adolescents [62], while research on anxiety in children is becoming more relevant [64]. The prevalence of anxiety is as relevant in adolescents as it is in children [61], so this implementation should be fair. Regarding the application, it stands out that awareness games are aimed at adolescents and detection games at children. Given that lack of emotional awareness can foster anxiety and depression symptoms at an early age [64], there should be no differences. Likewise, detection games should be extended to adolescence to avoid the serious development of these mental health conditions [62].

Finally, there is a bias concerning mental health purposes and applications of these serious games. Games that teach emotional regulation, emotional recognition, and social skills are those of prevention, since these skills serve to avoid both conditions [60]. On the other hand, therapy games focus on reducing anxiety symptoms and practicing CBT and meditation. Even so, all these purposes can be used for prevention or therapy. Therefore, a future line should be that more games serve prevention and therapy together, since there are only 3 of the 34 that do this. The only exception is stigma reduction, which is applied only in awareness games. It should also be noted that no game tested depressive symptoms on their own, so this would be an open research field.

Development of Serious Games

The game genres developed can be classified into arcade minigames, adventure worlds, and social simulations. Arcade minigames are used most often because the results are short games, easier to adjust to the implementation conditions. Since most tests are finished in an hour, these games can be repeated several times until the objective is achieved, also presenting a challenge so the user does not get bored. The next most popular genre is adventure world, which take advantage of the important ability of serious games to entertain so the learning is acquired. When a plot is played with a main character the user identifies with, this allows the recognition and assimilation of its emotions. Social simulations are used to generate empathy with characters and familiar environments that the users might experience in their real life. In this way, users can be directly taught how mental health conditions affect the lives of people who are diagnosed with them [20].

Regarding the interrelation with applications, awareness games use only social simulations to show the reality of mental health conditions to the participants. Adventure world is not a genre that is used for detection, probably because these are more playful games that focus on the character and not the player, so it is not possible to detect anything. Exploring mental health issues, emotional regulation is the most applicable to the 3 game genres. The techniques applicable in this learning are diverse to adapt them to many types of gameplay. Emotional recognition games do not develop adventure worlds as these do not seem to offer enough first-person interactivity. In addition, social simulation is not a genre in which anxiety reduction or CBT is applied. It is logical because this game genre is based on telling stories, although it is the most ideal for social skills and stigma reduction. Last, meditation is only taught with short arcade minigames as these techniques are performed the same in nondigital scenes.

Concerning devices tendencies, therapy games often take advantage of smartphone facilities. Most of these games have been designed to be played on a day-to-day basis, so it is more comfortable to use a portable device. On the other hand, prevention games are mostly performed in school classrooms where it is easier to access PCs. The way in which the experience is going to be introduced should be assessed when choosing the serious game device to save resources. In any case, it does not seem that the age at which the game is targeted affects this choice. Finally, future research should focus on VR because this review has only found 3 serious games in the recent years. This technology has proven to achieve positive results for learning and measuring user behavior [65].

Another important factor in game development is playing time. The vast majority of the analyzed games are played within 1 hour. This length responds to the duration of a class or a therapy session in which it is usually applied. Games with therapy and prevention applications tend to be longer because participants need more time to learn these techniques. It is also logical that longer games are developed for PC because the device is more comfortable than for smartphone or VR games. For VR, the duration is limited by the lack of familiarity with this device, which can cause dizziness or visual fatigue [65]. Therefore, the decision with this variable must be made according to the developmental capacities of the research team and its application.

It should be noted that certain data were searched for without success: video game development engine, software, applied game design, and availability. The lack of this more technical data was probably due to author specialization. Most of the authors are researchers in education or psychology who had not developed the game themselves. On the other hand, almost every study indicated a multidisciplinary team created the game but did not follow a specific game design. Finally, the search included the free public access to the developed game. Only 3 of the 34 games were available; the vast majority of articles did not indicate this aspect or pointed to websites that were not working. Since one of the great limitations of these studies is the lack of participants [18], it would be important to disclose and give access to serious games so psychologists, educators, or parents could try them. In this way, more data would be obtained, the game design would be assessed, and the ability of serious games to raise awareness, prevent, detect, or help within therapy would increase.

XSL•FO
RenderX

Evaluation of Serious Games

In this review, the game performance for the participants has been analyzed. It is highlighted that therapy games are not used as stand-alone tools because these people are already diagnosed and need complete treatment. Regarding the support or debriefing implementation, there is no statistical difference with respect to other applications. However, since the previous reviews find a lack of current and conclusive results of the usefulness of serious games [7,16], it may be better that all applications are not used as stand-alone tools. Serious games should serve as an appealing approach within an educational program for prevention, detection, and awareness of mental health conditions. Ravyse et al [66] determined in their review of different game designs that it was better to offer support through gameplay and perform a debriefing after the game. Therefore, all serious games should aim to offer this complete experience for further learning. Age should not matter, as has been seen in the analysis of the results, although it should be taken into account that children may need more help than adolescents.

Looking at the participant groups, prevention, detection, and awareness games are tested with the largest groups and therapy games with the smallest. It is a logical consequence of participant availability for these evaluations because the first 3 applications are usually performed in schools in which many students can participate. Therapy games, on the other hand, are played with small groups that have had depression or anxiety diagnosed.

Following with tools used for evaluation, most studies measure skills to test if participants improve or symptoms decrease after playing. These 2 evaluations are the most applied for prevention and detection since these verify the learning of techniques or if the game helps to improve their emotional state. Negative postcollection results would indicate if the user has an emotional issue that needs to be addressed. However, therapy games only measure symptoms because these games are used strictly in populations after diagnosis. In addition, awareness and therapy applications are the only games that verify serious games are a useful tool; awareness learning evaluates an immediate change of judgment, and therapy involves a process evaluated by a psychologist.

To measure these capacities, questionnaires are the most widely used tools, especially normalized ones. It is logical that proven tools are chosen, which also provides the opportunity to compare the results of a serious game with other nondigital programs. Even so, this review has not found that the same questionnaire models are used even if the same symptoms or skills are measured. As pointed out in previous studies [22], this should be standardized in the future, allowing for comparison of results between different game genres, age targets, group numbers, playing lengths, and even applications. So far, the evaluation of serious games is so different that a comparison of effectiveness is not possible. Furthermore, game data are evaluated in less than half of the studies, a circumstance that should change. Serious games should be designed to collect data on player behavior and choices. This would be the most objective measure, as questionnaires are affected by the subjectivity of the participant. Last, interviews are a good method of obtaining data, although many times these cannot be performed because of the time they take. It will depend on the researcher's capacity, but these cases are usually limited to therapy.

Regarding data comparison, 73% of articles evaluated before and after playing. This should be the standard to measure any changes caused by the game playing. Ideally, researchers should collect all types of data: pre-post with questionnaires or interviews and during the experience through game data. This would alleviate one of the problems found in these studies: missing data that prevent clearer results from being obtained [23]. At present, most games that measure symptoms already prefer this pre/during/post comparison. It makes sense that these are the ones that should most closely measure whether the player has a mental health condition. Finally, a summary of these trends and future lines is shown in Table 2.

XSL•FO
RenderX

Table 2. Summary of trends and future lines of serious games.

Analyzed processes in this serious games review
Implementation
Trends
The majority of games treat both depression and anxiety, with prevention and therapy being the most common applications.
Exclusive games for depression are targeted to adolescence and anxiety games to childhood.
Prevention and therapy games have the widest range of learning: emotional and social skills, meditation, cognitive behavioral therapy, and anxiety symptom reduction.
Future lines
More awareness and detection games are needed, and the 4 applications should be mixed to increase the effect.
Anxiety and depression should be equally addressed in all age ranges.
There are still no games that test depressive symptoms.
Development
Trends
The most common game genres are arcade minigames, adventure worlds, and social simulations, in this order.
Arcade minigames and adventure worlds can serve most types of learning, while social simulations are more valuable to teach social skills and stigma reduction.
Therapy games usually use smartphones and prevention games use PC, regardless of the target age.
Most games are played within 1 hour, but prevention and therapy games can be longer.
Future lines
There should be more games for VR to take advantage of the learning capabilities this device offers.
Game design should be standardized to better develop these games.
These serious games should be available to the general public to assess design and development.
Evaluation
Trends
Therapy games are the only ones not used as a stand-alone tool.
Prevention, detection, and awareness games are tested with the largest groups, and therapy games with the smallest.
Most studies measure skills improvement or symptom reduction in participants, commonly using normalized questionnaires.
Most studies collect and compare participant data before and after playing, which should be the standard.
Future lines
Serious games should not be the only tool for the learning and should always offer support and perform a debriefing.
Studies should use standardized questionnaires so the results could be compared between studies.
Games should be developed so playing data could be collected and analyzed to understand participant behavior and learning.

Comparison With Previous Reviews

One of the conclusions find in previous studies is that the research is limited [17]. The authors of this paper agree with this statement. This review has focused on finding all serious games published between 2015 and 2020, with a broad focus on their application for depression and anxiety in children and adolescents. The number of studies found was 34, which is not a large number. Previous reviews have analyzed between 5 and 34 articles but also included commercial games. Although the results show that more games are being developed each year, there is still much pending research. Nevertheless, this is a relatively new research field that largely depends on the technological advances of recent years. Therefore, it is expected that this development will be strengthened from now on. There are more and more free resources and learning tools that can be applied to research, facilitating workloads.

A lack of standardization of development guidelines has been encountered too [22]. This paper reaches the same conclusion by finding that no article followed a game design theory; all claimed to develop it with a multidisciplinary team. This is not to say that these experts have not used this theory but shows it is not standardized to create more games or compare studies. The review by Vajawat et al [22] affirmed the need for VR development and the incorporation of artificial intelligence into games. Our study has only found 3 VR serious games, a technology that should be studied more for its educational capabilities. In addition, the more artificial intelligence is applied

XSL•FO
RenderX

to these games, the better the experiences of children and adolescents can be measured to obtain gameplay data.

Another issue in this research field is the lack of effectiveness evaluations [7,16]. This study found that all articles when testing the serious game found positive results regarding the experience purpose. However, it was impossible to compare and assess these outcomes for several reasons. To begin with, these serious games were dedicated to depression, anxiety, or both conditions, measuring completely different skills or symptoms with different evaluation tools. In addition, the participant groups vary too much in number and conditions to be compared. Because of this, the same type of evaluation must be used to determine effectiveness. In addition, many of the articles confirmed the necessity to do more tests with larger groups taking into account more external aspects that may affect the results.

Only 3 of the 34 games included in this review obtained slight differences in their tests with respect to age. These results are strange since there are games that targeted wide age ranges, from 8 to 14 years or from 11 to 19 years. These are stages in which emotional maturation is critical [60]. Therefore, when therapy techniques are applied or social and emotional skills are learned, older children should have more success as they have greater development. The difference between ages should be one of the future research lines.

Last, there is a claim that serious games are not yet ready for dissemination as an independent treatment/prevention strategy [23]. Having already discussed the research issues of limited studies and lack of standardization in development and evaluation, the authors of this study agree. Nevertheless, results are positive and serious games development is experiencing great advances. In addition, most of these games are developed in teams with increasing experience, so their applications are constantly improving. The next few years will be decisive in implementing serious games such as treatment and prevention for depression and anxiety in childhood and adolescence.

Conclusions

This review studies 34 serious games published between 2015 and 2020 that target children and adolescents to address depression and anxiety. Analyzing their applications has shown a trend that most games are applied on both conditions for different age ranges. Even so, the exclusive anxiety games focus on childhood and the exclusive depression games focus on adolescence. Across all games, most apply prevention, which is usually played on a PC and tested in the largest groups, together with detection and awareness games. The second most used application is therapy, usually on smartphones, not used as a stand-alone tool, and tested in smaller groups. Prevention and therapy games tend to require sessions of 1 hour to complete; longer tests would use a PC. Regarding evaluation, most tests measure the acquisition of skills or the reduction of symptoms. The most used evaluation tools are normalized questionnaires completed before and after playing.

Furthermore, this paper agrees with the conclusions of previous reviews. This research is still limited, although more games are being developed and technological development will increase. A standardized game design theory is needed so all serious game developments follow comparable guidelines. Evaluation tools should be standardized to improve effectiveness evaluation, and studies should use larger groups of participants. Moreover, there is a lack of learning comparison within the age ranges. Once these games are tested, they should be accessible to the public for its benefit and to obtain more data for research. Resolution of these issues in the coming years will determine the implementation of serious games as treatment and prevention tools.

Finally, this review proposes future lines of serious games development after analyzing the missing applications. More awareness and detection games are necessary; in addition, the ideal way to implement these applications would be by mixing them. Another possible mix of these applications would be with prevention, and more games addressing prevention and therapy together are needed. Furthermore, there is a lack of detection games that measure depressive symptoms. All applications should be equally distributed by age, especially since awareness games for children and detection games for adolescents are missing. Regarding devices, it would be interesting to develop more games in VR due to its demonstrated capacity for learning and user behavior measurement. In addition, within game implementation, it is optimal to offer support to the participants and perform a debriefing. Last, when evaluating learning, normalized questionnaires and game data should be used to compare before, during, and after playing.

Acknowledgments

This work was partially supported by grants INVESTUN/18/0002 and FPU18/04688 from the GruaRV project of the Consejería de Empleo of the Junta de Castilla y León (Spain) and the Ministry of Science, Innovation, and Universities.

Authors' Contributions

KM and AB were responsible for conceptualization. AB was responsible for methodology. MIMM was responsible for formal analysis. KM, MIMM, and AB were responsible for funding acquisition. KM was responsible for investigation. AB was responsible for project administration. MIMM and AB were responsible for supervision. KM was responsible for writing and original draft. MIMM and AB were responsible for writing, review, and editing. All authors read and agreed to the published version of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of 34 serious games analyzed in this review.
[DOCX File , 30 KB - games_v9i4e30482_app1.docx]

References

1. Mental health. World Health Organization. 2019. URL: https://www.who.int/health-topics/mental-health#tab=tab_2 [accessed 2021-02-18]
2. Depression and other common mental disorders: global health estimates. World Health Organization. 2017. URL: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf [accessed 2021-02-18]
3. Zyda M. From visual simulation to virtual reality to games. Computer 2005 Sep;38(9):25-32. [doi: 10.1109/mc.2005.297]
4. Catalano CE, Luccini AM, Mortara M. Guidelines for an effective design of serious games. Int J Serious Games 2014 Feb 03;1(1):1. [doi: 10.17083/ijsg.v1i1.8]
5. Cannon-Bowers J, Bowers C, Procci K. Using video games as educational tools in healthcare. In: Tobias S, Fletcher J, editors. Computer Games and Instruction. Charlotte: IAP Information Age Publishing; 2011:47-72.
6. Papastergiou M. Exploring the potential of computer and video games for health and physical education: a literature review. Comput Educ 2009 Nov;53(3):603-622. [doi: 10.1016/j.compedu.2009.04.001]
7. 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]
8. Li J, Theng Y, Foo S. Game-based digital interventions for depression therapy: a systematic review and meta-analysis. Cyberpsychol Behav Soc Netw 2014 Aug;17(8):519-527 [FREE Full text] [doi: 10.1089/cyber.2013.0481] [Medline: 24810933]
9. Fleming TM, Cheek C, Merry SN, Thabrew H, Bridgman H, Stasiak K, et al. Juegos serios para el tratamiento o la prevención de la depresión: una revisión sistemática. Revista de Psicopatología y Psicología Clínica 2015 Jan 15;19(3):227. [doi: 10.5944/rppc.vol.19.num.3.2014.13904]
10. 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]
11. Lau HM, Smit JH, Fleming TM, Riper H. Serious games for mental health: are they accessible, feasible, and effective? A systematic review and meta-analysis. Front Psychiatry 2016 Jan;7:209 [FREE Full text] [doi: 10.3389/fpsyt.2016.00209] [Medline: 28149281]
12. Shah A, Kraemer KR, Won CR, Black S, Hasenbein W. Developing digital intervention games for mental disorders: a review. Games Health J 2018 Aug;7(4):213-224. [doi: 10.1089/g4h.2017.0150] [Medline: 30106642]
13. Nunnari F, Magliaro S, D'Errico G. Designing and assessing interactive virtual characters for children affected by ADHD. Virt Reality Aug Reality 2019:285-290. [doi: 10.1007/978-3-030-31908-3_17]
14. Barba M, Covino, de Luca. BRAVO: a gaming environment for the treatment of ADHD. Lect Notes Comput Sci 2019:394-407. [doi: 10.1007/978-3-030-25965-5_30]
15. Botella C, Fernández-Álvarez J, Guillén V, García-Palacios A, Baños R. Recent progress in virtual reality exposure therapy for phobias: a systematic review. Curr Psychiatry Rep 2017 Jul 24;19(7):42. [doi: 10.1007/s11920-017-0788-4] [Medline: 28540594]
16. Dias LPS, Barbosa JLV, Vianna HD. Gamification and serious games in depression care: a systematic mapping study. Telematics Informatics 2018 Apr;35(1):213-224. [doi: 10.1016/j.tele.2017.11.002]
17. Barnes S, Prescott J. Empirical evidence for the outcomes of therapeutic video games for adolescents with anxiety disorders: systematic review. JMIR Serious Games 2018 Feb 28;6(1):e3 [FREE Full text] [doi: 10.2196/games.9530] [Medline: 29490893]
18. Barnes S. Gaming against anxiety: user-feedback in the development of a therapeutic game for adolescents with anxiety disorders. 2019 Presented at: 9th World Congress of Behavioural and Cognitive Therapies (WCBCT); 2019; Berlin. [doi: 10.17605/OSF.IO/7ECPS]
19. Villani D, Carissoli C, Triberti S, Marchetti A, Gilli G, Riva G. Videogames for emotion regulation: a systematic review. Games Health J 2018 Apr;7(2):85-99. [doi: 10.1089/g4h.2017.0108] [Medline: 29424555]
20. Cheng VWS, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. JMIR Ment Health 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
21. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious Games and Gamification for Mental Health: Current Status and Promising Directions. Front Psychiatry 2016 Jan;7:215 [FREE Full text] [doi: 10.3389/fpsyt.2016.00215] [Medline: 28119636]

XSL•FO
RenderX

22. Vajawat B, Varshney P, Banerjee D. Digital gaming interventions in psychiatry: evidence, applications and challenges. Psychiatry Res 2021 Jan;295:113585. [doi: 10.1016/j.psychres.2020.113585] [Medline: 33303223]
23. David OA, Costescu C, Cardos R, Mogoaşe C. How effective are serious games for promoting mental health and health behavioral change in children and adolescents? A systematic review and meta-analysis. Child Youth Care Forum 2020 Jul 31;49(6):817-838. [doi: 10.1007/s10566-020-09566-1]
24. Aboujaoude E, Salame W, Naim L. Telemental health: a status update. World Psychiatry 2015 Jun 04;14(2):223-230 [FREE Full text] [doi: 10.1002/wps.20218] [Medline: 26043340]
25. Balzarotti S, Biassoni F, Villani D, Prunas A, Velotti P. Individual differences in cognitive emotion regulation: implications for subjective and psychological well-being. J Happiness Stud 2014 Oct 18;17(1):125-143. [doi: 10.1007/s10902-014-9587-3]
26. Lázaro E, Amayra I, López-Paz J, Martínez O, Pérez Alvarez M, Berrocoso S, et al. Using a virtual serious game (Deusto-e-motion1.0) to assess the theory of mind in primary school children: observational descriptive study. JMIR Serious Games 2020 Apr 02;8(2):e12971 [FREE Full text] [doi: 10.2196/12971] [Medline: 32238330]
27. Ceranoglu TA. Star Wars in psychotherapy: video games in the office. Acad Psychiatry 2010 Apr 29;34(3):233-236. [doi: 10.1176/appi.ap.34.3.233] [Medline: 20431107]
28. Boendermaker WJ, Veltkamp RC, Peeters M. Training behavioral control in adolescents using a serious game. Games Health J 2017 Dec;6(6):351-357. [doi: 10.1089/g4h.2017.0071] [Medline: 28926286]
29. Bossenbroek R, Wols A, Weerdmeester J, Lichtwarck-Aschoff A, Granic I, van Rooij MMJW. Efficacy of a virtual reality biofeedback game (DEEP) to reduce anxiety and disruptive classroom behavior: single-case study. JMIR Ment Health 2020 Mar 24;7(3):e16066 [FREE Full text] [doi: 10.2196/16066] [Medline: 32207697]
30. Cangas AJ, Navarro N, Parra JMA, Ojeda JJ, Cangas D, Piedra JA, et al. Stigma-stop: a serious game against the stigma toward mental health in educational settings. Front Psychol 2017 Aug 21;8:1385 [FREE Full text] [doi: 10.3389/fpsyg.2017.01385] [Medline: 28878702]
31. 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]
32. Carrasco AE. Acceptability of an adventure video game in the treatment of female adolescents with symptoms of depression. ResPsy 2016 Apr 18;19(1):10-18. [doi: 10.4081/ripppo.2016.182]
33. Cejudo J, Losada L, Feltrero R. Promoting social and emotional learning and subjective well-being: impact of the "Aislados" intervention program in adolescents. Int J Environ Res Public Health 2020 Jan 17;17(2):609 [FREE Full text] [doi: 10.3390/ijerph17020609] [Medline: 31963598]
34. Cheek C, Fleming T, Lucassen MF, Bridgman H, Stasiak K, Shepherd M, et al. Integrating health behavior theory and design elements in serious games. JMIR Ment Health 2015 Apr 21;2(2):e11 [FREE Full text] [doi: 10.2196/mental.4133] [Medline: 26543916]
35. Cheong Y, Khaled R, Yannakakis G. Serious games for teaching conflict resolution: modeling conflict dynamics. Conflict Multimodal Comm 2015:449-475. [doi: 10.1007/978-3-319-14081-0_21]
36. Christie GI, Shepherd M, Merry SN, Hopkins S, Knightly S, Stasiak K. Gamifying CBT to deliver emotional health treatment to young people on smartphones. Internet Interv 2019 Dec;18:100286 [FREE Full text] [doi: 10.1016/j.invent.2019.100286] [Medline: 31890633]
37. D'Amico A. The use of technology in the promotion of children's emotional intelligence: the multimedia program "Developing Emotional Intelligence". Int J Emot Educ 2018;10(1):47-67 [FREE Full text]
38. David OA, Cardo RAI, Matu S. Is REThink therapeutic game effective in preventing emotional disorders in children and adolescents? Outcomes of a randomized clinical trial. Eur Child Adolesc Psychiatry 2018 Jul 10;28(1):111-122. [doi: 10.1007/s00787-018-1192-2]
39. DeRosier ME, Thomas JM. A digital game for social skills training with young adolescents. Int J Comput Games Technol 2019 Jan 01;2019:1-12. [doi: 10.1155/2019/6981698]
40. Filella G, Cabello E, Pérez-Escola N, Ros A. Assessment of the emotional education program “Happy 8-12” for the assertive resolution of peer conflicts. Electr J Res Educ Psychol 2016 Dec 01;14(3):582-601. [doi: 10.14204/ejrep.40.15164]
41. Lobel A, Gotsis M, Reynolds E. Designing and utilizing biofeedback games for emotion regulation. Proc 2016 CHI Conf Ext Abstr Human Fact Comput Syst 2016:1945-1951. [doi: 10.1145/2851581.2892521]
42. Ong JG, Lim-Ashworth NS, Ooi YP, Boon JS, Ang RP, Goh DH, et al. An interactive mobile app game to address aggression (RegnaTales): pilot quantitative study. JMIR Serious Games 2019 May 08;7(2):e13242 [FREE Full text] [doi: 10.2196/13242] [Medline: 31066682]
43. Patwardhan M, Stoll R, Hamel D. Designing a mobile application to support the indicated prevention and early intervention of childhood anxiety. Proc Conf Wireless Health 2015:1. [doi: 10.1145/2811780.2811954]
44. Perry Y, Werner-Seidler A, Calear A, Mackinnon A, King C, Scott J, et al. Preventing depression in final year secondary students: school-based randomized controlled trial. J Med Internet Res 2017 Nov 02;19(11):e369 [FREE Full text] [doi: 10.2196/jmir.8241] [Medline: 29097357]
45. Pine R, Sutcliffe K, McCallum S, Fleming T. Young adolescents' interest in a mental health casual video game. Digit Health 2020 Aug 30;6:2055207620949391 [FREE Full text] [doi: 10.1177/2055207620949391] [Medline: 32944270]

XSL•FO
RenderX

46. Pramana G, Parmanto B, Lomas J, Lindhiem O, Kendall PC, Silk J. Using mobile health gamification to facilitate cognitive behavioral therapy skills practice in child anxiety treatment: open clinical trial. JMIR Serious Games 2018 May 10;6(2):e9 [FREE Full text] [doi: 10.2196/games.8902] [Medline: 29748165]
47. Rodriguez A, Rey B, Vara MD, Wrzesien M, Alcaniz M, Banos RM, et al. A VR-based serious game for studying emotional regulation in adolescents. IEEE Comput Grap Appl 2015 Jan;35(1):65-73. [doi: 10.1109/mcg.2015.8]
48. Scholten H, Malmberg M, Lobel A, Engels RCME, Granic I. A randomized controlled trial to test the effectiveness of an immersive 3D video game for anxiety prevention among adolescents. PLoS One 2016 Jan 27;11(1):e0147763 [FREE Full text] [doi: 10.1371/journal.pone.0147763] [Medline: 26816292]
49. Schoneveld EA, Wols A, Lichtwarck-Aschoff A, Otten R, Granic I. Mental health outcomes of an applied game for children with elevated anxiety symptoms: a randomized controlled non-inferiority trial. J Child Fam Stud 2020 Jun 24;29(8):2169-2185. [doi: 10.1007/s10826-020-01728-y]
50. Schuurmans AAT, Nijhof KS, Engels RCME, Granic I. Using a videogame intervention to reduce anxiety and externalizing problems among youths in residential care: an initial randomized controlled trial. J Psychopathol Behav Assess 2018 Nov 29;40(2):344-354 [FREE Full text] [doi: 10.1007/s10862-017-9638-2] [Medline: 29937625]
51. Schuurmans AAT, Nijhof KS, Scholte R, Popma A, Otten R. A novel approach to improve stress regulation among traumatized youth in residential care: feasibility study testing three game-based meditation interventions. Early Interv Psychiatry 2020 Aug 10;14(4):476-485. [doi: 10.1111/eip.12874] [Medline: 31502420]
52. Schuurmans AAT, Nijhof KS, Scholte R, Popma A, Otten R. Game-based meditation therapy to improve posttraumatic stress and neurobiological stress systems in traumatized adolescents: protocol for a randomized controlled trial. JMIR Res Protoc 2020 Sep 23;9(9):e19881 [FREE Full text] [doi: 10.2196/19881] [Medline: 32965226]
53. Spengler B, Hofer J, Busch H. A video game-based investigation of power stress moderators in children. Motiv Emot 2019 Aug 08;44(3):345-355. [doi: 10.1007/s11031-019-09790-w]
54. Tuijnman A, Kleinjan M, Hoogendoorn E, Granic I, Engels RC. A game-based school program for mental health literacy and stigma regarding depression (Moving Stories): protocol for a randomized controlled trial. JMIR Res Protoc 2019 Mar 14;8(3):e11255 [FREE Full text] [doi: 10.2196/11255] [Medline: 30869652]
55. Tunney C, Cooney P, Coyle D, O'Reilly G. Comparing young people's experience of technology-delivered v. face-to-face mindfulness and relaxation: two-armed qualitative focus group study. Br J Psychiatry 2017 Apr 02;210(4):284-289. [doi: 10.1192/bjp.bp.115.172783] [Medline: 27765773]
56. van der Meulen H, McCashin D, O'Reilly G, Coyle D. Using computer games to support mental health interventions: naturalistic deployment study. JMIR Ment Health 2019 May 09;6(5):e12430 [FREE Full text] [doi: 10.2196/12430] [Medline: 31094346]
57. Cejudo J, López-Delgado ML, Losada L. Effectiveness of the videogame “Spock” for the improvement of the emotional intelligence on psychosocial adjustment in adolescents. Computers in Human Behavior 2019 Dec;101:380-386. [doi: 10.1016/j.chb.2018.09.028]
58. DeRosier ME, Thomas JM. Establishing the criterion validity of Zoo U's game-based social emotional skills assessment for school-based outcomes. Journal of Applied Developmental Psychology 2018 Mar;55:52-61. [doi: 10.1016/j.appdev.2017.03.001]
59. Tuijnman A, Granic !, Whitkin J, Engels RCME. Developing and testing ScrollQuest: A video game targeting rejection sensitivity in adolescents. Ext Abstr Publ Ann Symp Comput Human Interact Play 2017:213-221 [FREE Full text] [doi: 10.1145/3130859.3131302]
60. Kranzler A, Young JF, Hankin BL, Abela JRZ, Elias MJ, Selby EA. Emotional awareness: a transdiagnostic predictor of depression and anxiety for children and adolescents. J Clin Child Adolesc Psychol 2016 Feb 06;45(3):262-269 [FREE Full text] [doi: 10.1080/15374416.2014.987379] [Medline: 25658297]
61. Garber J, Brunwasser SM, Zerr AA, Schwartz KT, Sova K, Weersing VR. Treatment and prevention of depression and anxiety in youth: test of cross-over effects. Depress Anxiety 2016 Oct 03;33(10):939-959 [FREE Full text] [doi: 10.1002/da.22519] [Medline: 27699941]
62. Nyquist AC, Luebbe AM. An emotion recognition-awareness vulnerability hypothesis for depression in adolescence: a systematic review. Clin Child Fam Psychol Rev 2019 Sep 10;23(1):27-53. [doi: 10.1007/s10567-019-00302-3]
63. Ghandour RM, Sherman LJ, Vladutiu CJ, Ali MM, Lynch SE, Bitsko RH, et al. Prevalence and treatment of depression, anxiety, and conduct problems in US children. J Pediatr 2019 Mar;206:256-267 [FREE Full text] [doi: 10.1016/j.jpeds.2018.09.021] [Medline: 30322701]
64. Sendzik L, Ö Schäfer J, C Samson A, Naumann E, Tuschen-Caffier B. Emotional awareness in depressive and anxiety symptoms in youth: a meta-analytic review. J Youth Adolesc 2017 Apr 18;46(4):687-700. [doi: 10.1007/s10964-017-0629-0] [Medline: 28101745]
65. Checa D, Bustillo A. A review of immersive virtual reality serious games to enhance learning and training. Multimed Tools Appl 2019 Dec 05;79(9-10):5501-5527. [doi: 10.1007/s11042-019-08348-9]
66. Ravyse WS, Seugnet Blignaut A, Leendertz V, Woolner A. Success factors for serious games to enhance learning: a systematic review. Virtual Reality 2016 Sep 20;21(1):31-58. [doi: 10.1007/s10055-016-0298-4]

XSL•FO
RenderX

Abbreviations

CBT: cognitive behavioral therapy
VR: virtual reality

Edited by N Zary; submitted 17.05.21; peer-reviewed by LT De Paolis, J Lebiedz', J Piqueras; comments to author 28.06.21; revised version received 28.07.21; accepted 10.09.21; published 16.12.21.

Please cite as:
Martinez K, Menéndez-Menéndez MI, Bustillo A
Awareness, Prevention, Detection, and Therapy Applications for Depression and Anxiety in Serious Games for Children and Adolescents: Systematic Review
JMIR Serious Games 2021;9(4):e30482
URL: https://games.jmir.org/2021/4/e30482
doi:10.2196/30482
PMID:34927589

©Kim Martinez, Maria Isabel Menéndez-Menéndez, Andres Bustillo. Originally published in JMIR Serious Games (https://games.jmir.org), 16.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Investigating the Use of a Serious Game to Improve Opioid Safety Awareness Among Adolescents: Quantitative Study

Olufunmilola Abraham[1], BPharm, MSc, PhD; Claire Rosenberger[1], BA; Kathleen Tierney[1], BSc; Jen Birstler[2], MSc

[1]Social and Administrative Sciences Division, School of Pharmacy, University of Wisconsin–Madison, Madison, WI, United States

[2]Biostatistics and Medical Informatics, School of Medicine and Public Health, University of Wisconsin–Madison, Madison, WI, United States

Corresponding Author:
Olufunmilola Abraham, BPharm, MSc, PhD
Social and Administrative Sciences Division
School of Pharmacy
University of Wisconsin–Madison
777 Highland Ave
Madison, WI, 53705
United States
Phone: 1 6082634498
Email: olufunmilola.abraham@wisc.edu

Abstract

Background: The misuse of opioid medications among adolescents is a serious problem in the United States. Serious games (SGs) are a novel way to promote the safe and responsible management of opioid medications among adolescents, thereby reducing the number of adolescent opioid misuse cases reported annually.

Objective: This study aimed to examine the effect of the SG MedSMART: Adventures in PharmaCity on adolescents' opioid safety knowledge, awareness, behavioral intent, and self-efficacy.

Methods: A nationally representative sample of adolescents aged 12 to 18 years were recruited online through Qualtrics panels from October to November 2020. Data collection consisted of a pregame survey, 30 minutes of gameplay, and a postgame survey. The pregame and postgame surveys included 66 items examining the participants' baseline opioid knowledge, safety, use, and demographic information. The postgame survey had 25 additional questions regarding the MedSMART game. Gameplay scenarios included 5 levels intended to mimic adolescents' daily life while educating the players about appropriate opioid storage and disposable practices, negative consequences of sharing opioid medications, and the use of Narcan. Survey questions were divided into 10 categories to represent key concepts and were summarized into concept scores. Differences in concept scores were described by overall mean (SD) when stratified by gender, race, school, grade, and age. Differences of change in concept score were analyzed using the Kruskal-Wallis and correlation tests.

Results: A total of 117 responses were analyzed. The results showed significant improvement on all concept scores except for Narcan knowledge (mean change -0.04, SD 0.29; P=.60) and safe storage (mean change 0.03, SD 0.20; P=.09). Female participants had greater improvement than males for safe disposal (female mean 0.12, SD 0.25 vs male mean 0.04, SD 0.17; P=.05). Male participants had higher improvement than female participants for misuse behavior (female mean 0.05, SD 0.28 vs male mean 0.14, SD 0.27; P=.04). Perceived knowledge for participants who had non-White or Hispanic racial backgrounds had higher improvement than for non-Hispanic White participants (non-White mean 1.10, SD 1.06 vs White mean 0.75, SD 0.91; P=.03). Older grades were associated with greater improvement in opioid knowledge (correlation coefficient -0.23, 95% CI -0.40 to -0.05; P=.01). There were 28 (23.9%) participants who played all 5 levels of the game and had better improvement in opioid use self-efficacy.

Conclusions: Our findings suggest MedSMART: Adventures in PharmaCity can be used as an effective tool to educate adolescents on the safe and responsible use of prescribed opioid medications. Future testing of the effectiveness of this SG should involve a randomized controlled trial. Additionally, the feasibility of implementing and disseminating MedSMART: Adventures in PharmaCity in schools and health care settings such as adolescent health or primary care clinics, emergency departments, and pharmacies should be investigated.

(JMIR Serious Games 2021;9(4):e33975) doi:10.2196/33975

XSL•FO
RenderX

KEYWORDS

opioids; adolescents; opioid safety; medication safety; opioid knowledge; serious games; naloxone; Narcan; youth; technology; safety; gaming; addiction; drug; young adult; teenager; survey; awareness

Introduction

Opioid overdose mortality has increased dramatically in the United States. Over the past 2 decades (1999 to 2019), nearly 500,000 deaths from opioid-related overdoses have been reported [1]. Moreover, the growth trend for opioid-related deaths continues to increase over time. Opioid-involved overdose deaths increased by approximately 136%, from 21,088 deaths in 2010 to 49,860 in 2019—over 6 times the number of opioid-related overdose deaths in 1999 [2]. Overdose fatalities are largely driven by heroin use, misuse of synthetic opioids such as fentanyl, and pain relievers available legally by prescription, including oxycodone (OxyContin, Percocet, and Roxicodone), hydrocodone (Vicodin), codeine, morphine, and tramadol (Ultram) [3]. Due to the considerable increase in opioid overdoses, the opioid epidemic has been declared a national public health crisis in the United States [4].

The opioid epidemic affects persons of all ages, genders, and racial and ethnic groups, with 7.6% of adolescents reporting opioid misuse in 2019 [4-6]. Although the reasons for adolescent opioid misuse are complex and vary from person to person, the existing pain management prescribing practices and the lack of guidelines for pain management in children are contributing factors for opioid misuse [7-9]. Currently, the Centers for Disease Control and Prevention does not have guidelines or recommendations for pain management in children, nor does it provide a clear definition for opioid misuse and safety [7,10]. Studies have found 64% of clinicians do not have a standardized protocol for prescribing pain management medications to adolescents, less than half of pediatric providers screen their adolescent patients for substance use, and only 30% offer an intervention, which is often short-lived [8,9]. Furthermore, the most recent data on the US pediatric opioid prescribing practices show almost half of pediatric opioid prescriptions were categorized as high-risk [11].

Several studies have demonstrated that adolescents have inadequate knowledge and understanding about opioid use and safety [12-14]. These studies suggest that adolescents are well-informed on how to use prescription opioids; however, they are underinformed of the addictive potential of opioids, the risks of overdose, and the availability of naloxone (Narcan, Emergent Operations Ireland) to reverse an opioid overdose [15,16]. Additionally, many adolescents are not able to correctly identify which medications are opioids and which are not [13]. It is important to educate adolescents on the safe, appropriate, and responsible management of opioid medications because they are in a developmental stage where they can learn about and implement healthy lifestyle behaviors. Research shows adolescents can be taught to avoid harmful health-related behaviors, especially if they are provided with evidence that correlates dangerous behaviors to potentially dangerous outcomes [17]. Therefore, beginning opioid safety communication and education at an early age is critical for curbing the opioid epidemic, decreasing opioid misuse among young people, and preventing future opioid-related deaths.

The use of serious games (SGs) is a promising approach for promoting proper opioid use and safety among adolescents [14,18-21]. SGs are video games designed not only for entertainment, but to educate persons on a specific topic or topics, change an attitude or behavior, or create awareness of a certain issue [22]. The global SG market was valued at US $2731 million in 2016 and is expected to reach US $9167 million by 2023 based on the games' desirability and ability to improve learning outcomes [23]. Moreover, 70% of people under the age of 18 regularly play video games in the United States [23]. Due to the significant number of adolescents who play video games, and the international acceptance of these games, utilizing SGs to educate adolescents on opioid use and safety is highly feasible.

Many SGs have had success in achieving their health-related educational purposes. The SG Aislados teaches adolescents skills to prevent drug addiction, sexist behavior, and other risk-related behaviors [22]. The game has been shown to improve adolescents' attitudes and change their behaviors [22]. Additionally, Recovery Warrior 2.0, a motion-activated video game prototype, which targets relapse prevention for adolescents, has preliminary data indicating that an SG for addiction recovery appears to be possible and appropriate for adolescents [24]. Although several interventions have shown promise in improving adolescents' health-related behaviors, there are limited studies that have examined the effect of an SG on adolescent's opioid-safety awareness [20]. This study aimed to investigate the use of the SG, MedSMART: Adventures in PharmaCity in improving adolescents' opioid safety knowledge, awareness, behavioral intent, and self-efficacy.

Methods

Survey Design

Data were collected through a pregame survey, 30 minutes of gameplay, and a postgame survey; a common method to evaluate SGs [25]. Pregame and postgame opioid-related survey questions were adapted from various validated surveys and scales or created by the investigators [26]. Questions from a Wisconsin statewide survey collecting perceptions, awareness, and use of prescription medications in Wisconsin residents were adapted for adolescent use and assessed the participants' knowledge of safe prescription opioid storage and disposal [14,27]. Questions assessing self-efficacy of prescription opioid use and other learning objectives from the SG were adapted from the MUSE (Medication Understanding and Use Self-Efficacy) scale and a survey assessing workplace safety and health knowledge in adolescents [28,29]. Questions measuring participants' knowledge about naloxone (Narcan) were adapted from the Maryland Opioid Survey Summary Report [30]. Attention check questions were included in the pregame and postgame surveys to ensure the participants were

XSL•FO
RenderX

thoroughly reading survey questions and to prevent straight lining. All surveys were hosted online through Qualtrics (Qualtrics LLC), and the data collected were securely stored through Qualtrics. Surveys were reviewed by Qualtrics staff and the study team to ensure functionality prior to distribution. This study was approved by the Institutional Review Board of University of Wisconsin-Madison.

Pregame Survey

The pregame survey consisted of 66 nonrandomized items, and 59 of the items were used to examine the participants' baseline opioid knowledge, knowledge of safety, disposal and misuse, and the perceived effect an SG may have on opioid safety awareness. These 59 items consisted of questions with "Yes," "No," and "Don't know" as answers or a 5-point Likert scale. Two items were also used as attention check questions to prevent straight lining by asking the participants to select a specific option within the 5-point Likert scale. Five demographics questions were asked to assess the participants' characteristics, including race and ethnicity, grade in school, age, gender, and the number of children under the age of 18 living in their household. The online survey consisted of 12 pages with up to 10 items on each page. The pregame survey questions are included in Multimedia Appendix 1.

Postgame Survey

The postgame survey consisted of the same 59-item baseline opioid-related questions in the pregame survey and 2 attention check questions, as well as 25 additional questions asking about the participants' perspectives on the MedSMART game. The postgame survey included a question asking the participants if they had ever been prescribed opioids by a doctor, and questions asking them to list the opioids they had been prescribed and any prescribed medications they perceived might be opioids. Additional personal health information was not collected in the survey per Qualtrics guidelines for surveys recruiting participants through their research panels. Response options included 5-point Likert scale questions, multiple choice questions, free response questions, and "yes," "no," and "unsure" answers. The online survey consisted of 17 pages with up to 10 items on each page, and the items were not randomized. The postgame survey is included in Multimedia Appendix 2.

Intervention

The SG MedSMART: Adventures in PharmaCity was designed to educate adolescents about safe opioid use and management as well as enhancing the players' ability to make informed decisions about proper opioid use in real life. MedSMART: Adventures in PharmaCity was developed over the course of several months by a team of researchers, health care personnel, and game design experts [18]. The key elements of the game were defined, including goals, levels, mechanics, game flow, story, and characters. The game was only playable on a computer or tablet with keyboard access and had simple commands (all player actions were made using keys "A," "D," "W," and spacebar). The game engine was built using Unity (Unity Technologies). The situations presented in the game were designed to be similar to scenarios adolescents may encounter in real life. The game provided immediate feedback to the player if they made a wrong decision, and the player could only progress in the game if they made the right decision and followed the correct sequence in the story line. After the prototype was built, the game was piloted with adolescents and pharmacy students to examine whether they liked the game, understood how to use it, and could navigate throughout the game [18,21]. Their feedback was incorporated into the final game design used in this study.

MedSMART: Adventures in PharmaCity is a task-guided game divided into 5 levels. Participants play as "Shan," the anthropomorphized sheep, who has been prescribed opioids after breaking their arm. The player's objective is to guide Shan to make the right choices regarding the proper use and care of opioids. The player navigates through the levels, each covering different opioid safety topics. Level 1, A Quiet Sunday Afternoon, is focused on teaching the player about the safe storage of opioids; the player is taught to lock up medication in a cabinet, so their friends do not take them (Figure 1). The second level, Monday Morning Bus Ride, enhances the game story line by adding a real-life scene about being in pain and forgetting there was an important assignment due that day (Figure 2). Level 3, A Persuasive Speech at School, focuses on teaching the player not to take others pain medication and the negative consequences of taking someone else's medications; even if 2 people are prescribed a pain killer, that does not mean it is the same kind of medication or dose (Figure 3). Level 4, Bus Ride Home, focuses on teaching the player not to share their medication with others, the negative consequences of sharing medications with others, and what Narcan is used for (Figure 4). Level 5, Last Minute Chore, shows the player the correct way to dispose of an opioid medication and why they should not throw medications away or flush them down the toilet (Figure 5). The scenarios in each level were intended to be realistic and relevant to the adolescents' daily life.

XSL•FO
RenderX

Figure 1. Screenshot from level 1, A Quiet Sunday Afternoon.

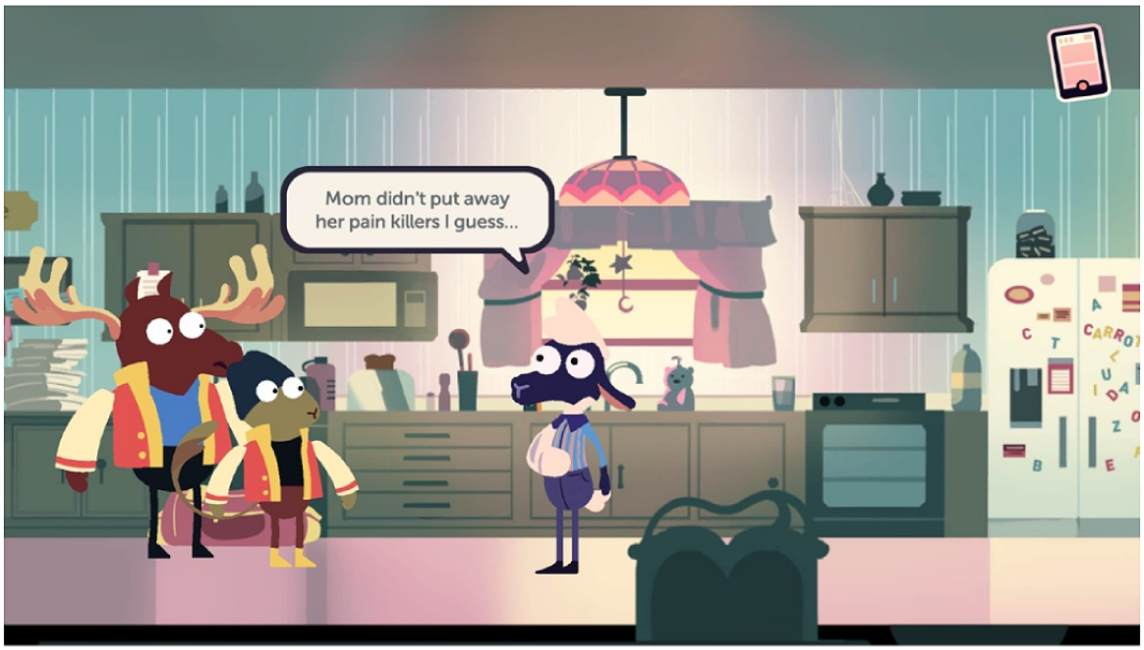
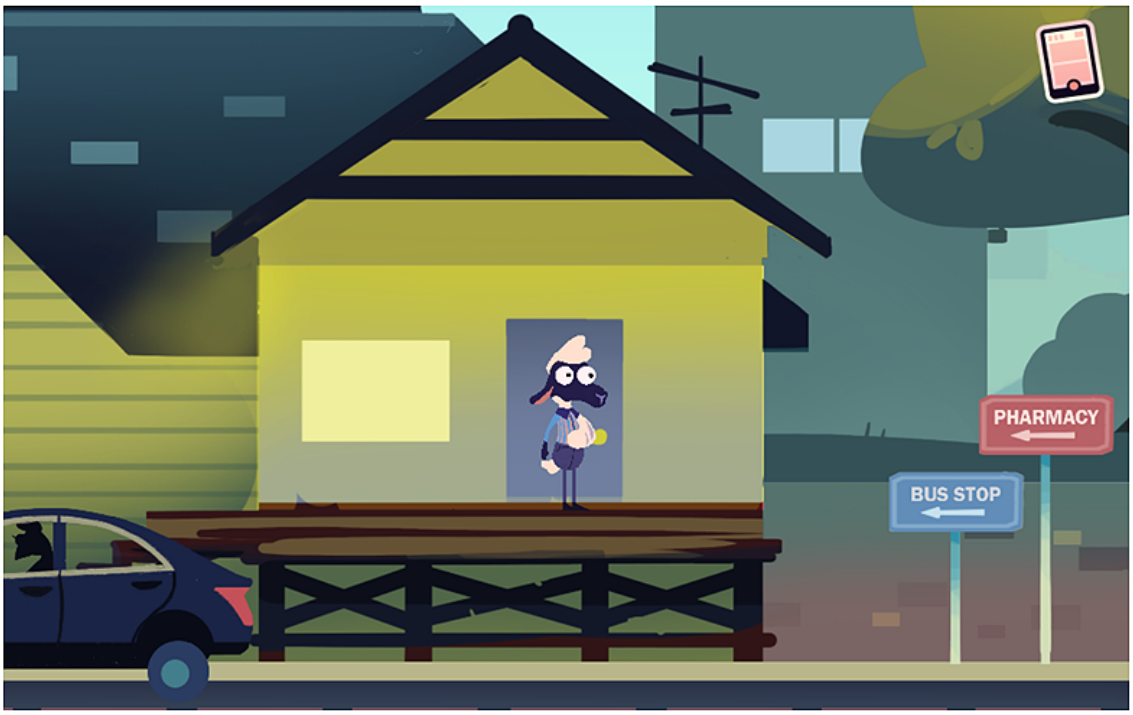


Figure 2. Screenshot from level 2, Monday Morning Bus Ride.

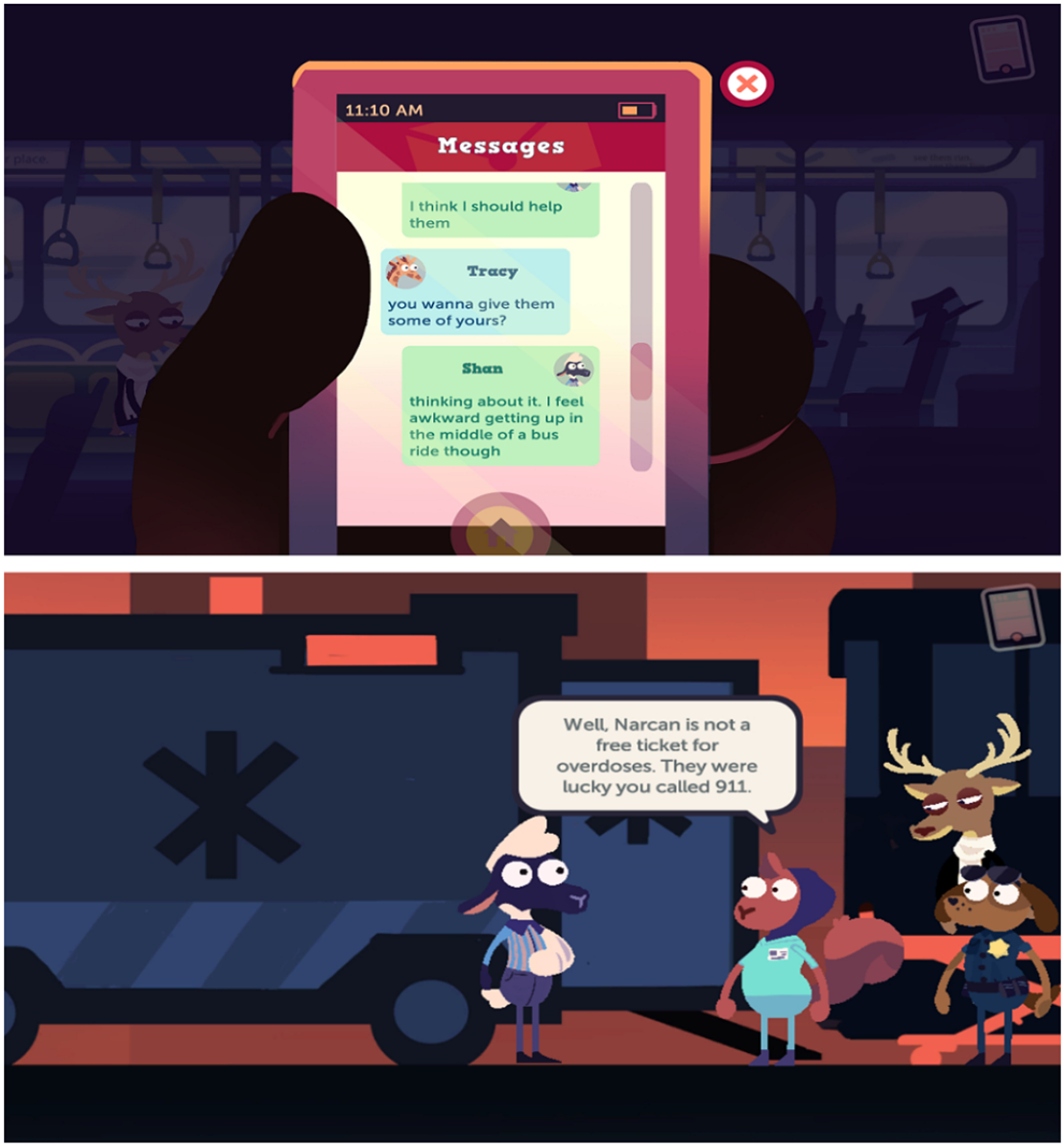


XSL•FO
RenderX

Figure 3. Screenshots from level 3, A Persuasive Speech at School.

XSL•FO
RenderX

Figure 4. Screenshots from level 4, Bus Ride Home.



XSL•FO
RenderX

Figure 5. Screenshot from level 5, Last Minute Chore.

Procedure

From October to November 2020, a nationally representative (quotas were set for race or ethnicity and gender based on the US census data) sample of adolescents was recruited via Qualtrics research panels. Parents of eligible adolescents were targeted. Eligible participants were adolescents aged 12 to 18 years who lived in the United States and could speak and understand English. After screening for eligibility, the parents were provided with a link to the online consent document and asked to provide consent for their child to participate in the pregame survey. The child was also provided with an online assent document and asked to provide assent. The parents were then directed to the pregame survey for their child to complete independently. Upon completion of the pregame survey, the parents of participants received compensation based on their predetermined agreement with Qualtrics. The parents were then provided with a link to the online consent document and asked to consent to their child's participation in gameplay and a postgame survey. The children were also provided with a link to an online assent document and asked to provide assent for participation. Once consent and assent were confirmed, the parents received a link with instructions for their child to play the SG for 30 minutes and complete a postgame survey. Upon verification of gameplay and completion of the postgame survey, a US $10 amazon e-gift card was emailed to the parent of the adolescent for participation. All consent and assent documents detailed the study activities, time commitment, purpose of study, principal investigator's information, and confidentiality and data security measures.

Data Analysis

Survey questions were divided into 10 categories to represent key concepts and summarized into concept scores. Pregame and postgame surveys had Likert scores on a 1 through 5 scale and knowledge scores on a 0 to 100% scale; changes in concept scores, from presurvey to postsurvey, were calculated for each participant and were the primary efficacy outcomes. Concept score changes were described by overall mean (SD) and were stratified by gender, race, school, grade, and age. Differences in concept score changes were analyzed using the Kruskal-Wallis tests for categorical characteristics and correlation tests for continuous characteristics. Primary analysis included participants who accessed the game based on IP (located within the United States) and satisfied attention checks for both pregame and postgame surveys. Secondary analyses aimed to (1) assess the differences between participants included in primary analysis vs those who did not meet full attention criteria; and (2) assess the associations between the levels played and the length of play with concept scores. The participants who did not meet full attention criteria either did not access the game or did not self-report playing the game but met other attention criteria (ie, select "A great deal" when asked). No *P* value adjustments were made to account for inflated type 1 error rate. Significance was assessed at the α=.05 level. All statistical analyses were performed using R statistical software, version 4.0.5 (The R Foundation).

Results

There were 117 participants who met full attention criteria on pregame and postgame surveys and accessed the game. Of these participants, 55.4% (n=59) identified as male, 48.72% (n=57) identified as White, and the mean age was 14.62 (SD 1.62). The characteristics are shown in Table 1.

XSL•FO
RenderX

Table 1. Participant demographic characteristics.

Characteristics	Values
Grade, n (%)	
7	20 (17)
8	19 (16)
9	23 (20)
10	21 (18)
11	22 (19)
12	12 (10)
Age (years), mean (SD)	14.62 (1.62)
Gender,[a] n (%)	
Female	57 (48.7)
Male	59 (50.4)
Nonbinary	1 (0.9)
Race or ethnicity,[b] n (%)	
White or Caucasian	57 (48.72)
Black or African American	23 (19.66)
Hispanic or Latinx	6 (5.13)
Asian	6 (5.13)
American Indian or Alaskan Native	1 (0.85)
Native Hawaiian or other Pacific Islander	0 (0)
More than one selected	24 (20.51)
Other, please specify	0 (0)

[a]Three options were presented to the participants to select for their gender: “Male,” “Female,” and “Other.”

[b]If participants only selected one category, that was their defined race; all other combinations of selections were defined as “Other.”

These participants significantly improved on all concept scores except for Narcan knowledge (mean change -0.04, SD 0.29; *P*=.60) and safe storage (mean change 0.03, SD 0.20; *P*=.09) (Table 2). Female participants had greater improvement than males for safe disposal (female mean 0.12, SD 0.25 vs male mean 0.04, SD 0.17; *P*=.05) but not for misuse behavior (female mean 0.05, SD 0.28 vs male mean 0.14, SD 0.27; *P*=.04). Perceived knowledge for participants who did not identify as White or were Hispanic had greater improvements than White participants (non-White mean 1.10, SD 1.06 vs White mean 0.75, SD 0.91; *P*=.03). Older grades were associated with greater improvements in opioid knowledge (correlation coefficient -0.23, 95% CI -0.40 to -0.05; *P*=.01). No other associations were significant.

XSL•FO
RenderX

Table 2. Concept scores among participants.

Characteristics	Self-efficacy: MUSE[a]	Self-efficacy: opioid safety	Perceived knowledge	Misuse harm	Behavioral intent	Safe storage	Safe disposal	Opioid knowledge	Narcan knowledge	Misuse behavior
Overall										
N	117	117	117	117	117	117	117	117	45	117
Mean (SD)	0.28 (0.60)	0.39 (0.71)	0.93 (1.00)	0.22 (0.94)	0.25 (0.62)	0.03 (0.20)	0.08 (0.22)	0.06 (0.16)	-0.04 (0.29)	0.09 (0.28)
Kruskal-Wallis *P* value	<.001	<.001	<.001	.01	<.001	.09	<.001	<.001	.6	<.001
Gender										
Female, mean (SD)	0.35 (0.67)	0.37 (0.70)	1.14 (0.97)	0.15 (0.72)	0.33 (0.59)	0.02 (0.22)	0.12 (0.25)	0.06 (0.16)	0.00 (0.35)	0.05 (0.28)
Male, mean (SD)	0.21 (0.51)	0.39 (0.74)	0.74 (1.00)	0.27 (1.11)	0.19 (0.65)	0.04 (0.17)	0.04 (0.17)	0.05 (0.16)	-0.06 (0.25)	0.14 (0.27)
Kruskal-Wallis *P* value	.42	.99	.07	.37	.33	.21	.05	.88	.32	.04
Race (grouping 1)										
White or Caucasian, mean (SD)	0.23 (0.41)	0.30 (0.65)	0.75 (0.91)	0.20 (0.84)	0.23 (0.66)	0.02 (0.14)	0.05 (0.18)	0.03 (0.13)	-0.10 (0.30)	0.10 (0.27)
Black or African American, mean (SD)	0.27 (0.75)	0.20 (0.80)	0.76 (0.90)	0.15 (1.02)	0.12 (0.65)	0.02 (0.24)	0.12 (0.26)	0.07 (0.22)	0.15 (0.24)	0.01 (0.18)
Hispanic or Latinx, mean (SD)	0.25 (0.61)	0.58 (0.82)	1.27 (1.22)	0.29 (1.08)	0.40 (0.59)	0.11 (0.29)	0.10 (0.19)	0.09 (0.15)	-0.11 (0.17)	0.18 (0.34)
Other or missing, mean (SD)	0.54 (0.89)	0.71 (0.45)	1.37 (0.95)	0.30 (1.00)	0.30 (0.47)	-0.02 (0.07)	0.10 (0.32)	0.12 (0.18)	0.17 (0.24)	0.05 (0.33)
Kruskal-Wallis *P* value	.76	.04	.02	.78	.72	.36	.58	.23	.04	.38
Race (grouping 2)[b]										
White or Caucasian, mean (SD)	0.23 (0.41)	0.30 (0.65)	0.75 (0.91)	0.20 (0.84)	0.23 (0.66)	0.02 (0.14)	0.05 (0.18)	0.03 (0.13)	-0.10 (0.30)	0.10 (0.27)
Non-White, mean (SD)	0.33 (0.73)	0.47 (0.76)	1.10 (1.06)	0.24 (1.03)	0.27 (0.59)	0.05 (0.24)	0.11 (0.25)	0.09 (0.18)	0.06 (0.24)	0.09 (0.29)
Kruskal-Wallis *P* value	.73	.21	.03	.42	.63	.73	.23	.26	.12	.80
School										
High school, mean (SD)	0.26 (0.56)	0.34 (0.68)	0.86 (0.94)	0.19 (0.80)	0.23 (0.68)	0.03 (0.21)	0.06 (0.19)	0.04 (0.14)	-0.06 (0.30)	0.12 (0.27)
Middle school, mean (SD)	0.31 (0.67)	0.48 (0.78)	1.07 (1.11)	0.27 (1.17)	0.30 (0.49)	0.03 (0.17)	0.13 (0.26)	0.10 (0.19)	0.03 (0.25)	0.03 (0.28)
Kruskal-Wallis *P* value	.42	.21	.23	.51	.45	.77	.06	.34	.39	.15
Grade										
Correlation coefficient	-0.17	-0.14	-0.17	-0.08	-0.09	0	-0.18	-0.23	-0.1	0.12
95% CI	-0.34 to 0.02	-0.31 to 0.05	-0.34 to 0.02	-0.25 to 0.11	-0.26 to 0.10	-0.18 to 0.18	-0.35 to 0.00	-0.40 to -0.05	-0.38 to 0.20	-0.06 to 0.30
Pearson's correlation *P* value	.07	.15	.07	.41	.35	.98	.06	.01	.52	.18

XSL•FO
RenderX

Characteristics	Self-efficacy: MUSE[a]	Self-efficacy: opioid safety	Perceived knowledge	Misuse harm	Behavioral intent	Safe storage	Safe disposal	Opioid knowledge	Narcan knowledge	Misuse behavior
Age										
Correlation coefficient	-0.15	-0.12	-0.16	-0.07	-0.06	0	-0.15	-0.18	-0.15	0.1
95% CI	-0.32 to 0.03	-0.29 to 0.07	-0.33 to 0.03	-0.25 to 0.11	-0.24 to 0.12	-0.18 to 0.19	-0.32 to 0.03	-0.35 to 0.00	-0.42 to 0.15	-0.08 to 0.28
Pearson's correlation *P* value	.11	.22	.09	.44	.52	.96	.11	.05	.34	.27

[a]MUSE: Medication Understanding and Use Self-Efficacy.

[b]Race grouping 2 examines participants who selected “White” as one category and every other option as “non-White.”

Of the 117 participants who played the game, 28 (23.9%) played all 5 levels. The participants who partially completed the game played an average of 2.42 (SD 1.18) levels (Table 3). Compared with the participants who partially completed the game, completers had a worse change in Narcan knowledge (mean -0.10, SD 0.16) than noncompleters (mean 0.05, SD 0.21; *P*=.04) but better improvement in self-efficacy: opioid safety (completers mean 0.74, SD 0.73 vs noncompleters mean 0.30, SD 0.73; *P*=.01). Completers also had a longer length of play (completers mean 26.94, SD 8.16 minutes vs noncompleters mean 17.41, SD 13.16 minutes; *P*<.001).

XSL•FO
RenderX

Table 3. Examining the differences between the participants who played all 5 levels of the game and those who did not.

Variables	Noncompletes	Played all 5 levels	*P* value
n	67	28	__[a]
Behavioral intent, mean (SD)	0.24 (0.61)	0.26 (0.54)	.88
Misuse behavior, mean (SD)	0.10 (0.28)	0.11 (0.28)	.39
Misuse behavior 2, mean (SD)[b]	0.10 (0.35)	0.04 (0.19)	.32
Misuse harm, mean (SD)	0.22 (1.04)	0.21 (0.78)	.90
Narcan knowledge, mean (SD)	0.05 (0.21)	-0.10 (0.16)	.04
Opioid knowledge: harming teens, mean (SD)[c]	0.09 (0.42)	0.21 (0.42)	.20
Opioid knowledge, mean (SD)	0.07 (0.15)	0.05 (0.16)	.68
Perceived knowledge, mean (SD)	0.84 (1.00)	1.04 (0.98)	.43
Safe disposal, mean (SD)	0.10 (0.22)	0.06 (0.15)	.93
Safe storage, mean (SD)	0.02 (0.13)	0.03 (0.20)	.66
Self-efficacy: opioid safety, mean (SD)	0.30 (0.73)	0.74 (0.73)	.01
Self-efficacy: MUSE,[d] mean (SD)	0.26 (0.55)	0.35 (0.66)	.58
Length of play, mean (SD)	17.41 (13.16)	26.94 (8.16)	<.001
Number of levels played, mean (SD)	2.42 (1.18)	—	—
Gender, n (%)			.32
Female	36 (53.7)	11 (39.3)	—
Male	30 (44.8)	17 (60.7)	—
Nonbinary	1 (1.5)	0 (0.0)	—
Age, mean (SD)	14.36 (1.60)	14.54 (1.62)	.68
Race (grouping 1), n (%)			.14
A: White or Caucasian	29 (43.3)	13 (46.4)	—
B: Black or African American	17 (25.4)	4 (14.3)	—
C: Hispanic or Latinx	15 (22.4)	4 (14.3)	—
D: Other or missing	6 (9.0)	7 (25.0)	—
Race (grouping 2) = B: non-White, n (%)[e]	38 (56.7)	15 (53.6)	.78

[a]Not applicable.

[b]Individual question asking, “Is it okay to take someone else's opioid medication if you have had the same prescription in the past?”

[c]Individual question asking, “Is the opioid crisis harming teenagers in the United States?”

[d]MUSE: Medication Understanding and Use Self-Efficacy.

[e]Race grouping 2 examines participants who selected “White” as one category and every other option as “non-White.”

There were 39 (33%) participants who met full attention criteria on the pregame and postgame surveys but either did not access the game or reported playing the wrong game in the postgame survey. Compared to those who played the game, the participants who did not meet the criteria had worse changes in perceived knowledge (*P*=.05); self-efficacy: opioid safety (*P*=.05); self-efficacy: MUSE (*P*=.01); and when asked the question, “Is it okay to take someone else's opioid medication if you have had the same prescription in the past?” (*P*=.01) (Table 4). Additionally, those who attempted the game but reported playing the wrong game had lower length of play (poor attention mean 7.02, SD 10.23 minutes vs full attention with game mean 20.22, SD 12.64; *P*=.04) and played fewer levels (poor attention mean 1, no SD vs full attention with game mean 3.18, SD 1.54; *P*=.01).

XSL•FO
RenderX

Table 4. Examining the differences among participants who accessed the game and had good attention vs those who did not access the game and had poor attention.

Variables	Poor attention, did not access game	Good attention, accessed game	*P* value
n	39	117	__[a]
Behavioral intent, mean (SD)	0.17 (0.75)	0.25 (0.62)	.25
Misuse behavior, mean (SD)	0.05 (0.31)	0.09 (0.28)	.39
Misuse behavior 2, mean (SD)[b]	-0.08 (0.35)	0.09 (0.32)	.01
Misuse harm, mean (SD)	-0.04 (0.82)	0.22 (0.94)	.10
Narcan knowledge, mean (SD)	0.05 (0.28)	-0.04 (0.29)	.44
Opioid knowledge: harming teens, mean (SD)[c]	0.13 (0.41)	0.13 (0.41)	.99
Opioid knowledge, mean (SD)	0.09 (0.21)	0.06 (0.16)	.87
Perceived knowledge, mean (SD)	0.52 (1.13)	0.93 (1.00)	.05
Safe disposal, mean (SD)	0.09 (0.22)	0.08 (0.22)	.95
Safe storage, mean (SD)	0.05 (0.15)	0.03 (0.20)	.16
Self-efficacy: opioid safety, mean (SD)	0.14 (0.71)	0.39 (0.71)	.05
Self-efficacy: MUSE,[d] mean (SD)	-0.03 (0.81)	0.28 (0.60)	.01
Length of play, mean (SD)	7.02 (10.23)	20.22 (12.64)	.04
Number of levels played, mean (SD)	1.00 (0.00)	3.18 (1.54)	.01
Gender, n (%)			.42
Female	15 (38.5)	57 (48.7)	—
Male	23 (59.0)	59 (50.4)	—
Nonbinary	1 (2.6)	1 (0.9)	—
Age, mean (SD)	14.90 (1.48)	14.62 (1.62)	.32
Race (grouping 1), n (%)			.07
A: White or Caucasian	26 (66.7)	57 (48.7)	—
B: Black or African American	8 (20.5)	23 (19.7)	—
C: Hispanic or Latinx	5 (12.8)	23 (19.7)	—
D: Other or missing	0 (0.0)	14 (12.0)	—
Race (grouping 2) = B: non-White, n (%)[e]	13 (33.3)	60 (51.3)	.05

[a]Not applicable.

[b]Individual questions asking, “Is it okay to take someone else's opioid medication if you have had the same prescription in the past?”

[c]Individual questions asking, “Is the opioid crisis harming teenagers in the United States?”

[d]MUSE: Medication Understanding and Use Self-Efficacy.

[e]Race grouping 2 examines participants who selected “White” as one category and every other option as “non-White.”

There were 13 (11%) participants who failed level 3 but succeeded in level 4, indicating they learned not to share medications with others. These players had longer length of play (mean 27.11, SD 9.56), played more levels (mean 4.38, SD 0.87), and displayed more opioid misuse behaviors during gameplay (mean 3.77, SD 1.88) than all other players (Table 5).

XSL•FO
RenderX

Table 5. Players who learned not to share meds vs players with all other patterns.

Variables	All other patterns	Failure in level 3, success in level 4	*P* value
n	104	13	__[a]
Behavioral intent, mean (SD)	0.27 (0.64)	0.13 (0.42)	.38
Misuse behavior, mean (SD)	0.10 (0.29)	0.03 (0.21)	.69
Misuse behavior 2, mean (SD)[b]	0.10 (0.33)	0.08 (0.28)	.83
Misuse harm, mean (SD)	0.18 (0.84)	0.55 (1.50)	.13
Narcan knowledge, mean (SD)	-0.03 (0.30)	-0.08 (0.17)	.47
Opioid knowledge: harming teens, mean (SD)[c]	0.12 (0.41)	0.15 (0.38)	.83
Opioid knowledge, mean (SD)	0.05 (0.14)	0.14 (0.25)	.25
Perceived knowledge, mean (SD)	0.91 (1.00)	1.08 (1.04)	.64
Safe disposal, mean (SD)	0.08 (0.21)	0.12 (0.26)	.35
Safe storage, mean (SD)	0.04 (0.20)	0.00 (0.10)	.72
Self-efficacy: opioid safety, mean (SD)	0.34 (0.70)	0.73 (0.74)	.05
Self-efficacy: MUSE,[d] mean (SD)	0.27 (0.56)	0.33 (0.84)	.89
Length of play, mean (SD)	19.13 (12.77)	27.11 (9.55)	.02
Number of levels played, mean (SD)	2.99 (1.54)	4.38 (0.87)	.002
Opioid failures, mean (SD)[e]	2.38 (2.36)	3.77 (1.88)	.03
Gender, n (%)			.88
Female	50 (48.1)	7 (53.8)	—
Male	53 (51.0)	6 (46.2)	—
Nonbinary	1 (1.0)	0 (0.0)	—
Age, mean (SD)	14.69 (1.62)	14.08 (1.55)	.19
Race (grouping 1), n (%)			.54
A: White or Caucasian	51 (49.0)	6 (46.2)	—
B: Black or African American	19 (18.3)	4 (30.8)	—
C: Hispanic or Latinx	22 (21.2)	1 (7.7)	—
D: Other or missing	12 (11.5)	2 (15.4)	—
Race (grouping 2) = B: non-White, n (%)[f]	53 (51.0)	7 (53.8)	.84

[a]Not applicable.

[b]Individual question asking, “Is it okay to take someone else's opioid medication if you have had the same prescription in the past?”

[c]Individual question asking, “Is the opioid crisis harming teenagers in the United States?”

[d]MUSE: Medication Understanding and Use Self-Efficacy.

[e]Players failed the level due to an opioid misuse behavior (failure to lock up opioids, sharing opioids, etc) vs a non–opioid-related failure (failure to get a hall pass, forgetting notes, etc).

[f]Race grouping 2 examines participants who selected “White” as one category and every other option as “non-White.”

Discussion

Principal Findings

This paper describes the use of an SG designed to educate adolescents on opioid safety. The game was evaluated in terms of impact on the players’ pregame and post survey results for self-efficacy, perceived knowledge, misuse harm and behavior, behavioral intent, safe storage and disposal, opioid knowledge, and Narcan knowledge. The results indicate significant improvement in all areas except for Narcan knowledge and safe storage. However, not all remaining improvements were still significant when compared to a self-selecting control group of participants who either did not play the game or reported playing the wrong game. Those who accessed the game had better improvement in misuse behavior, perceived knowledge, and self-efficacy. Thus, the findings from this study suggest the potential effectiveness of the SG, MedSMART: Adventures in PharmaCity on improving opioid medication safety awareness among adolescents.

XSL•FO
RenderX

The findings from this study revealed female participants had better improvement for safe disposal, but males had greater improvement for misuse behavior. The participants who were non-White or Hispanic had higher improvements in perceived knowledge than White participants, and older participants were associated with greater improvements in opioid knowledge. There are limited studies that examine why these differences exist amongst gender, race, and age [31,32]. Those that examine these differences typically only study adults and may not apply a pre-post survey study design [33]. Racial and ethnic disparities in pain management are well documented in literature and may contribute to these differences as well [34-38]. Studies show that not only are non-White patients less likely to receive opioid prescriptions for certain conditions, but also that disparities exist in access to specialized care, such as mental health, addiction, or pain specialists [34-38]. Data from this study and current literature suggest that research focused on these aspects is warranted in the adolescent population.

Implications

Studies have indicated that SGs possess the potential for increasing students' learning, motivation, and engagement as well as the ability to develop their minds and improve their learning efficiency [39-41]. Additionally, the Pediatric Pharmacy Advocacy Group recommends pharmacists educate adolescents about the proper administration, storage, and disposal of opioids [42]. Current literature suggests that reducing opioid prescribing by health care professionals, specifically dentists and surgeons, could substantially lower prescription opioid exposure in adolescents [11]. There is a need for initiatives that target high-volume prescribers and provide medication safety education for the patients.

An opportunity exists to implement MedSMART: Adventures in PharmaCity in multiple settings: (1) a school setting where students can play the game in a health course or a similar setting; (2) a doctor's office where the adolescent is being prescribed the opioid and could play the game in the waiting room; and (3) at the pharmacy where the adolescent is picking up their opioid prescription. The addition of an SG in any of these settings may help proactively curb opioid misuse in adolescents.

Limitations

This single-arm study did not have a randomized researcher-blinded control group. Although the participants who did not play the game or who reported playing the wrong game were compared to those who did play the game (Table 4), this was a self-selecting control group and not randomized. Efficacy outcomes relied on self-completion questionnaires, which may have been affected by participant dishonesty and reporting bias.

Additionally, although the sample demographics were similar to those of the US demographics, the majority of the participants were White. Thus, generalizability of study findings may be somewhat limited, as there are racial and ethnic differences in access to prescription opioids and pain management.

The participants could not ask for assistance navigating the game or clarification on survey questions because they were not monitored by research staff during data collection. Moreover, the participants' attentiveness during gameplay was not observed. The study design relied on trusting the parent to not influence their child's survey responses or gameplay. To ensure accuracy, future designs should confirm that adolescents are taking the surveys and playing the game and that parents are not affecting the responses. A focus group or interview could also be incorporated to a study's design to understand what players learned from the game or to identify potential game improvements.

Conclusion

Health messaging surrounding opioid safety requires novel and engaging strategies to be effective in increasing knowledge, changing behavior, and preventing prescription opioid misuse. The SG MedSMART: Adventures in PharmaCity was designed to engage adolescents in real-life scenarios and provide them with the information needed to correctly use opioids. The findings from this study show that the SG significantly improved all concept scores except for Narcan knowledge and safe storage. Future research should examine how the game can improve adolescents' understanding for safe opioid storage and the use of Narcan. MedSMART: Adventures in PharmaCity could be a useful and effective tool for preventive opioid misuse intervention programs and should replicate this study using a randomized controlled trial with a prepost survey study design. Further research is needed to assess potential benefits of dissemination and implementation of the game in health care or school settings.

Acknowledgments

The authors appreciate Dr Laura Stephenson from the University of Wisconsin–Madison School of Pharmacy for her assistance with creating and managing the Institutional Review Board protocol, survey development, and data collection. We would also like to thank Lisa Szela from the University of Wisconsin–Madison School of Pharmacy for editing the manuscript and providing feedback throughout the writing process.

This study was supported by KL2 grant KL2 TR002374-03 and grant UL1TR002373 to the University of Wisconsin–Madison Institute for Clinical and Translational Research by the Clinical and Translational Science Award program, through the National Institutes of Health National Center for Advancing Translational Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey questions, composite scores, and response types in the pregame survey.
[DOCX File , 22 KB - games_v9i4e33975_app1.docx]

Multimedia Appendix 2
Survey questions, composite scores, and response types in the postgame survey.
[DOCX File , 23 KB - games_v9i4e33975_app2.docx]

References

1. Wide-ranging online data for epidemiologic research (WONDER). Centers for Disease Control and Prevention. URL: http://wonder.cdc.gov [accessed 2021-07-15]
2. Overdose Death Rates. National Institute on Drug Abuse. URL: https://www.drugabuse.gov/drug-topics/trends-statistics/overdose-death-rates [accessed 2021-07-15]
3. Singh GK, Kim Jr IE, Girmay M, Perry C, Daus GP, Vedamuthu IP, et al. Opioid Epidemic in the United States: Empirical Trends, and A Literature Review of Social Determinants and Epidemiological, Pain Management, and Treatment Patterns. Int J MCH AIDS 2019 Sep 15;8(2):89-100. [doi: 10.21106/ijma.284]
4. Rudd RA, Aleshire N, Zibbell JE, Gladden RM. Increases in Drug and Opioid Overdose Deaths — United States, 2000–2014. MMWR Morb. Mortal. Wkly. Rep 2016 Jan 01;64(50-51):1378-1382. [doi: 10.15585/mmwr.mm6450a3]
5. Hedegaard H, Miniño A, Warner M. Drug Overdose Deaths in the United States, 1999-2017. NCHS Data Brief;(329) 2018:1-8. [doi: 10.15620/cdc:112340]
6. High School YRBS. Centers for Disease Control and Prevention. URL: https://nccd.cdc.gov/youthonline/App/Results.aspx? [accessed 2021-07-14]
7. Abraham O, Thakur T, Brown R. Prescription opioid misuse and the need to promote medication safety among adolescents. Research in Social and Administrative Pharmacy 2019 Jul;15(7):841-844. [doi: 10.1016/j.sapharm.2019.01.003]
8. Huang JS, Kuelbs CL. Clinician opioid prescribing practices and patient utilization of prescribed opioids in pediatrics. J of Opioid Management 2018 Sep 01;14(5):309-316. [doi: 10.5055/jom.2018.0463]
9. Harris BR. Talking about screening, brief intervention, and referral to treatment for adolescents: An upstream intervention to address the heroin and prescription opioid epidemic. Preventive Medicine 2016 Oct;91:397-399. [doi: 10.1016/j.ypmed.2016.08.022]
10. Napier T, Persons A. Using modern neuroscience to inform opioid use and abuse liability in adolescents. Orthop Nurs 2019;38:166-171. [doi: 10.1097/nor.0000000000000527]
11. Chua K, Brummett CM, Conti RM, Bohnert AS. Opioid Prescribing to US Children and Young Adults in 2019. Pediatrics 2021 Aug 16;148(3):e2021051539. [doi: 10.1542/peds.2021-051539]
12. Smith LN, Caballero NS, Cintron C, Toklu HZ, Smillov M. Impact of an educational debate on the knowledge of college students on opioids, and factors affecting their perception about addiction. jrp 2019 Sep 03;23(5):792-796. [doi: 10.35333/jrp.2019.26]
13. Abraham O, Szela L, Thakur T, Brasel K, Brown R. Adolescents' Perspectives on Prescription Opioid Misuse and Medication Safety. J Pediatr Pharmacol Ther 2021 2021;26(2):133-143. [doi: 10.5863/1551-6776-26.2.133]
14. Abraham O, Szela L, Norton D, Stafford H, Hoernke M, Brown R. Adolescents' awareness about prescription opioid misuse and preferences for educational interventions. Journal of the American Pharmacists Association 2020 Nov;60(6):978-985. [doi: 10.1016/j.japh.2020.07.029]
15. Moore SK, Grabinski M, Bessen S, Borodovsky JT, Marsch LA. Web-Based Prescription Opioid Abuse Prevention for Adolescents: Program Development and Formative Evaluation. JMIR Form Res 2019 Jul 19;3(3):e12389. [doi: 10.2196/12389]
16. Unger JB, Molina GB, Baron MF. Opioid knowledge and perceptions among Hispanic/Latino residents in Los Angeles. Substance Abuse 2020 Aug 19:1-7. [doi: 10.1080/08897077.2020.1806185]
17. Bachman JG, Johnson LD, O'Malley PM. Explaining recent increases in students' marijuana use: impacts of perceived risks and disapproval, 1976 through 1996. Am J Public Health 1998 Jun;88(6):887-892. [doi: 10.2105/ajph.88.6.887]
18. Abraham O, Thakur T, Brown R. Developing a Theory-Driven Serious Game to Promote Prescription Opioid Safety Among Adolescents: Mixed Methods Study. JMIR Serious Games 2020 Jul 3;8(3):e18207. [doi: 10.2196/18207]
19. Abraham O, Rosenberger CA, Birstler J, Tierney K. Examining adolescents' opioid knowledge and likelihood to utilize an educational game to promote medication safety. Research in Social and Administrative Pharmacy 2021 Sep:1-8. [doi: 10.1016/j.sapharm.2021.09.005]
20. Abraham O, LeMay S, Bittner S, Thakur T, Stafford H, Brown R. Investigating Serious Games That Incorporate Medication Use for Patients: Systematic Literature Review. JMIR Serious Games 2020 Apr 29;8(2):e16096. [doi: 10.2196/16096]

XSL•FO
RenderX

21. Abraham O, Tidd M, Buechel M, Thakur T, Brown R. Student Pharmacists' Assessment of a Serious Game on Opioid Medication Safety. Innov Pharm 2020 Dec 08;11(4):19. [doi: 10.24926/iip.v11i4.2937]
22. Alonso-Fernandez C, Calvo A, Freire M, Martinez-Ortiz I, Fernandez-Manjon B. Systematizing game learning analytics for serious games. 2017 Presented at: IEEE Global Engineering Education Conference (EDUCON); 2017; Athens, Greece p. 1111-1118. [doi: 10.1109/educon.2017.7942988]
23. Serious Games Market Outlook. Allied Market Research. URL: https://www.alliedmarketresearch.com/serious-games-market#:~:text=The%20global%20serious%20games%20market,19.2%25%20from%202017%20to%202023.&text=The%20primary%20function%20of%20serious,users%2C%20rather%20than%20mere%20entertainment [accessed 2021-08-31]
24. Abroms LC, Fishman M, Vo H, Chiang SC, Somerville V, Rakhmanov L, et al. A Motion-Activated Video Game for Prevention of Substance Use Disorder Relapse in Youth: Pilot Randomized Controlled Trial. JMIR Serious Games 2019 May 23;7(2):e11716. [doi: 10.2196/11716]
25. Calderón A, Ruiz M. A systematic literature review on serious games evaluation: An application to software project management. Computers & Education 2015 Sep;87:396-422. [doi: 10.1016/j.compedu.2015.07.011]
26. Abraham O, Thakur T, Brasel K, Norton D, Rosenberger CA. Development of the Adolescent Opioid Safety and Learning (AOSL) scale using exploratory factor analysis. Research in Social and Administrative Pharmacy 2021 Jun:1-8. [doi: 10.1016/j.sapharm.2021.06.007]
27. Linnan S, Walsh-Fez A, Moberg P. Perceptions, awareness, and use of substances in Wisconsin: results of a survey conducted for the Wisconsin Strategic Prevention Framework Partnerships for Success 2015. University of Wisconsin Population Health Institute; February 2018 Feb:1-52 [FREE Full text]
28. Cameron KA, Ross EL, Clayman ML, Bergeron AR, Federman AD, Bailey SC, et al. Measuring patients' self-efficacy in understanding and using prescription medication. Patient Education and Counseling 2010 Sep;80(3):372-376. [doi: 10.1016/j.pec.2010.06.029]
29. Guerin RJ, Toland MD, Okun AH, Rojas-Guyler L, Bernard AL. Using a Modified Theory of Planned Behavior to Examine Adolescents' Workplace Safety and Health Knowledge, Perceptions, and Behavioral Intention: A Structural Equation Modeling Approach. J Youth Adolescence 2018 Mar 31;47(8):1595-1610. [doi: 10.1007/s10964-018-0847-0]
30. Admassu H, Tom S, Sealfon N, Gibson M, Pradel F. Maryland Public Opinion Survey on Opioids: Summary Report. Baltimore (MD), US: University of Maryland, Baltimore; 2015:22-30.
31. Serdarevic M, Striley CW, Cottler LB. Sex differences in prescription opioid use. Curr Opin Psychiatry 2017 Jul;30(4):238-246 [FREE Full text] [doi: 10.1097/YCO.0000000000000337] [Medline: 28426545]
32. McCauley JL, Back SE, Brady KT. Pilot of a brief, web-based educational intervention targeting safe storage and disposal of prescription opioids. Addictive Behaviors 2013 Jun;38(6):2230-2235. [doi: 10.1016/j.addbeh.2013.01.019]
33. McCabe SE, West BT, Veliz P, McCabe VV, Stoddard SA, Boyd CJ. Trends in Medical and Nonmedical Use of Prescription Opioids Among US Adolescents: 1976–2015. Pediatrics 2017 Mar 20;139(4):e20162387. [doi: 10.1542/peds.2016-2387]
34. Hausmann L, Gao S, Lee E, Kwoh K. Racial disparities in the monitoring of patients on chronic opioid therapy. Pain 2013;154(1):46-52. [doi: 10.1016/j.pain.2012.07.034]
35. Santoro T, Santoro J. Racial Bias in the US Opioid Epidemic: A review of the history of systemic bias and implications for care. Cureus 2018;10(12):e3733. [doi: 10.7759/cureus.3733]
36. Singhal A, Tien Y, Hsia RY. Racial-Ethnic Disparities in Opioid Prescriptions at Emergency Department Visits for Conditions Commonly Associated with Prescription Drug Abuse. PLoS ONE 2016 Aug 8;11(8):e0159224. [doi: 10.1371/journal.pone.0159224]
37. King C, Liu X. Racial and Ethnic Disparities in Opioid Use Among US Adults With Back Pain. Spine (Phila Pa 1976) 2020 Aug 01;45(15):1062-1066. [doi: 10.1097/BRS.0000000000003466] [Medline: 32675613]
38. Badreldin N, Grobman W, Yee L. Racial disparities in postpartum pain management. Obstet Gynecol 2019;134(6):1147-1153. [doi: 10.1097/aog.0000000000003561]
39. Anastasiadis T, Lampropoulos G, Siakas K. Digital Game-based Learning and Serious Games in Education. IJASRE 2018;4(12):139-144. [doi: 10.31695/ijasre.2018.33016]
40. Papastergiou M. Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. Computers & Education 2009 Jan;52(1):1-12. [doi: 10.1016/j.compedu.2008.06.004]
41. Huang W, Johnson T. Instructional game design using cognitive load theory. In: Handbook of Research on Effective Electronic Gaming in Education. Hershey, Pennsylvania, USA: IGI Global; 2008:1143-1165.
42. Matson K, Johnson P, Tran V, Horton E, Sterner-Allison J, Advocacy Committee on behalf of Pediatric Pharmacy Advocacy Group. Opioid Use in Children. J Pediatr Pharmacol Ther 2019;24(1):72-75 [FREE Full text] [doi: 10.5863/1551-6776-24.1.72] [Medline: 30837819]

Abbreviations

MUSE: Medication Understanding and Use Self-Efficacy
SG: serious game

XSL•FO
RenderX

Edited by N Zary; submitted 01.10.21; peer-reviewed by N Khalili-Mahani; comments to author 26.10.21; revised version received 27.10.21; accepted 06.11.21; published 23.12.21.

Please cite as:
Abraham O, Rosenberger C, Tierney K, Birstler J
Investigating the Use of a Serious Game to Improve Opioid Safety Awareness Among Adolescents: Quantitative Study
JMIR Serious Games 2021;9(4):e33975
URL: https://games.jmir.org/2021/4/e33975
doi:10.2196/33975
PMID:34747712

©Olufunmilola Abraham, Claire Rosenberger, Kathleen Tierney, Jen Birstler. Originally published in JMIR Serious Games (https://games.jmir.org), 23.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effect of Video Game–Based Interventions on Performance and Cognitive Function in Older Adults: Bayesian Network Meta-analysis

Chao Yang[1*], MD; Xiaolei Han[1*], MD; Mingxue Jin[1*], MD; Jianhui Xu[1], MD; Yiren Wang[2], BA; Yajun Zhang[2], BA; Chonglong Xu[3], BA; Yingshi Zhang[2], MD, PhD; Enshi Jin[1,4], MD; Chengzhe Piao[1,4], MD

[1]Department of Ethnic Culture and Vocational Education, Liaoning National Normal College, Shenyang, China

[2]Department of Clinical Pharmacy, Shenyang Pharmaceutical University, Shenyang, China

[3]Chi Chi Technology LLC, Shenyang, China

[4]Information Construction Department, Liaoning National Normal College, Shenyang, China

[*]these authors contributed equally

Corresponding Author:
Enshi Jin, MD
Department of Ethnic Culture and Vocational Education
Liaoning National Normal College
No. 45 Chongshan Dong Road
Shenyang, 110032
China
Phone: 86 18642065021
Email: jes61@163.com

Abstract

Background: The decline in performance of older people includes balance function, physical function, and fear of falling and depression. General cognitive function decline is described in terms of processing speed, working memory, attention, and executive functioning, and video game interventions may be effective.

Objective: This study evaluates the effect of video game interventions on performance and cognitive function in older participants in terms of 6 indicators: balance function, executive function, general cognitive function, physical function, processing speed, and fear of falling and depression.

Methods: Electronic databases were searched for studies from inception to June 30, 2020. Randomized controlled trials and case-controlled trials comparing video game interventions versus nonvideo game control in terms of performance and cognitive function outcomes were incorporated into a Bayesian network meta-analysis. All data were continuous variables.

Results: In total, 47 studies (3244 participants) were included. In pairwise meta-analysis, compared with nonvideo game control, video game interventions improved processing speed, general cognitive function, and depression scores. In the Bayesian network meta-analysis, interventions with video games improved balance function time (standardized mean difference [SMD] –3.34, 95% credible interval [CrI] –5.54 to –2.56), the cognitive function score (SMD 1.23, 95% CrI 0.82-1.86), processing speed time (SMD –0.29, 95% CrI –0.49 to –0.08), and processing speed number (SMD 0.72, 95% CrI 0.36-1.09), similar to the pairwise meta-analysis. Interventions with video games with strong visual senses and good interactivity ranked first, and these might be more beneficial for the elderly.

Conclusions: Our comprehensive Bayesian network meta-analysis provides evidence that video game interventions could be considered for the elderly for improving performance and cognitive function, especially general cognitive scores and processing speed. Games with better interactivity and visual stimulation have better curative effects. Based on the available evidence, we recommend video game interventions for the elderly.

Trial Registration: PROSPERO CRD42020197158; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=197158

(JMIR Serious Games 2021;9(4):e27058) doi:10.2196/27058

KEYWORDS

video game; performance; cognitive function; older; Bayesian network meta-analysis

XSL•FO
RenderX

Introduction

Background

Since the mid-20th century, with steady improvements in living standards and the growing affluence of global societies, greater longevity has led to great concern over the decline in performance and cognitive abilities that accompanies normal and neuropathological aging [1]. The decline in performance includes balance function, physical function, and fear of falling and depression. General cognitive function decline is described in terms of processing speed, working memory, attention, and executive functioning [2-5]. These demographic changes and neurological aging compromise the sustainability of health care systems, with more health care resources needed to care for the aging population [6]. Therefore, it is essential to propose feasible and acceptable interventions to promote active aging, intended to preserve and optimize opportunities for health, participation, and security in order to enhance the quality of life as people age.

Objective

With the rapid development of computer technology and the video game industry, game types and game experience have greatly improved. A video game is any game played on a digital device, encompassing a wide range of interfaces, including web-based programs and apps for mobile devices [7]. Exergames are video games that require physical activity and movement when played [8]. Video games are promising and adaptable cognitive training tools in the active aging process that can help stave off the negative effects of aging in performance and cognitive function. Moreover, video games are inexpensive and interesting and can be employed in hospitals as well as in the community [9]. Furthermore, older adults are now more digitally connected than ever, and most older people aged 65 years or more have a computer with an internet connection [10]. Therefore, the use and implementation of video game interventions are high.

Although off-the-shelf video games were not developed for serious purposes or these specific interventions, many studies have determined that video game interventions may benefit the elderly in terms of serious purposes, including cognitive function, fall prevention, and other benefits [1,11,12]. No previous systematic review has provided a comprehensive overview with meta-regression and Bayesian network meta-analysis evaluating which type of video game intervention has the best effect on performance and cognitive function.

Methods

This systematic review conformed to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) and the PRISMA extension statement for network meta-analysis [13,14]. We registered the protocol for this Bayesian network meta-analysis with the International Prospective Register of Systematic Reviews (PROSPERO; CRD42020197158) [15].

Search Strategy and Selection Criteria

We conducted systematic literature searches of PubMed, EMBASE, and the Cochrane Library from their inception to June 30, 2020. The MeSH search terms were as follows in the full-text search: “active video game,” “active game,” “older adults,” and “elderly.” We included both randomized controlled trials (RCTs) and case-controlled trials (CCTs) that met the following criteria: participants were older adults (aged ≥60 years) without dementia; the interventions were all kinds of active video games (such as those on Xbox 360 or Nintendo Wii, computer-based games, virtual reality–based games), which means that the participants perform the video game intervention in an active state rather than a static state, and the whole body needs to be involved in the video games; and controls included exercise, puzzle games, visual stimulation, and no game play, which means that even patients who play games are still in a static state. The above declared the difference between the intervention group and the control group in order to study the effect of active video games in the intervention group. Comparisons of interventions with video games versus nonvideo game control were made with reported outcomes of performance and cognitive function.

Outcome Assessment

We assessed the effectiveness of 6 performance and cognitive outcomes by comparing the intervention group and the control group at the final point. The first outcome was balance function, tested using the Berg Balance Scale (BBS), balance test time (s), and balance test speed (m/s). The second outcome was executive function, tested using performance on the trail-making test B (TMT-B, s), delta (s), Stroop word (s), attention, working memory, and the Corsi block test. The third outcome was related to general cognition, tested using the Mini-Mental State Exam (MMSE; score) and the Montreal Cognitive Assessment (MoCA; score). Physical function was the fourth outcome, tested using everyday function and function tests (s and cm). The fifth outcome was processing speed tested using the TMT-A (s) and processing speed (number). The final outcomes were the fear of falling and depression, tested using the Efficacy Scale International (score) and the Geriatric Depression Scale (score). All these outcomes were continuous data.

Data Extraction and Quality Assessment

Two investigators (CY and XH) independently conducted the electronic literature search. The reference lists of relevant publications were also checked, and no language restrictions were set. These 2 researchers evaluated eligible titles, abstracts, and full texts, and disagreements between them were resolved by discussion with a third researcher (author YZ [Yingshi Zhang] or PC). A preset table was designed to extract details of potentially relevant papers, including the first author, publishing year, study type, region, sample size, gender, age, body mass index (BMI), education year, MMSE score, video game type, frequency, period, follow-up, control type, other associated diseases, and community or hospital. Two investigators (authors CY and EJ) extracted all continuous data independently onto a Microsoft Excel spreadsheet. The quality of the included RCTs was evaluated according to Cochrane Collaboration’s tool for assessing the risk of bias [16], and the

XSL•FO
RenderX

quality of the included CCTs was assessed using the Newcastle-Ottawa Scale (NOS) score [17]. We also used the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) scale [18] to evaluate the quality of outcomes. Disagreements were resolved by a third researcher (YZ [Yingshi Zhang] or PC).

Data Synthesis and Statistical Analysis

We applied standardized mean differences (SMDs) and 95% CIs to summarize the 6 outcomes from pairwise meta-analysis. To determine the efficacy of video game intervention versus nonvideo game control, we performed subgroup analysis and meta-regression of various intervention types (ie, Nintendo Wii, Xbox 360, and other video games), types of activity control activities (eg, physical activity, visual stimulation, puzzle games, and no game play), and periods (0-4 weeks, 4-8 weeks, 8-12 weeks, and more than 12 weeks). To determine the heterogeneity among our included studies, $P \leq .05$ or $I^2 > 50\%$ indicated heterogeneity in the outcome. $P < .10$ revealed that a grouping method was a source of heterogeneity. The random effects model was used to ensure the accuracy of the summarized data. Publication bias was assessed using the Begg and Egger tests, where $P \leq .05$ indicated the existence of publication bias [19,20].

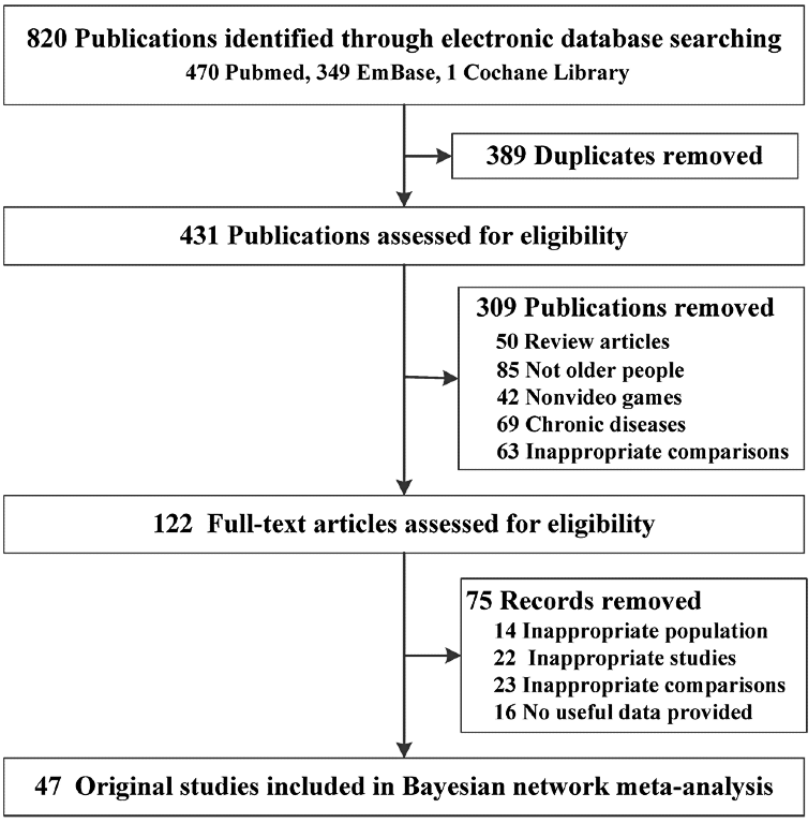
We performed a Bayesian random effects network meta-analysis composed of 4 chains with 100,000 iterations after an initial burnin of 10,000 and a thinning of 2.5 in order to determine the most suitable video game intervention. We calculated the SMDs and corresponding 95% credible intervals (CrIs), and the mean rank and surface under the cumulative ranking curve (SUCRA) values were produced from network meta-analysis estimates with a consistent model. We also produced comparison-adjusted funnel plots to explore publication bias from network meta-analysis. All the aforementioned analyses were performed using StataMP, version 14.0 (StataCorp) and WinBUGS, version 1.4.3 (MRC Biostatistics Unit and Imperial College School of Medicine).

Results

Description of Included Studies

Figure 1 shows details of the selection process. The search strategy generated 820 citations in total. After duplication removal and preliminary screening, 122 potentially publications were scrutinized for eligibility. Finally, we identified 47 original studies [21-66] that met the inclusion criteria. Only 1 (2.1%) of them was a CCT, and the remaining 46 (97.9%) were RCTs. Overall, 1651 of 3244 (50.9%) participants were assigned to the intervention group and the remaining 1593 (49.1%) were assigned to the control group. The sample sizes ranged from 12 to 977 (Table 1 and Multimedia Appendix 1). Baseline characteristics were balanced except that the intervention group was notably older (Table 2). Risk-of-bias assessment was performed for each RCT and CCT, and all included studies had acceptable quality (Multimedia Appendices 2 and 3).

Figure 1. Study selection flowchart.



XSL•FO
RenderX

Table 1. Summarized baseline characteristics of included studies.

Control type	Source, year	Study type	Sample size (I/C)[a]
Intervention type: Xbox 360			
No game play	Rica et al, 2020 [21]	RCT[b]	25/25
No game play	Sápi et al, 2019 [29]	RCT	30/22
No game play	Sato et al, 2015 [51]	RCT	28/26
Normal exercises of upper and lower limbs	Amjad et al, 2019 [22]	RCT	20/18
Adventures and sports	Sápi et al, 2019 [29]	RCT	30/23
Conventional physical therapy	Bacha et al, 2018 [32]	RCT	23/23
Home exercise group	Karahan et al, 2015 [49]	RCT	48/42
Intervention type: Nintendo Wii			
Tai chi chuan	Gatica-Rojas et al, 2019 [25]	CCT	7/5
Traditional exercise	Li et al, 2018 [34]	RCT	49/53
No game play	Montero-Alía et al, 2019 [27]	RCT	508/469
No game play	Zadro et al, 2019 [31]	RCT	30/30
No game play	Gomes et al, 2018 [33]	RCT	15/15
No game play	Franco et al, 2012 [59]	RCT	11/10
No game play	Maillot et al, 2012 [60]	RCT	15/15
No game play	Singh et al, 2012 [64]	RCT	18/18
Fall prevention education	Lee et al, 2017 [37]	RCT	21/19
Same movements	Monteiro-Junior et al, 2017 [38]	RCT	10/9
Gym exercise class	Kwok et al, 2016 [42]	RCT	40/40
Ethylene vinyl acetate copolymer insoles	Jorgensen et al, 2013 [55]	RCT	28/30
Therapeutic balance exercise group	Singh et al, 2013 [57]	RCT	18/18
Seated exercise group	Daniel et al, 2012 [58]	RCT	8/8
Physical activity	Daniel et al, 2012 [58]	RCT	8/7
Completed exercises	Franco et al, 2012 [59]	RCT	11/11
Intervention type: video game training			
Insight process-based intervention	Belchior et al, 2019 [23]	RCT	17/19
No game play	Belchior et al, 2019 [23]	RCT	17/18
No game play	Sosa et al, 2019 [30]	RCT	20/15
No game play	Ordnung et al, 2017 [39]	RCT	14/15
No game play	Toril et al, 2016 [45]	RCT	19/20
No game play	Schoene et al, 2015 [56]	RCT	47/43
No game play	Belchior et al, 2013 [54]	RCT	14/13
No game play	Pichierri et al, 2012 [62]	RCT	11/11
Simulation strategy games	Szelag et al, 2018 [35]	RCT	30/25
Visual stimulation	Buitenweg et al, 2017 [36]	RCT	56+33/50
Common puzzle games	Souders et al, 2017 [40]	RCT	30/30
Balance and stretching training	Eggenberger et al, 2016 [41]	RCT	19/14
Balance and stretching training	Schättin et al, 2016 [44]	RCT	13/14
Knowledge quiz training game	Nouchi et al, 2016 [43]	RCT	36/36
Knowledge quiz training game	van Muijden et al, 2012 [65]	RCT	53/19

XSL•FO
RenderX

Control type	Source, year	Study type	Sample size (I/C)[a]
Education booklet	Gschwind et al, 2015 [48]	RCT	78/75
Community center–based activities	Kim et al, 2015 [50]	RCT	14/14
Physical activity	Whyatt et al, 2015 [53]	RCT	40/42
Placebo training condition	Belchior et al, 2013 [54]	RCT	14/15
Placebo training condition	Nouchi et al, 2012 [61]	RCT	14/14
Usual activities	Schoene et al, 2013 [56]	RCT	15/17
Intervention type: computer-based games			
No game play	Faust et al, 2019 [24]	RCT	25/19
No game play	Perrot et al, 2019 [28]	RCT	12/11
Kawashima brain training	Perrot et al, 2019 [28]	RCT	12/12
Met researchers	Ballesteros et al, 2014 [35]	RCT	17/13
Typical rehabilitation program	Szturm et al, 2011 [66]	RCT	13/14
Intervention type: virtual reality–based games			
Combined physical and cognitive training	Liao et al, 2019 [26]	RCT	18/16
Conventional exercise	Yeşilyaprak et al, 2016 [46]	RCT	7/11
Treadmill memory training	Eggenberger et al, 2015 [47]	RCT	24/22
Treadmill walking	Eggenberger et al, 2015 [47]	RCT	24/25
No game play	Rendon et al, 2012 [63]	RCT	20/20

[a]I: intervention group; C: control group.
[b]RCT: randomized controlled trial.

Table 2. Balance status of baseline characteristics (italics indicate a significant difference).

Baseline indicator	SMD[a]/OR[b] (95% CI)	*P* value	I^2 (%)	Baseline balanced
Age	*0.236 (0.018-0.454)*	<.001	87.1	No
Gender	0.939 (0.757-1.163)	.42	2.9	Yes
Body mass index	–0.113 (–0.311 to 0.084)	<.001	62.6	Yes
Education	0.026 (–0.161 to 0.213)	.24	19.1	Yes
Mini-Mental State Exam	0.023 (–0.172 to 0.217)	.19	25.2	Yes

[a]SMD: standardized mean difference.
[b]OR: odds ratio.

Outcomes of Pairwise Meta-analysis

For the outcomes of balance function, we first compared balance scores (BBS, etc) pre- and postintervention. No significant differences were found overall and in subgroups by intervention and control types. Significant differences were only found at 8-12 weeks. Substantial heterogeneity was found overall and in all subgroups. The source of heterogeneity was not determined from meta-regression, and publication bias was occasionally found. For the balance test time (s), no significant differences were found overall and in all subgroup meta-analyses, with frequent substantial heterogeneity and no publication bias. In balance test speed (m/s) assessment, significant differences were found only for no game play as the control type (odds ratio [OR] 0.611, 95% CI 0.048-1.175). These findings suggest that video game interventions may improve balance in the elderly (Table 3); however, the most suitable intervention has not yet been determined.

For assessment of executive function, we first evaluated the results of the TMT-B (s). No significant differences were found in all subgroups, with frequent substantial heterogeneity. Second, no significant difference was found in the outcome of delta (s). Similarly, no significant differences were found in the subgroup and overall outcomes of the Stroop word test (s), with substantial heterogeneity. When we evaluated the attention in video games compared with nonvideo games, significant differences were found only in the period of more than 12 weeks, with only 1 original study. In the evaluation of working memory, significant differences were detected for overall outcomes (OR=1.034, 95% CI 0.305-1.763), other video games as the intervention type (OR=1.076, 95% CI 0.295-1.858), no game play as the control type (OR 1.023, 95% CI 0.133-1.914),

XSL•FO
RenderX

and 4-8 weeks as the intervention period (OR 1.401, 95% CI 0.559-2.243), with substantial moderate-grade heterogeneity in these subgroups. For the Corsi block test, significant differences were found for no game play and intervention periods of 0-4 and 8-12 weeks. Only the outcome for 8-12 weeks in 4 studies showed no heterogeneity (OR 0.429, 95% CI 0.080-0.778, P=.42, I^2=0.0%) with high quality (Table 4). In summary, video game interventions had little effect on executive function, except for memory-related functions.

Table 3. Summarized outcomes from pairwise meta-analysis of performance and cognitive data for balance function (italics indicate a significant difference).

Variable	Subgroup	Studies (n)	OR[a] (95% CI)	*P* value, I^2	Meta-regression	Grade	Publication bias, *P* value
BBS[b]							
Overall	N/A[c]	26	0.213 (–0.025 to 0.451)	<.001, 83.5%[d]	0.362	Moderate	0.499, .02
Intervention type	Nintendo Wii	14	0.187 (–0.059 to 0.432)	<.001, 74.0%[d]	—[e]	Moderate	0.547, .02
Intervention type	Xbox 360	3	0.419 (–0.133 to 0.971)	.03, 71.1%[d]	—	Low	0.602, .35
Intervention type	Other video games	9	0.093 (–0.664 to 0.850)	<.001, 89.7%[d]	—	Moderate	0.420, .85
Control type	Activity control	13	0.190 (–0.290 to 0.669)	<.001, 85.5%[d]	0.812	Moderate	0.870, .46
Control type	No game play	13	0.159 (–0.095 to 0.413)	<.001, 75.5%[d]		Moderate	0.493, .03
Period	0-4 weeks	7	0.127 (–0.192, 0.446)	.73, 0.0%	0.902	High	0.881, .41
Period	4-8 weeks	16	0.362 (–0.103 to 0.826)	<.001, 87.1%[d]	—	Moderate	0.719, .99
Period	8-12 weeks	3	*–0.178 (–0.265 to 0.090)*	.73, 0.0%	—	Moderate	0.602, .31
Balance test (s)							
Overall	N/A	22	–0.090 (–0.348 to 0.168)	<.001, 72.4%	—	—	0.310, .26
Intervention type	Nintendo Wii	9	–0.364 (–0.846 to 0.119)	<.001, 78.4%[d]	0.766	Moderate	0.144, .27
Intervention type	Xbox 360	4	0.084 (–0.613 to 0.780)	<.001, 85.7%[d]	—	Moderate	0.404, .36
Intervention type	Other video games	9	0.125 (–0.135 to 0.385)	.23, 24.2%	—	High	1.000, .66
Control type	Activity control	16	–0.081 (–0.317 to 0.156)	.001,59.2%[d]	0.968	Moderate	0.458, .32
Control type	No game play	6	–0.156 (–1.128 to 0.816)	<.001, 89.1%[d]	—	Moderate	0.624, .53
Period	0-4 weeks	1	–0.515 (–1.780 to 0.750)	—	0.565	Low	—
Period	4-8 weeks	13	–0.105 (–0.372 to 0.162)	.02, 50.9%[d]	—	Moderate	0.903, .90
Period	8-12 weeks	5	–0.100 (–0.923 to 0.723)	<.001, 91.2%[d]	—	Moderate	0.881, .59
Period	More than 12 weeks	3	0.192 (-0.116 to 0.499)	.36, 2.4%	—	Moderate	—
Balance test speed (m/s)							
Overall (all 4-8 weeks)	N/A	5	–0.046 (–0.970 to 0.878)	<.001, 84.5%[d]	—	Moderate	0.355, <.001
Intervention type	Nintendo Wii	3	–1.101 (–3.627 to 1.426)	<.001, 91.9%[d]	0.551	Moderate	0.117, .03
Intervention type	Xbox 360	2	0.394 (–0.021 to 0.808)	.29,10.9%	—	Moderate	—
Control type	Activity control	4	–0.449 (–1.796 to 0.898)	<.001, 88.1%[d]	0.660	Moderate	0.174, .22
Control type	No game play	1	*0.611 (0.048-1.175)*	—	—	Low	—

[a]OR: odds ratio.

[b]BBS: Berg Balance Scale.

[c]N/A: not applicable.

[d]Substantial heterogeneity.

[e]—: not applicable.

XSL•FO
RenderX

Table 4. Summarized outcomes from pairwise meta-analysis of performance and cognitive data for executive function (italics indicate a significant difference).

Variable	Subgroup	Studies (n)	OR[a] (95% CI)	*P* value, I^2	Meta-regression	Grade	Publication bias, *P* value
TMT-B[b] (s)							
Overall	N/A[c]	16	–0.203 (–0.807 to 0.402)	<.001,93.6%[d]	—[e]	Moderate	0.392, .75
Intervention type	Other video games	15	0.077 (–0.461 to 0.615)	<.001, 91.9%[d]	—	Moderate	0.729, .52
Control type	Activity control	13	–0.209 (–0.972 to 0.554)	<.001, 94.9%[d]	0.995	Moderate	0.542, .78
Control type	No game play	3	–0.100 (–0.424 to 0.223)	.98, 0.0%	—	Moderate	0.602, .29
Period	0–4 weeks	2	1.095 (–2.738 to 4.927)	<.001, 98.1%[d]	0.246	Low	0.317, —
Period	4–8 weeks	5	–0.141 (–2.476 to 2.193)	<.001, 96.7%[d]	—	Moderate	0.624, .88
Period	8–12 weeks	4	–0.083 (–0.371 to 0.205)	.52, 0.0%	—	Moderate	1.000, .22
Period	More than 12 weeks	4	–0.684 (–1.487 to 0.119)	<.001, 91.3%[d]	—	Moderate	0.042, .03
Delta (s)							
Overall	N/A	3	–1.780 (–3.758 to 0.198)	<.001, 93.8%[d]	—	Low	0.296, .12
Stroop word (s)							
Overall (all other video games)	N/A	14	0.050 (–0.275 to 0.374)	<.001, 76.6%[d]	—	Moderate	0.261, .83
Control type	Activity control	10	0.135 (–0.281 to 0.552)	<.001, 80.6%[d]	0.541	Moderate	0.245, .95
Control type	No game play	4	–0.169 (–0.640 to 0.301)	.01, 54.7%[c]	—	Moderate	1.000, .82
Period	0–4 weeks	1	0.530 (0.060-1.001)	—	0.669	—	0.317, —
Period	4–8 weeks	6	0.015 (–0.787 to 0.818)	<.001, 87.2%[d]	—	Moderate	0.624, .15
Period	8–12 weeks	5	0.039 (–0.475 to 0.554)	.02, 66.3%[d]	—	Moderate	0.327, .21
Period	More than 12 weeks	2	–0.017 (–0.268 to 0.235)	.92,0.0%	—	Moderate	0.317, —
Attention							
Overall (all other video games)	N/A	9	0.185 (–1.607 to 1.976)	<.001, 97.7%[d]	—	Moderate	0.137, .32
Control type	Activity control	7	1.317 (–0.158 to 2.792)	<.001, 96.8%[d]	0.116	Moderate	0.099, .37
Control type	No game play	2	–17.991 (–64.346 to 28.364)	<.001, 99.3%[d]	—	Low	0.317, —
Period	0–4 weeks	1	0.298 (–0.424 to 1.019)	—	0.568	—	—
Period	4–8 weeks	3	–7.497 (–13.386 to 0.608)	<.001, 98.5%[d]	—	Moderate	0.117, .02
Period	8–12 weeks	4	2.130 (–1.520 to 5.781)	<.001, 98.0%[d]	—	Moderate	0.042, .02
Period	More than 12 weeks	1	*0.381 (0.061-0.701)*	—	—	—	—
Working memory							
Overall	N/A	12	*1.034 (0.305-1.763)*	<.001, 92.3%[d]		Moderate	0.501, .38
Intervention type	Nintendo Wii	1	0.559 (–0.361 to 1.479)	—	—	—	—
Intervention type	Other video games	11	*1.076 (0.295-1.858)*	<.001, 93.0%[d]		Moderate	0.139, .17
Control type	Activity control	8	1.045 (–0.012 to 2.102)	<.001, 94.2%[d]	0.985	Moderate	0.322, .28
Control type	No game play	4	*1.023 (0.133-1.914)*	<.001, 85.4%[d]	—	Moderate	0.497, .82

XSL•FO
RenderX

Variable	Subgroup	Studies (n)	OR[a] (95% CI)	*P* value, I^2	Meta-regression	Grade	Publication bias, *P* value
Period	0-4 weeks	2	1.885 (–1.713 to 5.482)	<.001, 98.2%[d]	0.660	Low	0.317, —
Period	4-8 weeks	3	*1.401 (0.559-2.243)*	.01, 79.4%[d]	—	Moderate	0.602, .86
Period	8-12 weeks	4	0.016 (–0.841 to 0.874)	.002, 79.9%[d]	—	Moderate	1.000, .94
Period	More than 12 weeks	2	1.961 (–1.186 to 5.107)	.002, 79.9%[d]	—	Low	0.317, —
Corsi block test							
Overall (all other video games)	N/A	5	1.120 (–0.077 to 2.316)	<.001, 92.3%[d]	—	—	0.221, .38
Control type	Activity control	4	1.137 (–0.352 to 2.627)	<.001, 94.2%[d]	0.971	Moderate	0.174, .44
Control type	No game play	1	*1.064 (0.185-1.944)*	—	—	—	—
Period	0–4 weeks	1	*3.738 (2.883-4.594)*	—	0.006[f]	—	0.497, .54
Period	8–12 weeks	4	*0.429 (0.080-0.778)*	.42, 0.0%	—	High	—

[a]OR: odds ratio.

[b]TMT: trail-making test.

[c]N/A: not applicable.

[d]Substantial heterogeneity.

[e]—: not applicable.

[f]Sources of heterogeneity.

Regarding general cognition, the first item assessed was the MMSE score. A significant difference was found with Xbox 360 as the intervention type and 4-8 weeks as the intervention period. The second item evaluated was the MoCA score. A significant difference was only found for activity control as the control type (OR 1.826, 95% CI 0.043-3.609), with substantial heterogeneity (Table 5). With respect to physical function, first, in the everyday function outcome, a significant difference was found only in the intervention period of 4-8 weeks (OR –1.045, 95% CI –1.866 to –0.223), with substantial heterogeneity. A significant difference was found in the function test (cm) outcome (OR 0.725, 95% CI 0.235-1.214), with low heterogeneity (*P*=.26, I^2=25.5%). No significant difference was found for the function test (s) outcome (Table 6). In summary, video game intervention may have an effect on improving general cognitive function, but it had little effect on physical function.

We considered processing speed after video game intervention versus nonvideo game control. First, when we evaluated the TMT-A (s), significant differences were found for overall outcomes (OR –0.833, 95% CI –1.463 to –0.204), other video games (OR –0.874, 95% CI –1.558 to –0.190), activity control (OR –1.033, 95% CI –1.830 to –0.235), and for more than 12 weeks (OR –2.395, 95% CI –4.272 to –0.519), with substantial heterogeneity. In terms of processing speed (number), significant differences were found overall and in subgroups, while substantial heterogeneity was found overall (OR 1.084, 95% CI 0.765-1.402, *P*<.001, I^2=95.5%) and for activity control (Table 7). In Falls Efficacy Scale assessment, no significant differences were found overall and in all subgroups. In the Geriatric Depression Scale assessment, significant differences were found for Xbox 360 and other video games (OR –0.651, 95% CI –1.164 to –0.138) as the intervention types and 0-4 and 8-12 weeks as the intervention periods (OR –1.800, 95% CI –2.745 to –0.854), with substantial heterogeneity (Table 8). In summary, video game intervention had a significant advantage in terms of processing speed and had a tendency to reduce depression scores.

XSL•FO
RenderX

Table 5. Summarized outcomes from pairwise meta-analysis of performance and cognitive data for general cognition (italics indicate a significant difference).

Variable	Subgroup	Studies (n)	OR[a] (95% CI)	*P* value, I^2	Meta-regression	Grade	Publication bias, *P* value
MMSE[b] (score)							
Overall (all activity control)	N/A[c]	3	1.557 (–0.459 to 3.572)	<.001, 95.5%[d]	—[e]	Low	0.296, .37
Intervention type	Xbox 360	1	*4.606 (3.366-5.846)*	—	0.095[f]	—	—
Intervention type	Other video games	2	0.215 (–0.175 to 0.605)	.60, 0.0%	—	Moderate	0.317, —
Period	0–4 weeks	2	0.215 (–0.175 to 0.605)	.60, 0.0%	0.095[f]	Moderate	0.317, —
Period	4–8 weeks	1	*4.606 (3.366-5.846)*	—	—	—	—
MoCA[g] (score)							
Overall	N/A	5	1.296 (–0.102 to 2.693)	<.001, 93.9%[d]	—	Moderate	0.624, .01
Intervention type	Nintendo Wii	1	–0.416 (–1.140 to 0.308)	—	0.419	—	—
Intervention type	Xbox 360	2	3.796 (–4.023 to 11.616)	<.001, 98.4%[d]	—	Low	0.317, —
Intervention type	Other video games	2	0.442 (–0.052 to 0.936)	.99, 0.0%	—	Moderate	0.317, —
Control type	Activity control	4	*1.826 (0.043-3.609)*	<.001, 95.1%[d]	0.594	Moderate	0.317, —
Control type	No game play	1	–0.416 (–1.140 to 0.308)	—	—	—	—
Period	0-4 weeks	1	–0.139 (–0.717 to 0.440)	—	0.649	—	0.734, .04
Period	4-8 weeks	4	1.800 (–0.126 to 3.727)	<.001, 95.2%[d]	—	Moderate	—

[a]OR: odds ratio.

[b]MMSE: Mini-Mental State Exam.

[c]N/A: not applicable.

[d]Substantial heterogeneity.

[e]—: not applicable.

[f]Sources of heterogeneity.

[g]MoCA: Montreal Cognitive Assessment.

XSL•FO
RenderX

Table 6. Summarized outcomes from pairwise meta-analysis of performance and cognitive data for physical function (italics indicate a significant difference).

Variable	Subgroup	Studies (n)	OR[a] (95% CI)	*P* value, I^2	Meta-regression	Grade	Publication bias, *P* value
Everyday function							
Overall	N/A[b]	12	–0.014 (–0.538 to 0.510)	<.001, 86.3%[c]	—[d]	Moderate	0.837, .90
Intervention type	Nintendo Wii	3	–1.024 (–2.645 to 0.597)	<.001, 87.3%[c]	0.113	Low	0.602, .94
Intervention type	Xbox 360	1	–0.016 (–0.594 to 0.562)	—	—	—	—
Intervention type	Other video games	8	0.306 (–0.283 to 0.895)	<.001, 86.1%[c]	—	Moderate	0.322, .41
Control type	Activity control	9	0.138 (–0.418 to 0.695)	<.001, 83.9%[c]	—	Moderate	0.532, .64
Control type	No game play	3	–0.544 (–2.187 to 1.098)	<.001, 93.3%[c]	—	Low	0.602, .20
Period	0-4 weeks	2	0.568 (–0.634 to 1.769)	.01, 83.6%[c]	0.760	Low	0.317, —
Period	4-8 weeks	4	*–1.045 (–1.866 to 0.223)*	.003, 78.8%[c]	—	Moderate	0.042, .01
Period	8-12 weeks	3	1.076 (–0.125 to 2.277)	<.001, 88.3%[c]	—	Low	0.117, .14
Period	More than 12 weeks	3	–0.035 (–0.324 to 0.255)	.61, 0.0%	—	Moderate	0.117, .62
Function test (s)							
Overall	N/A	4	–0.284 (–1.735 to 1.168)	<.001, 95.3%[c]	—	Moderate	0.308, .22
Function test (cm)							
Overall	N/A	3	*0.725 (0.235-1.214)*	0.26, 25.5%	—	Moderate	1.000, .85

[a]OR: odds ratio.

[b]N/A: not applicable.

[c]Substantial heterogeneity.

[d]—: not applicable.

XSL•FO
RenderX

Table 7. Summarized outcomes from pairwise meta-analysis of performance and cognitive data for processing speed (italics indicate a significant difference).

Variable	Subgroup	Studies (n)	OR[a] (95% CI)	*P* value, I^2	Meta-regression	Grade	Publication bias, *P* value
TMT-A[b] (s)							
Overall	N/A[c]	11	–0.833 (–1.463 to *–0.204)*	<.001, 93.0%[d]	—[e]	Moderate	0.020, .01
Intervention type	Xbox 360	1	–0.531 (–1.180 to 0.117)	—	0.838	—	—
Intervention type	Other video games	10	–0.874 (–1.558 to *–0.190)*	<.001, 93.6%[d]	—	Moderate	0.009, .01
Control type	Activity control	9	–1.033 (–1.830 to *–0.235)*	<.001, 94.4%[d]	0.551	Moderate	0.007, .01
Control type	No game play	2	–0.163 (–0.516 to 0.189)	.87, 0.0%	—	Moderate	0.317, —
Period	4–8 weeks	4	–0.272 (–0.616 to 0.072)	.49, 0.0%	0.086[f]	High	0.497, .85
Period	8–12 weeks	3	0.132 (–0.133 to 0.398)	.53, 0.0%	—	Moderate	0.117, .40
Period	More than 12 weeks	4	*–2.395 (–4.272* to *0.519)*	<.001, 97.7%[d]	—	Moderate	0.042, .004
Processing speed (number)							
Overall (all other video games)	N/A	7	*1.084 (0.765-1.402)*	<.001, 95.5%[d]	—	Moderate	0.096, <.001
Control type	Activity control	5	*1.374 (0.960-1.789)*	<.001, 96.9%[d]	0.413	Moderate	0.142, .001
Control type	No game play	2	*0.665 (0.168-1.163)*	.42, 0.0%	—	Moderate	0.317, —
Period	0–4 weeks	1	*0.842 (0.091-1.592)*	—	0.024[f]	—	—
Period	8–12 weeks	4	*0.536 (0.170-0.903)*	.74, 0.0%	—	High	0.497, .55
Period	More than 12 weeks	2	*8.169 (6.916-9.423)*	.94,0.0%	—	Moderate	0.317, —

[a]OR: odds ratio.

[b]TMT: trail-making test.

[c]N/A: not applicable.

[d]Substantial heterogeneity.

[e]—: not applicable.

[f]Sources of heterogeneity.

XSL•FO
RenderX

Table 8. Summarized outcomes from pairwise meta-analysis of performance and cognitive data for fear of falling and depression (italics indicate a significant difference).

Variable	Subgroup	Studies (n)	OR[a] (95% CI)	*P* value, I^2	Meta-regression	Grade	Publication bias, *P* value
Falls Efficacy Scale International (score)							
Overall	N/A[b]	7	–0.990 (–3.003 to 1.022)	<.001, 99.2%[c]	0.628	—[d]	0.776, .11
Intervention type	Nintendo Wii	4	–1.305 (–4.232 to 1.621)	<.001, 99.4%[c]	0.628	Moderate	1.000, .04
Intervention type	Other video games	3	–0.539 (–1.651 to 0.574)	<.001, 87.7%[c]	—	Low	0.602, .50
Control type	Activity control	4	–0.450 (–1.143 to 0.242)	.001, 81.9%[c]	0.446	Moderate	0.497, .28
Control type	No game play	3	–1.641 (–5.240 to 1.959)	<.001, 99.4%[c]	—	Low	0.602, .20
Period	4–8 weeks	5	–0.314 (–0.858 to 0.229)	.003, 75.6%[c]	0.156	Moderate	0.327, .26
Period	8–12 weeks	2	–2.534 (–6.909 to 1.841)	<.001, 99.6%[c]	—	Low	0.317, —
Geriatric Depression Scale (score)							
Overall	N/A	10	–0.393 (–1.058 to 0.273)	<.001, 91.5%[c]	—	—	1.000, .41
Intervention type	Nintendo Wii	4	0.683 (–0.797 to 2.164)	<.001, 94.8%[c]	0.876	Moderate	0.174, .02
Intervention type	Xbox 360	1	*–2.742 (–3.521 to 1.962)*	—	—	—	—
Intervention type	Other video games	5	*–0.651 (–1.164 to 0.138)*	.004, 74.0%[c]	—	Moderate	0.142, .43
Control type	Activity control	5	0.135 (–0.941 to 1.211)	<.001, 93.3%[c]	0.352	Moderate	0.142, .07
Control type	No game play	5	–0.866 (–1.825 to 0.094)	<.001, 91.3%[c]	—	Moderate	0.050, .34
Period	0–4 weeks	1	*–0.637 (–1.111 to 0.163)*	—	0.403	—	—
Period	4–8 weeks	5	0.552 (–0.604 to 1.708)	<.001,93.3%[c]	—	Moderate	0.327, .004
Period	8–12 weeks	3	*–1.800 (–2.745 to 0.854)*	.01,79.1%[c]	—	Moderate	0.117, .35
Period	More than 12 weeks	1	–0.289 (–0.705 to 0.127)	—	—	—	—

[a]OR: odds ratio.

[b]N/A: not applicable.

[c]Substantial heterogeneity.

Outcomes of Network Meta-analysis

Using pairwise meta-analysis, we observed that the video game intervention improved clinical performance and cognitive function, especially processing speed and depression scores, in the elderly. However, the most suitable video game intervention has not been determined. Moreover, many subgroups included fewer studies, which may have affected the accuracy of the results. Therefore, we selected the outcomes of the included studies to conduct Bayesian network meta-analysis among the 6 indicators to identify the most suitable type of video game intervention for the elderly. Figure 2 provides network plots of balance time with the intervention period (Figure 2A) and without the intervention period (Figure 2B), because most studies were included in this outcome.

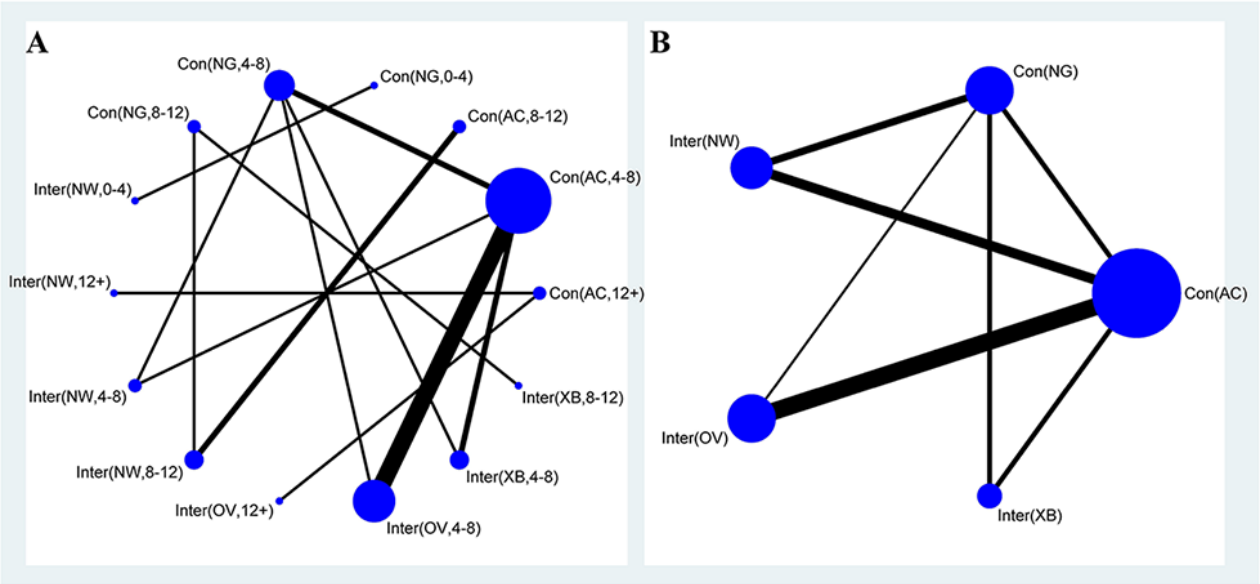
For the indicator of balance function, first, in terms of balance time (s), compared with no game play as the control group, Xbox 360 as the intervention type ranked first, with a significant difference (SMD –3.34, 95% CrI –5.54 to –2.56), followed by other video games (SMD –1.15, 95% CrI –2.69 to –0.64), Nintendo Wii, and activity control. Significant differences could also be found in Xbox 360 versus other video games as the intervention type (SMD –1.62, 95% CrI –4.89 to –1.03), Nintendo Wii (SMD –4.14, 95% CrI –9.64 to –0.28), and activity control. For the BBS, significant differences were found only in other video games compared with Nintendo Wii as the intervention type (SMD 0.38, 95% CrI 0.03-1.79), with some publication bias (Multimedia Appendix 4). In summary, the intervention typical of Xbox 360 and other video games may be the best intervention method to maintain balance function over time (Figure 2A).

For general cognition, we combined the scores of MMSE and MoCA. Compared with no game play as the control type, which ranked the lowest, other video games as the intervention type ranked the highest, with a significant difference (SMD 1.23, 95% CrI 0.82-1.86), followed by Xbox 360 (SMD 1.13, 95% CrI 0.74-1.72), Nintendo Wii (SMD 0.90, 95% CrI 0.59–1.37), and activity control (SMD 1.44, 95% CrI 0.94-2.20). Significant

XSL•FO
RenderX

differences were found in all comparisons, which suggests that the ordering from our Bayesian network meta-analysis was reasonable. No significant differences were found among the network outcomes from the TMT-B (s) to determine executive function (Figure 2B). In summary, video game intervention improved the cognitive function of elderly patients but had no effect on executive function.

Figure 2. Network plot for all interventions and controls for balance time with (A) and without (B) the intervention period in older adults. Each circular node represents an intervention/control type. The circle size is proportional to the total number of participants, while the line width is proportional to the number of studies used in the head-to-head comparisons. AC: activity control; NG: no game play; NW: Nintendo Wii; OV: other video games; XB: Xbox.



In the assessment of processing speed, the first item examined was the TMT-A (s). Compared with no game play, other video games ranked first, with a significant difference (SMD –0.29, 95% CrI –0.49 to –0.08), followed by activity control (SMD –0.36, 95% CrI –0.57 to –0.15), virtual reality–based games, and Xbox 360. Significant differences were also found in comparisons of other video games and activity control versus virtual reality–based games and Xbox 360. The second item assessed was processing speed (number). Other video games as the intervention type also ranked first, with a significant difference compared with no game play as the control (SMD 0.72, 95% CrI 0.36-1.09), followed by virtual reality–based games (SMD 0.60, 95% CrI 0.26-0.94) and activity control (SMD 0.42, 95% CrI 0.06-0.77). In general, video game intervention improved processing speed (Figure 3A). For the evaluation of depression scales, compared with no game play, Xbox 360 ranked the highest, followed by other video games, activity control, and Nintendo Wii, with no significant difference. In terms of the Falls Efficacy Scale, there was no significant difference (Figure 3B).

In conclusion, based on Bayesian network meta-analysis, we determined that video game intervention improves balance function, cognitive function, and processing speed, which were similar results to those obtained using pairwise meta-analysis.

XSL•FO
RenderX

Figure 3. Summary effects from Bayesian network meta-analysis for balance time and BBS (A) and cognitive score and executive function (B) are ranked by the mean rank and the SUCRA score. Information relating to the SMDs and 95% CrIs is listed in the columns, with the rows displaying the intervention identity. SMD values higher than 0 favor the column-defining intervention (ie, the left-most in order), indicating improvement in effectiveness. *Statistical significance. AC: activity control; BBS: Berg Balance Scale; CrI: credible interval; NG: no game play; NW: Nintendo Wii; OV: other video games; SMD: standardized mean difference; SUCRA: surface under the cumulative ranking curve; trail-making test; XB: Xbox.

A

Con(NG)	1.01 (-1.16, 2.30)	-0.54 (-1.02, 1.27)	1.01 (-1.16, 2.30)	-11.67 (-30.49, 34.53)
-2.16 (-4.42, 0.37)	Con(AC)	-11.19 (-31.51, 33.76)	-0.80 (-2.84, 3.02)	11.10 (-36.76, 30.69)
0.86 (-2.39, 4.10)	3.07 (-0.25, 5.92)	Inter(NW)	0.38 (0.03, 1.79)*	11.19 (-33.76, 31.51)
-1.15 (-2.69, -0.64)*	0.48 (-1.01, 2.78)	-2.53 (-4.75, 0.75)	Inter(OV)	10.88 (-33.92, 29.72)
-3.34 (-5.54, -2.56)*	-1.31 (-5.53, 1.75)	-4.14 (-9.64, -0.28)*	-1.62 (-4.89, -1.03)*	Inter(XB)

Intervention/Control Balance time Balance scale

B

Con(NG)	-0.30 (-0.80,0.19)	-	-0.18 (-0.66,0.29)	-0.33 (-0.80,0.13)
1.44 (0.94,2.20)*	Con(AC)	-	0.12 (-0.35,0.59)	-0.03 (-0.52,0.46)
0.90 (0.59,1.37)*	0.63 (0.41,0.96)*	Inter(NW)	-	-
1.13 (0.74,1.72)*	0.78 (0.51,1.19)*	0.80 (0.52,1.22)*	Inter(XB)	-0.15 (-0.62,0.32)
1.23 (0.82,1.86)*	0.85 (0.56,1.29)*	1.37 (0.90,2.07)*	1.09 (0.72,1.65)*	Inter(OV)

Intervention/Control Cognitive score Executive functions(TMT-B)

Discussion

Principal Findings

Our Bayesian network meta-analysis quantified the comparative effectiveness of video games based on 47 studies including 3244 elderly participants, with acceptable quality. We comprehensively summarized the comparative efficacy of video games in improving performance and cognitive function in 6 domains: balance function, executive function, general cognition, physical function, processing speed, and fear of falling and depression. The results suggested that, first, on pairwise meta-analysis, video game interventions are beneficial for cognition scores, processing speed, and depression scores. There are tendencies toward benefits for balance function, executive function, and physical function. Second, on Bayesian network meta-analysis, interventions with video games may improve balance function, cognitive scores, and processing speed in the elderly, which was similar to the results of pairwise meta-analysis. Third, from the ranking of the Bayesian network meta-analysis, Xbox 360 and other video games always ranked first, while Nintendo Wii always ranked last of all interventions. This was accompanied by having the most outcomes of moderate GRADE with low publication bias in both pairwise and Bayesian network meta-analyses.

In this systematic review, we used a comprehensive search with clear inclusion and exclusion criteria and carefully examined the efficacy of video game interventions in improving performance and cognitive function in 6 domains and 18 outcomes. Generally, the consistency of Bayesian network meta-analysis was similar to that of pairwise meta-analysis. On Bayesian network meta-analysis, compared with no game play as the control type, significant differences were found in balance time (s), cognitive scores, processing speed (TMT-A), and processing speed (number). For expert balance time (s), significant differences were also found in terms of cognitive scores, processing speed (TMT-A), and processing speed (number) on pairwise meta-analysis (Table 2; Figures 3 and 4). Based on the results of our study, video game interventions had the most obvious benefit for cognitive scores and processing speed. Processing speed is defined as the time spent completing mental tasks. It relates to the patient's speed of understanding the information they obtain, whether it is visual (letters and numbers), auditory (language), or mobile. Similar results were found by Ozdogar et al [67], suggesting that video-based exergaming is almost as effective as conventional rehabilitation with respect to improving walking, upper- and lower-extremity functions, cognitive function, fatigue, depression, and health-related quality of life. Shin et al [68] found that participants who frequently played video games showed enhanced processing speed, which could be an effect of game practice. Mansor et al [2] determined limited effects of video games on cognitive function, and another valuable research published by Wang et al [69] also proved that game-based brain training can be considered a supplementary intervention for improving cognitive function in community-dwelling older adults. Moreover, Vázquez et al's [70] research indicates that video game–based interventions may assist adults in active aging processes and prevent secondary aging. The above valuable studies all support our results.

In our study, we found that other video games and Xbox 360 are more effective than Nintendo Wii. Other types of video games were defined as exergames, video games, computer-based games, and virtual reality–based games. The main reason for these results may be that Xbox 360 and other video game screens require equipment to play, which have good platforms, a strong visual sense, and good interactivity. However, they are not easy to carry because of the requirement of external equipment.

XSL•FO
RenderX

Figure 4. Summary effects from Bayesian network meta-analysis for processing speed (A) and depression and falls efficacy (B) are ranked by the mean rank and the SUCRA score. Information relating to the SMDs and 95% CrIs is listed in the columns, with the rows displaying the intervention identity. SMD values higher than 0 favor the column-defining intervention (ie, the left-most in order), indicating improvement in effectiveness. *Statistical significance. AC: activity control; CrI: credible interval; NG: no game play; NW: Nintendo Wii; OV: other video games; SMD: standardized mean difference; SUCRA: surface under the cumulative ranking curve; trail-making test; XB: Xbox.

A

Con(NG)	-	0.13 (-0.23,0.48)	0.31 (-0.06,0.68)	0.72 (0.36,1.09)*
-0.05 (-0.26,0.17)	Inter(XB)	-	-	-
-0.07 (-0.29,0.14)	-0.02 (-0.23,0.18)	Inter(VR)	0.18 (-0.17,0.53)	0.60 (0.26,0.94)*
-0.36 (-0.57,-0.15)*	-0.31 (-0.52,-0.11)*	-0.29 (-0.48,-0.08)*	Con(AC)	0.42 (0.06,0.77)*
-0.29 (-0.49,-0.08)*	-0.24 (-0.45,-0.03)*	-0.21 (-0.43,0.00)*	0.08 (-0.13,0.28)	Inter(OV)

Intervention/Control | Processing speed (TMT-A) | Processing speed (No.)

B

Con(NG)	-0.05 (-0.32,0.21)	-0.21 (-0.49,0.06)	-0.20 (-0.47,0.07)	-
-0.02 (-0.67,0.63)	Inter(NW)	-0.16 (-0.44,0.12)	-0.15 (-0.42,0.12)	-
-0.28 (-0.95,0.38)	-0.26 (-0.92,0.41)	Con(AC)	0.01 (-0.26,0.28)	-
-0.28 (-0.94,0.37)	-0.26 (-0.91,0.39)	0.00 (-0.67,0.66)	Inter(OV)	-
-0.33 (-0.98,0.33)	-0.31 (-0.96,0.35)	-0.05 (-0.71,0.62)	-0.04 (-0.70,0.61)	Inter(XB)

Intervention/Control | Depression scale | Falls efficacy scale

A previous meta-analysis did not distinguish between different video games; therefore, this is an innovation of our study. A previous study suggested that after Nintendo Wii therapy, patients experienced motor learning retention, achieving a sustained benefit using the technique [71]. We observed that Nintendo Wii only benefited patients in general; therefore, when we choose video game interventions for elderly participants, we should choose better-intervention video game types with better interaction and visual stimulation for older adults. Video games could also improve mental health among older adults [72].

The mechanisms underlying changes following video game interventions remain unclear, although they might be related to tonic/phasic activation or inhibition of affected brain regions during video game playing [73] that make the participants feel satisfied. Specifically, psychological satisfaction and pleasure might be related to the various feedback mechanisms provided to the player by the active video game. This elaborate reinforcement and reward schedule has the potential to maximize motivation [74]. Video games are able to maintain flexibility of striatal responses to reward, a mechanism that might be extremely important to keep motivation high and therefore might be of critical value for many different applications, including cognitive training and therapeutic possibilities [75]. In the studies we included, only a few participants felt fatigue or leg muscle soreness; however, the players could tolerate and relieve themselves, suggesting that video game interventions are safe. Dankbaar et al [76] showed that video lectures from a serious game are effective for specific topics, such as patient safety.

Limitations

This study had several limitations. First, we predesigned the inclusion criteria of video game intervention versus nonvideo game control. The definition of "video game" was broad, including Xbox 360, Nintendo Wii, virtual reality–based games, and computer-based games. However, the degree of stimulation, interaction, and pleasure of different video games for participants differed, resulting in clinical heterogeneity. Second, substantial heterogeneity was frequently determined on pairwise meta-analysis. One reason is the existence of clinical heterogeneity (many subgroups included less than 3 studies), and the other is the different types of combined scales for outcomes, leading to methodological heterogeneity. Third, some publication bias was detected in both pairwise and Bayesian network meta-analyses, which may be due to the difficulty in publishing negative outcomes. Lastly, to partially maintain the slight stability of the results under the Bayesian framework, we only chose outcomes that included more studies. Therefore, although the Bayesian network meta-analysis was not comprehensive, the results were more accurate.

Conclusion

In summary, impaired cognitive function is a highly prevalent condition that can profoundly influence the quality of life and accounts for major health care expenditures among the elderly. Our comprehensive Bayesian network meta-analysis provided evidence that video game interventions could be considered for the elderly to improve their performance and cognitive function, especially general cognitive scores and processing speed. Video games with better interactivity and visual stimulation have better curative effects. Based on the available evidence, we recommend video game interventions for the elderly. Future studies should be designed as multicenter RCTs, involving more subjects and providing more detailed description of the types of video games, in order to determine the most appropriate type of video game for older adults.

XSL•FO
RenderX

Authors' Contributions

CY was the principal investigator. CY, YZ (Yingshi Zhang), EJ, and CP were involved in study conception, design, data capture, data analysis, interpretation of the results, and editing of the manuscript and were the primary writers of the manuscript. CY, XH, MJ, and JX were involved in data capture, data analysis, and editing of the manuscript. CY, CX, YW, and YZ (Yajun Zhang) were involved in study conception, data analysis, and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Baseline characteristics of 47 studies.
[DOC File , 211 KB - games_v9i4e27058_app1.doc]

Multimedia Appendix 2
Risk-of-bias summary.
[PDF File (Adobe PDF File), 873 KB - games_v9i4e27058_app2.pdf]

Multimedia Appendix 3
Risk-of-bias assessment using the NOS for CCTs. CCT: case-controlled trial; NOS: Newcastle-Ottawa Scale.
[DOC File , 14 KB - games_v9i4e27058_app3.doc]

Multimedia Appendix 4
Net funnel of balance time.
[PDF File (Adobe PDF File), 453 KB - games_v9i4e27058_app4.pdf]

References

1. Ng NF, Schafer RJ, Simone CM, Osman AM. Perceptions of brain training: public expectations of cognitive benefits from popular activities. Front Hum Neurosci 2020;14:15 [FREE Full text] [doi: 10.3389/fnhum.2020.00015] [Medline: 32256323]
2. Mansor NS, Chow CM, Halaki M. Cognitive effects of video games in older adults and their moderators: a systematic review with meta-analysis and meta-regression. Aging Ment Health 2020 Jun;24(6):841-856. [doi: 10.1080/13607863.2019.1574710] [Medline: 30784296]
3. Gates NJ, Rutjes AW, Di Nisio M, Karim S, Chong L, March E, et al. Computerised cognitive training for maintaining cognitive function in cognitively healthy people in midlife. Cochrane Database Syst Rev 2019 Mar 13;3:CD012278 [FREE Full text] [doi: 10.1002/14651858.CD012278.pub2] [Medline: 30864746]
4. Amidfar M, de Oliveira J, Kucharska E, Budni J, Kim Y. The role of CREB and BDNF in neurobiology and treatment of Alzheimer's disease. Life Sci 2020 Sep 15;257:118020. [doi: 10.1016/j.lfs.2020.118020] [Medline: 32603820]
5. Ali AA, Abo El-Ella DM, El-Emam SZ, Shahat AS, El-Sayed RM. Life Sci 2019 Jul 15;229:21-35. [doi: 10.1016/j.lfs.2019.05.006] [Medline: 31063734]
6. 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]
7. Baranowski T, Buday R, Thompson DI, Baranowski J. Playing for real: video games and stories for health-related behavior change. Am J Prev Med 2008 Jan;34(1):74-82 [FREE Full text] [doi: 10.1016/j.amepre.2007.09.027] [Medline: 18083454]
8. Simmich J, Deacon AJ, Russell TG. Active video games for rehabilitation in respiratory conditions: systematic review and meta-analysis. JMIR Serious Games 2019 Feb 25;7(1):e10116 [FREE Full text] [doi: 10.2196/10116] [Medline: 30801256]
9. Morrison S, Simmons R, Colberg SR, Parson HK, Vinik AI. Supervised balance training and Wii Fit-based exercises lower falls risk in older adults with type 2 diabetes. J Am Med Dir Assoc 2018 Feb;19(2):185.e7-185.e13. [doi: 10.1016/j.jamda.2017.11.004] [Medline: 29289541]
10. Ryan C, Lewis J. Computer and Internet Use in the United States: 2015. 2017. URL: https://www.census.gov/library/publications/2017/acs/acs-37.html [accessed 2021-11-23]
11. Adcock M, Sonder F, Schättin A, Gennaro F, de Bruin ED. A usability study of a multicomponent video game-based training for older adults. Eur Rev Aging Phys Act 2020 Jan 11;17(1):3 [FREE Full text] [doi: 10.1186/s11556-019-0233-2] [Medline: 31938075]
12. Vázquez FL, Torres Á, Otero P, Blanco V, López L, García-Casal A, et al. Cognitive-behavioral intervention via interactive multimedia online video game for active aging: study protocol for a randomized controlled trial. Trials 2019 Dec 09;20(1):692 [FREE Full text] [doi: 10.1186/s13063-019-3859-5] [Medline: 31815656]

XSL•FO
RenderX

13. 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]
14. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med 2015 Jun 02;162(11):777-784 [FREE Full text] [doi: 10.7326/M14-2385] [Medline: 26030634]
15. Centre for Reviews and Dissemination. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care. URL: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf [accessed 2021-05-08]
16. Higgins JPT, Altman DG, Sterne JAC. Chapter 8: Assessing risk of bias in included studies. In: Higgins J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. London, UK: Cochrane; 2011:A.
17. Wells G. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. Symposium on Systematic Reviews: Beyond the Basics 2014:1-1.
18. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008 Apr 26;336(7650):924-926 [FREE Full text] [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]
19. Yang C, Gong G, Jin E, Han X, Zhuo Y, Yang S, et al. Topical application of honey in the management of chemo/radiotherapy-induced oral mucositis: a systematic review and network meta-analysis. Int J Nurs Stud 2019 Jan;89:80-87. [doi: 10.1016/j.ijnurstu.2018.08.007] [Medline: 30352321]
20. Zhang Y, Cai J, An L, Hui F, Ren T, Ma H, et al. Does music therapy enhance behavioral and cognitive function in elderly dementia patients? A systematic review and meta-analysis. Ageing Res Rev 2017 May;35:1-11. [doi: 10.1016/j.arr.2016.12.003] [Medline: 28025173]
21. Rica RL, Shimojo GL, Gomes MC, Alonso AC, Pitta RM, Santa-Rosa FA, et al. Effects of a Kinect-based physical training program on body composition, functional fitness and depression in institutionalized older adults. Geriatr Gerontol Int 2020 Mar;20(3):195-200. [doi: 10.1111/ggi.13857] [Medline: 31923924]
22. Amjad I, Toor H, Niazi IK, Pervaiz S, Jochumsen M, Shafique M, et al. Xbox 360 Kinect cognitive games improve slowness, complexity of EEG, and cognitive functions in subjects with mild cognitive impairment: a randomized control trial. Games Health J 2019 Apr;8(2):144-152. [doi: 10.1089/g4h.2018.0029] [Medline: 30239213]
23. Belchior P, Yam A, Thomas KR, Bavelier D, Ball KK, Mann WC, et al. Computer and videogame interventions for older adults' cognitive and everyday functioning. Games Health J 2019 Apr 29;8(2):129-143 [FREE Full text] [doi: 10.1089/g4h.2017.0092] [Medline: 30273002]
24. Faust ME, Multhaup KS, Ong MS, Demakis GJ, Balz KG. Exploring the specificity, synergy, and durability of auditory and visual computer gameplay transfer effects in healthy older adults. J Gerontol B Psychol Sci Soc Sci 2020 Jun 02;75(6):1170-1180 [FREE Full text] [doi: 10.1093/geronb/gbz096] [Medline: 31410475]
25. Gatica-Rojas V, Cartes-Velásquez R, Albornoz-Verdugo ME, Soto-Poblete A, Monteiro-Junior RS, Elgueta-Cancino E. Effects of a Nintendo Wii exercise program versus tai chi chuan on standing balance in older adults: a preliminary study. J Phys Ther Sci 2019 Jan;31(1):1-4 [FREE Full text] [doi: 10.1589/jpts.31.1] [Medline: 30774195]
26. Liao Y, Chen I, Lin Y, Chen Y, Hsu W. Effects of virtual reality-based physical and cognitive training on executive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial. Front Aging Neurosci 2019 Jul 16;11:162 [FREE Full text] [doi: 10.3389/fnagi.2019.00162] [Medline: 31379553]
27. Montero-Alía P, Miralles-Basseda R, López-Jiménez T, Muñoz-Ortiz L, Jiménez-González M, Prat-Rovira J, et al. Controlled trial of balance training using a video game console in community-dwelling older adults. Age Ageing 2019 Jul 01;48(4):506-512 [FREE Full text] [doi: 10.1093/ageing/afz047] [Medline: 31081504]
28. Perrot A, Maillot P, Hartley A. Cognitive training game versus action videogame: effects on cognitive functions in older adults. Games Health J 2019 Feb;8(1):35-40. [doi: 10.1089/g4h.2018.0010] [Medline: 30376364]
29. Sápi M, Domján A, Fehérné Kiss A, Pintér S. Is Kinect training superior to conventional balance training for healthy older adults to improve postural control? Games Health J 2019 Feb;8(1):41-48. [doi: 10.1089/g4h.2018.0027] [Medline: 30153062]
30. Sosa GW, Lagana L. The effects of video game training on the cognitive functioning of older adults: a community-based randomized controlled trial. Arch Gerontol Geriatr 2019 Jan;80:20-30. [doi: 10.1016/j.archger.2018.04.012] [Medline: 30316185]
31. Zadro JR, Shirley D, Simic M, Mousavi SJ, Ceprnja D, Maka K, et al. Video-game-based exercises for older people with chronic low back pain: a randomized controlled table trial (GAMEBACK). Phys Ther 2019 Jan 01;99(1):14-27. [doi: 10.1093/ptj/pzy112] [Medline: 30247715]
32. Bacha JMR, Gomes GCV, de Freitas TB, Viveiro LAP, da Silva KG, Bueno GC, et al. Effects of Kinect adventures games versus conventional physical therapy on postural control in elderly people: a randomized controlled trial. Games Health J 2018;7(1):24-36. [doi: 10.1080/17521740701702115]
33. Gomes GCV, Simões MDS, Lin SM, Bacha JMR, Viveiro LAP, Varise EM, et al. Feasibility, safety, acceptability, and functional outcomes of playing Nintendo Wii Fit Plus for frail older adults: A randomized feasibility clinical trial. Maturitas 2018 Dec;118:20-28. [doi: 10.1016/j.maturitas.2018.10.002] [Medline: 30415751]

XSL•FO
RenderX

34. Li J, Theng Y, Foo S, Xu X. Exergames vs. traditional exercise: investigating the influencing mechanism of platform effect on subthreshold depression among older adults. Aging Ment Health 2018 Dec;22(12):1634-1641. [doi: 10.1080/13607863.2017.1385722] [Medline: 28984486]
35. Ballesteros S, Mayas J, Prieto A, Ruiz-Marquez E, Toril P, Reales JM. Effects of video game training on measures of selective attention and working memory in older adults: results from a randomized controlled trial. Front Aging Neurosci 2017 Nov;9:354 [FREE Full text] [doi: 10.3389/fnagi.2017.00354] [Medline: 29163136]
36. Buitenweg JIV, van de Ven RM, Prinssen S, Murre JMJ, Ridderinkhof KR. Cognitive flexibility training: a large-scale multimodal adaptive active-control intervention study in healthy older adults. Front Hum Neurosci 2017;11:529 [FREE Full text] [doi: 10.3389/fnhum.2017.00529] [Medline: 29209183]
37. Lee Y, Choi W, Lee K, Song C, Lee S. Virtual reality training with three-dimensional video games improves postural balance and lower extremity strength in community-dwelling older adults. J Aging Phys Act 2017 Oct 01;25(4):621-627. [doi: 10.1123/japa.2015-0271] [Medline: 28290746]
38. Monteiro-Junior RS, da Silva Figueiredo LF, Maciel-Pinheiro PDT, Abud ELR, Braga AEMM, Barca ML, et al. Acute effects of exergames on cognitive function of institutionalized older persons: a single-blinded, randomized and controlled pilot study. Aging Clin Exp Res 2017 Jun;29(3):387-394. [doi: 10.1007/s40520-016-0595-5] [Medline: 27256080]
39. Ordnung M, Hoff M, Kaminski E, Villringer A, Ragert P. No overt effects of a 6-week exergame training on sensorimotor and cognitive function in older adults. A preliminary investigation. Front Hum Neurosci 2017;11:160 [FREE Full text] [doi: 10.3389/fnhum.2017.00160] [Medline: 28420973]
40. Souders DJ, Boot WR, Blocker K, Vitale T, Roque NA, Charness N. Evidence for narrow transfer after short-term cognitive training in older adults. Front Aging Neurosci 2017 Feb;9:41 [FREE Full text] [doi: 10.3389/fnagi.2017.00041] [Medline: 28293188]
41. Eggenberger P, Wolf M, Schumann M, de Bruin ED. Exergame and balance training modulate prefrontal brain activity during walking and enhance executive function in older adults. Front Aging Neurosci 2016 Apr 12;8:66 [FREE Full text] [doi: 10.3389/fnagi.2016.00066] [Medline: 27148041]
42. Kwok BC, Pua YH. Effects of WiiActive exercises on fear of falling and functional outcomes in community-dwelling older adults: a randomised control trial. Age Ageing 2016 Sep;45(5):621-627. [doi: 10.1093/ageing/afw108] [Medline: 27496921]
43. Nouchi R, Saito T, Nouchi H, Kawashima R. Small acute benefits of 4 weeks processing speed training games on processing speed and inhibition performance and depressive mood in the healthy elderly people: evidence from a randomized control trial. Front Aging Neurosci 2016;8:302 [FREE Full text] [doi: 10.3389/fnagi.2016.00302] [Medline: 28066229]
44. Schättin A, Arner R, Gennaro F, de Bruin ED. Adaptations of prefrontal brain activity, executive functions, and gait in healthy elderly following exergame and balance training: a randomized-controlled study. Front Aging Neurosci 2016;8:278 [FREE Full text] [doi: 10.3389/fnagi.2016.00278] [Medline: 27932975]
45. Toril P, Reales JM, Mayas J, Ballesteros S. Video game training enhances visuospatial working memory and episodic memory in older adults. Front Hum Neurosci 2016 May;10:206 [FREE Full text] [doi: 10.3389/fnhum.2016.00206] [Medline: 27199723]
46. Yeşilyaprak SS, Yıldırım M, Tomruk M, Ertekin Ö, Algun ZC. Comparison of the effects of virtual reality-based balance exercises and conventional exercises on balance and fall risk in older adults living in nursing homes in Turkey. Physiother Theory Pract 2016 Apr 06;32(3):191-201. [doi: 10.3109/09593985.2015.1138009] [Medline: 27049879]
47. Eggenberger P, Schumacher V, Angst M, Theill N, de Bruin ED. Does multicomponent physical exercise with simultaneous cognitive training boost cognitive performance in older adults? A 6-month randomized controlled trial with a 1-year follow-up. Clin Interv Aging 2015;10:1335-1349 [FREE Full text] [doi: 10.2147/CIA.S87732] [Medline: 26316729]
48. Gschwind YJ, Eichberg S, Ejupi A, de Rosario H, Kroll M, Marston HR, et al. ICT-based system to predict and prevent falls (iStoppFalls): results from an international multicenter randomized controlled trial. Eur Rev Aging Phys Act 2015;12:10 [FREE Full text] [doi: 10.1186/s11556-015-0155-6] [Medline: 26865874]
49. Karahan AY, Tok F, Taşkın H, Kuçuksaraç S, Başaran A, Yıldırım P. Effects of exergames on balance, functional mobility, and quality of life of geriatrics versus home exercise programme: randomized controlled study. Cent Eur J Public Health 2015 Nov;23 Suppl:S14-S18 [FREE Full text] [doi: 10.21101/cejph.a4081] [Medline: 26849537]
50. Kim K, Choi Y, You H, Na DL, Yoh M, Park J, et al. Effects of a serious game training on cognitive functions in older adults. J Am Geriatr Soc 2015 Mar 20;63(3):603-605. [doi: 10.1111/jgs.13304] [Medline: 25800915]
51. Sato K, Kuroki K, Saiki S, Nagatomi R. Improving walking, muscle strength, and balance in the elderly with an exergame using Kinect: a randomized controlled trial. Games Health J 2015 Jun;4(3):161-167. [doi: 10.1089/g4h.2014.0057] [Medline: 26182059]
52. Schoene D, Valenzuela T, Toson B, Delbaere K, Severino C, Garcia J, et al. Interactive cognitive-motor step training improves cognitive risk factors of falling in older adults: a randomized controlled trial. PLoS One 2015;10(12):e0145161 [FREE Full text] [doi: 10.1371/journal.pone.0145161] [Medline: 26673919]
53. Whyatt C, Merriman NA, Young WR, Newell FN, Craig C. A Wii bit of fun: a novel platform to deliver effective balance training to older adults. Games Health J 2015 Dec;4(6):423-433 [FREE Full text] [doi: 10.1089/g4h.2015.0006] [Medline: 26469308]

XSL•FO
RenderX

54. Bieryla K, Dold N. Feasibility of Wii Fit training to improve clinical measures of balance in older adults. CIA 2013 Jun:775. [doi: 10.2147/cia.s46164]
55. Jorgensen MG, Laessoe U, Hendriksen C, Nielsen OBF, Aagaard P. Efficacy of Nintendo Wii training on mechanical leg muscle function and postural balance in community-dwelling older adults: a randomized controlled trial. J Gerontol A Biol Sci Med Sci 2013 Jul;68(7):845-852. [doi: 10.1093/gerona/gls222] [Medline: 23114461]
56. Schoene D, Lord SR, Delbaere K, Severino C, Davies TA, Smith ST. A randomized controlled pilot study of home-based step training in older people using videogame technology. PLoS One 2013;8(3):e57734 [FREE Full text] [doi: 10.1371/journal.pone.0057734] [Medline: 23472104]
57. Singh DKA, Rajaratnam BS, Palaniswamy V, Raman VP, Bong PS, Pearson H. Effects of balance-focused interactive games compared to therapeutic balance classes for older women. Climacteric 2013 Feb 28;16(1):141-146. [doi: 10.3109/13697137.2012.664832] [Medline: 22640573]
58. Daniel K. Wii-hab for pre-frail older adults. Rehabil Nurs 2012;37(4):195-201. [doi: 10.3109/13697137.2012.664832] [Medline: 22640573]
59. Franco JR, Jacobs K, Inzerillo C, Kluzik J. The effect of the Nintendo Wii Fit and exercise in improving balance and quality of life in community dwelling elders. Technol Health Care 2012;20(2):95-115. [doi: 10.3233/THC-2011-0661] [Medline: 22508022]
60. Maillot P, Perrot A, Hartley A. Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. Psychol Aging 2012 Sep;27(3):589-600. [doi: 10.1037/a0026268] [Medline: 22122605]
61. Nouchi R, Taki Y, Takeuchi H, Hashizume H, Akitsuki Y, Shigemune Y, et al. Brain training game improves executive functions and processing speed in the elderly: a randomized controlled trial. PLoS One 2012 Jan;7(1):e29676 [FREE Full text] [doi: 10.1371/journal.pone.0029676] [Medline: 22253758]
62. Pluchino A, Lee SY, Asfour S, Roos BA, Signorile JF. Pilot study comparing changes in postural control after training using a video game balance board program and 2 standard activity-based balance intervention programs. Arch Phys Med Rehabil 2012 Jul;93(7):1138-1146. [doi: 10.1016/j.apmr.2012.01.023] [Medline: 22414490]
63. Rendon AA, Lohman EB, Thorpe D, Johnson EG, Medina E, Bradley B. The effect of virtual reality gaming on dynamic balance in older adults. Age Ageing 2012 Jul;41(4):549-552 [FREE Full text] [doi: 10.1093/ageing/afs053] [Medline: 22672915]
64. Singh DK, Rajaratnam BS, Palaniswamy V, Pearson H, Raman VP, Bong PS. Participating in a virtual reality balance exercise program can reduce risk and fear of falls. Maturitas 2012 Nov;73(3):239-243. [doi: 10.1016/j.maturitas.2012.07.011] [Medline: 22884437]
65. van Muijden J, Band GPH, Hommel B. Online games training aging brains: limited transfer to cognitive control functions. Front Hum Neurosci 2012;6:221 [FREE Full text] [doi: 10.3389/fnhum.2012.00221] [Medline: 22912609]
66. Szturm T, Betker AL, Moussavi Z, Desai A, Goodman V. Effects of an interactive computer game exercise regimen on balance impairment in frail community-dwelling older adults: a randomized controlled trial. Phys Ther 2011 Oct;91(10):1449-1462. [doi: 10.2522/ptj.20090205] [Medline: 21799138]
67. Ozdogar AT, Ertekin O, Kahraman T, Yigit P, Ozakbas S. Effect of video-based exergaming on arm and cognitive function in persons with multiple sclerosis: a randomized controlled trial. Mult Scler Relat Disord 2020 May 1;40:101966. [doi: 10.1016/j.msard.2020.101966] [Medline: 32045868]
68. Shin KS, Yim YK, Kim Y, Park S, Lee J. Effects of visual game experience on auditory processing speed. Cyberpsychol Behav Soc Netw 2017 Mar;20(3):187-194. [doi: 10.1089/cyber.2016.0171] [Medline: 28263679]
69. Wang G, Zhao M, Yang F, Cheng LJ, Lau Y. Game-based brain training for improving cognitive function in community-dwelling older adults: a systematic review and meta-regression. Arch Gerontol Geriatr 2021;92:104260 [FREE Full text] [doi: 10.1016/j.archger.2020.104260] [Medline: 32980574]
70. Vázquez FL, Otero P, García-Casal JA, Blanco V, Torres Á, Arrojo M. Efficacy of video game-based interventions for active aging. A systematic literature review and meta-analysis. PLoS One 2018;13(12):e0208192 [FREE Full text] [doi: 10.1371/journal.pone.0208192] [Medline: 30533015]
71. Carregosa AA, Aguiar Dos Santos LR, Masruha MR, Coêlho MLDS, Machado TC, Souza DCB, et al. Virtual rehabilitation through Nintendo Wii in poststroke patients: follow-up. J Stroke Cerebrovasc Dis 2018 Feb;27(2):494-498. [doi: 10.1016/j.jstrokecerebrovasdis.2017.09.029] [Medline: 29100855]
72. Li J, Theng Y, Foo S. Exergames for older adults with subthreshold depression: does higher playfulness lead to better improvement in depression? Games Health J 2016 Jun;5(3):175-182. [doi: 10.1089/g4h.2015.0100] [Medline: 27135146]
73. Li Y, Zhang L, Long K, Gong H, Lei H. Real-time monitoring prefrontal activities during online video game playing by functional near-infrared spectroscopy. J Biophotonics 2018 Sep;11(9):e201700308. [doi: 10.1002/jbio.201700308] [Medline: 29451742]
74. Green CS, 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]
75. Lorenz RC, Gleich T, Gallinat J, Kühn S. Video game training and the reward system. Front Hum Neurosci 2015;9:40 [FREE Full text] [doi: 10.3389/fnhum.2015.00040] [Medline: 25698962]

XSL•FO
RenderX

76. Dankbaar MEW, Richters O, Kalkman CJ, Prins G, Ten Cate OTJ, van Merrienboer JJG, et al. Comparative effectiveness of a serious game and an e-module to support patient safety knowledge and awareness. BMC Med Educ 2017 Feb 02;17(1):30 [FREE Full text] [doi: 10.1186/s12909-016-0836-5] [Medline: 28148296]

Abbreviations

BBS: Berg Balance Scale
BMI: body mass index
CCT: case-controlled trial
CrI: credible interval
GRADE: Grading of Recommendations Assessment, Development and Evaluation
MMSE: Mini-Mental State Exam
MoCA: Montreal Cognitive Assessment
NOS: Newcastle-Ottawa Scale
OR: odds ratio
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
RCT: randomized controlled trial
SMD: standardized mean difference
SUCRA: surface under the cumulative ranking curve
TMT: trail-making test

Edited by N Zary; submitted 09.01.21; peer-reviewed by P Haring, R Mandryk, AS Lu; comments to author 19.03.21; revised version received 22.03.21; accepted 12.10.21; published 30.12.21.

Please cite as:
Yang C, Han X, Jin M, Xu J, Wang Y, Zhang Y, Xu C, Zhang Y, Jin E, Piao C
The Effect of Video Game–Based Interventions on Performance and Cognitive Function in Older Adults: Bayesian Network Meta-analysis
JMIR Serious Games 2021;9(4):e27058
URL: https://games.jmir.org/2021/4/e27058
doi:10.2196/27058
PMID:34967759

©Chao Yang, Xiaolei Han, Mingxue Jin, Jianhui Xu, Yiren Wang, Yajun Zhang, Chonglong Xu, Yingshi Zhang, Enshi Jin, Chengzhe Piao. Originally published in JMIR Serious Games (https://games.jmir.org), 30.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Gamifying Sexual Education for Adolescents in a Low-Tech Setting: Quasi-Experimental Design Study

Hussein Haruna[1,2*], PhD; Kingsley Okoye[1*], PhD; Zamzami Zainuddin[2], PhD; Xiao Hu[2], PhD; Samuel Chu[2*], PhD; Samira Hosseini[1,3], PhD

[1]Writing Lab, Institute for the Future of Education, Tecnologico de Monterrey, Monterrey, Mexico
[2]Faculty of Education, The University of Hong Kong, Hong Kong, Hong Kong
[3]School of Engineering and Sciences, Tecnologico de Monterrey, Monterrey, Nuevo Leon, Mexico
[*]these authors contributed equally

Corresponding Author:
Hussein Haruna, PhD
Writing Lab, Institute for the Future of Education
Tecnologico de Monterrey
Avenida Eugenio Garza Sada 2501 Sur
Monterrey, 64849
Mexico
Phone: 52 18134162370
Email: harunahussein@gmail.com

Abstract

Background: Sexual education has become increasingly important as unhealthy sexual practices and subsequent health risks become more prevalent during adolescence. Traditional sex education teaching methodologies are limiting for digital natives exposed to various digital technologies. Harnessing the power of technology applications attractive to the younger generation may be a useful approach for teaching sex education.

Objective: The aim of this study was to improve sexual health knowledge and understanding of the problems associated with unhealthy sexual practices and address sexual and reproductive health challenges experienced in a low-tech setting.

Methods: A participatory design approach was used to develop the digital gamified methodology. A sample of 120 secondary school students aged 11-15 were randomly assigned to either experimental or control group for each of the 3 teaching approaches: (1) gamified instruction (actual serious games [SG] in teaching); (2) gamification (GM; making nongames, such as game-like learning); and (3) traditional teaching (TT) methods.

Results: The SG and GM approaches were more effective than TT methods in teaching sexual health education. Specifically, the average scores across groups demonstrated an increase of mean scores from the pre- to posttest (25.10 [SD 5.50] versus 75.86 [SD 13.16]; t_{119}=41.252; P<.001 [2 tailed]). Analysis of variance indicated no significant differences across groups for pretest scores ($F_{2,117}$=1.048, P=.35). Significant differences across groups were evident in the posttest scores. Students in the SG and GM groups had higher average scores than the TT group ($F_{2,117}$=83.98; P<.001). Students reported increased learning motivation, attitude, know-how, and participation in learning (P<.001) when using SG and GM approaches.

Conclusions: Digital health technologies (particularly teaching and learning through gamified instruction and other novel approaches) may improve sexual health education. These findings may also be applied by practitioners in health care settings and by researchers wishing to further the development of sex education.

(JMIR Serious Games 2021;9(4):e19614) doi:10.2196/19614

KEYWORDS

gamified instruction; serious gaming; gamification; educational innovation; teenage students; digital generation; e-learning; low-tech setting

XSL•FO
RenderX

Introduction

Background

Unhealthy sexual activity and its related diseases have increased globally. Nowhere is the effect of sexually transmitted diseases more apparent than in the countries of sub-Saharan Africa (SSA). In SSA countries, many adolescents are exposed early to sexual intercourse and sexual and emotional abuse [1]. Consequently, adolescents are vulnerable to unsafe sexual intercourse practices, sexual encounters with many partners, forced sexual contact, exploitive sexual activity and relationships, and influence from sexually active friends [2-5]. Exposure to these types of sexual practice has led to an increase in sexually transmitted infections (STIs), such as HIV/AIDS and Chlamydia [6-8], and other consequences. A plethora of research studies have supported the increase of sexual health literacy as a way to reduce the spate of unhealthy sexual practices and curb the current increase in sexually transmitted diseases [3,9-11]. A variety of sexual health education programs for adolescents have been implemented globally. However, the efficacy of pedagogy plays a crucial role in fostering sexual health knowledge acquisition. An effective pedagogy supports a host of academic achievement paradigms [7,8,12-15]. Although effective pedagogy has been given less emphasis in sexual health education than in core curriculum subjects [16], initiatives are being undertaken for more effective sexual health education in the digital era.

Digital health games designed to target sexual health practices have increasingly demonstrated their capabilities, appeal, and influence on educating digital native adolescents [17]. Gamified learning (serious games [SG] and gamification [GM]) platforms provide unique methods for delivering educational objectives, increasing knowledge, and reducing sex-related problems faced by adolescents [18-20]. The capacity of outreach for digital games is higher than that for traditional teaching (TT) methods [21]. Approximately 97% of adolescents normally engage in digital games, whereas 50% spend more than 1 hour per day on one kind of gaming equipment or platform or another. The Speak Up Project for Digital Learning revealed a higher preference for digital gamified learning platforms for instruction over traditional ways of learning [22]. When considering the sensitive nature of sexual health knowledge dissemination, digital games are attractive because they offer a discreet, interactive, and confidential environment for learning. This makes a difference for conservative societies [3,23].

Digital games facilitate role playing and offer challenging approaches to learning improving attitude and decision-making skills applicable to real-life scenarios. Because digital platforms offer an engaging approach for learners, they promote knowledge acquisition [8]. The novelty of this study is highlighted by the exploration into sexual health education in SSA countries, which have limited technology use. Only one study has investigated attitude changes and sexual health knowledge acquisition in a country with a similar low-tech environment [24]. Appeals for data supporting the use of technology to disseminate sexual health knowledge in low-tech settings have been documented [25]. This study employs a participatory research approach. It does this to design 2 digital health interventions (SG and GM) that assess 4 aspects (motivation, attitude, knowledge, and engagement [MAKE]) among adolescents in SSA countries.

Objective of This Study

The aims of this study are to (1) add to the limited existing knowledge of game-based technologies and (2) address the interest in using this novel kind of technology as the teaching approach in a low-tech setting in Africa. First, we hypothesized that the application of game elements and mechanics in learning would enhance the sexual health literacy of teenage students. Second, we hypothesized that the teenage students would develop an attitude toward gamified instruction (SG and GM) that was more favorable and receptive than that toward the traditional learning approaches. Henceforth, this study looks at how gamified instruction can improve the sexual health education of adolescents, address their sexual health challenges, and help them overcome those challenges, all of this in developing countries, which tend to be far less invested in digital technologies than developed countries [26].

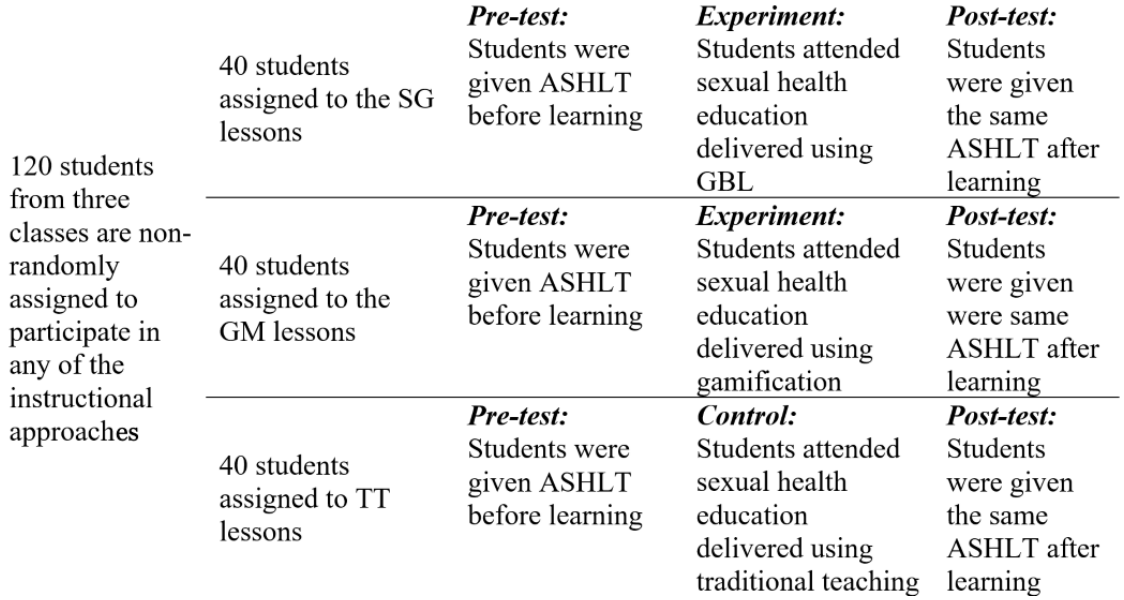
Methods

Study Design

This study employed a quasi-experimental research design. The design guided the sexual health literacy interventions for students clustered in 3 classes. Sexual health education was made mandatory for all students to sanction the randomization technique [26,27]. The study was also in line with previous empirical studies and other publications that guide quasi-experimental research design [28-30]. It evaluated learning outcomes using pre- and posttest evaluations across the 3 teaching approaches (SG, GM, and TT). The students' perceptions were compared to determine which instructional approach was the most effective in motivating students to learn, change attitudes, acquire knowledge, and to become engaged in the courses. Figure 1 presents the quasi-experimental research design employed in this study.

XSL•FO
RenderX

Figure 1. Quasi-experimental research design employed in this study. Students in their existing 3 classes were nonrandomly selected to participate in one of the 3 instructional approaches (SG, GM, and TT). ASHLT: Adolescent Sexual Health Literacy Test; GM: gamification; SG: serious game; TT: traditional teaching.



Participants

The study involved teenage students (n=120) aged 11-15 who were enrolled in a secondary school at the time of the study. The research was carried out in a school in Dar es Salaam, Tanzania. This school was selected because it had 2 computer laboratories, each able to accommodate around 40 students. They had internet connectivity, a power supply, and a standby generator. The 3 classes had around 40 students each. Everyone in a class was in the same grade. The participants were not chosen randomly. Each participant was assigned to either an experimental group or a control group based on their intact classroom settings. The researchers had no authority to form or annihilate the existing study population setting. The research team randomly designated 2 of the classes as experimental (digital game and GM) and 1 class as control (traditional).

Each group was unaware of the other groups. The learning materials were the same for all the 3 groups. The only differences were in the instructional approaches. The participants were from different schools that had taken part in the revision and participatory design process of the interventions. Finally, there were 5 topics, each covered in one 40-minute class per week, as reported in Table 1.

Table 1. Sexual health education lessons covered per week and their time length (n=120). The columns and individual cells represent records per lesson, per week, and per class.

Topics	Week	Duration (minutes)	Serious game	Gamification	Traditional teaching
1. Personal hygiene and good manners	1	40	40	40	40
2. Sexual responsibility and decision making	2	40	40	40	40
3. Dealing with peer pressure during adolescence	3	40	40	40	40
4. Prevention of sexually transmitted infections, including HIV/AIDS	4	40	40	40	40
5. Dealing with harmful practices and sexual violence	5	40	40	40	40

Study Conditions

Interventions

SG and GM interventions were developed following “activity theory” [31], “design-based research” [32], and participatory design approach [15,33]. All of these emphasize the involvement of stakeholders in developing instructional interventions for addressing the intended needs of the study population. As this was the third round of intervention testing, the games were refined based on outcomes from the second round. Students from this group shared their comments for making further improvement. The participatory research design approach employed led to the refinement of the 2 gamified interventions with a view to addressing the challenges faced by the adolescents [18,25]. While the revisions were carried out, the intended users of the systems/learning platforms and other stakeholders (eg, pediatricians; sexual and reproductive health specialists; sexual health teachers from participating schools; computer and information science specialists, including a game designer who was a computer engineer; and targeted secondary school student end users) were all involved in the study. These stakeholders were invited to participate in a series of design workshops during the refining of the intervention. This study also reports the research conducted during the third iteration. Further details of the SG and GM design and development have

XSL•FO
RenderX

been published in another research [26]. The descriptions of each study condition are presented below.

Traditional Teaching Class

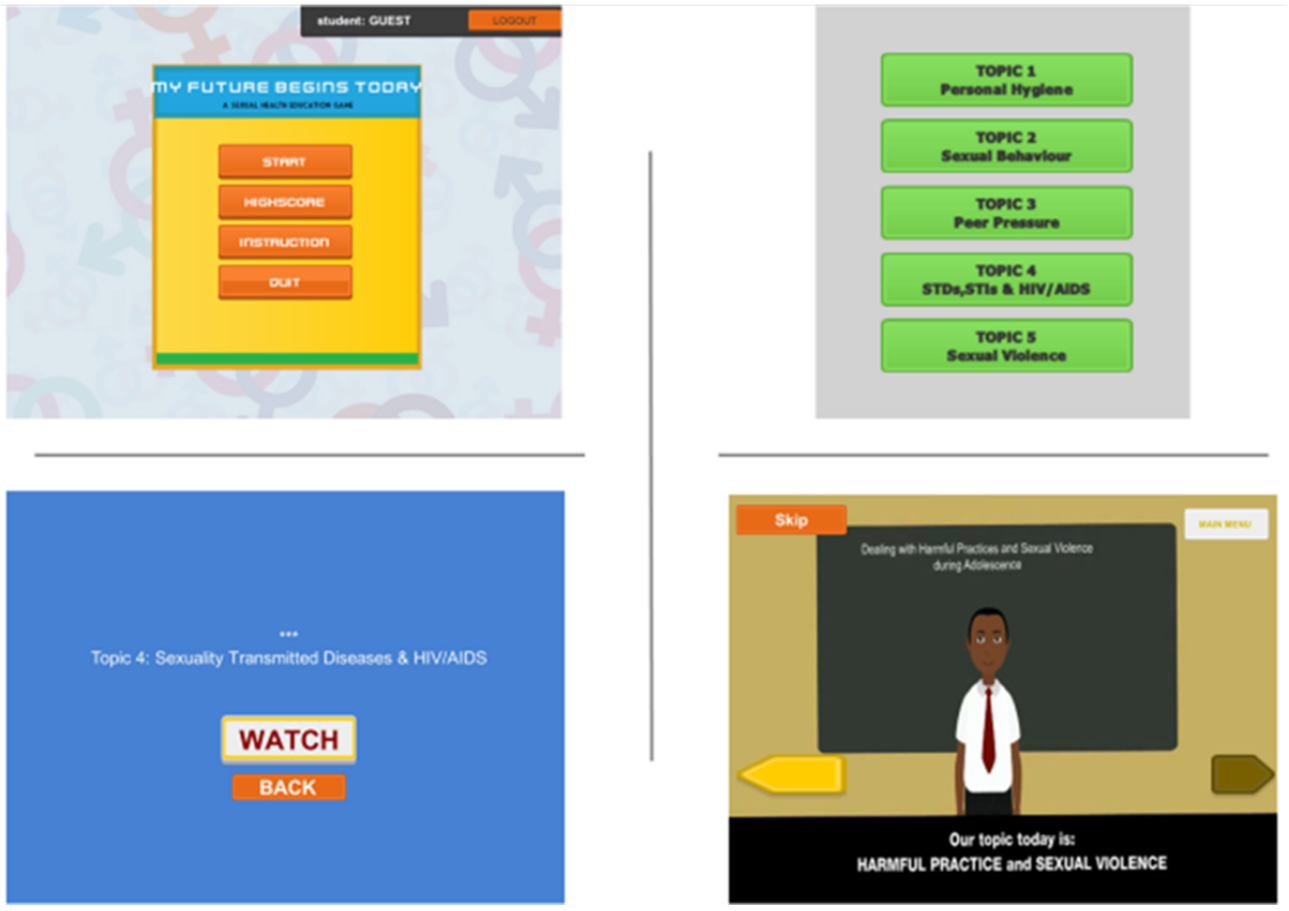
Students assigned to receive TT were treated as the control group. They were taught in a conventional classroom manner. Their teacher taught 1 day a week for 40 minutes for 5 weeks. Students were given hand-outs for further reading after each session. No digital technology was used.

Serious Gaming Teaching Class

Students in this group received sexual health education using an SG approach (Figures 2 and 3). A week before the classes started, SG students were oriented on the game in the school computer laboratory. Students played the "My Future Begins Today" game individually under the watchful eye of a teacher and the researchers after the classes have commenced. There were 5 topics arranged in chronological order. Each topic took about 40 minutes per week. The students were also allowed to use the game during free time.

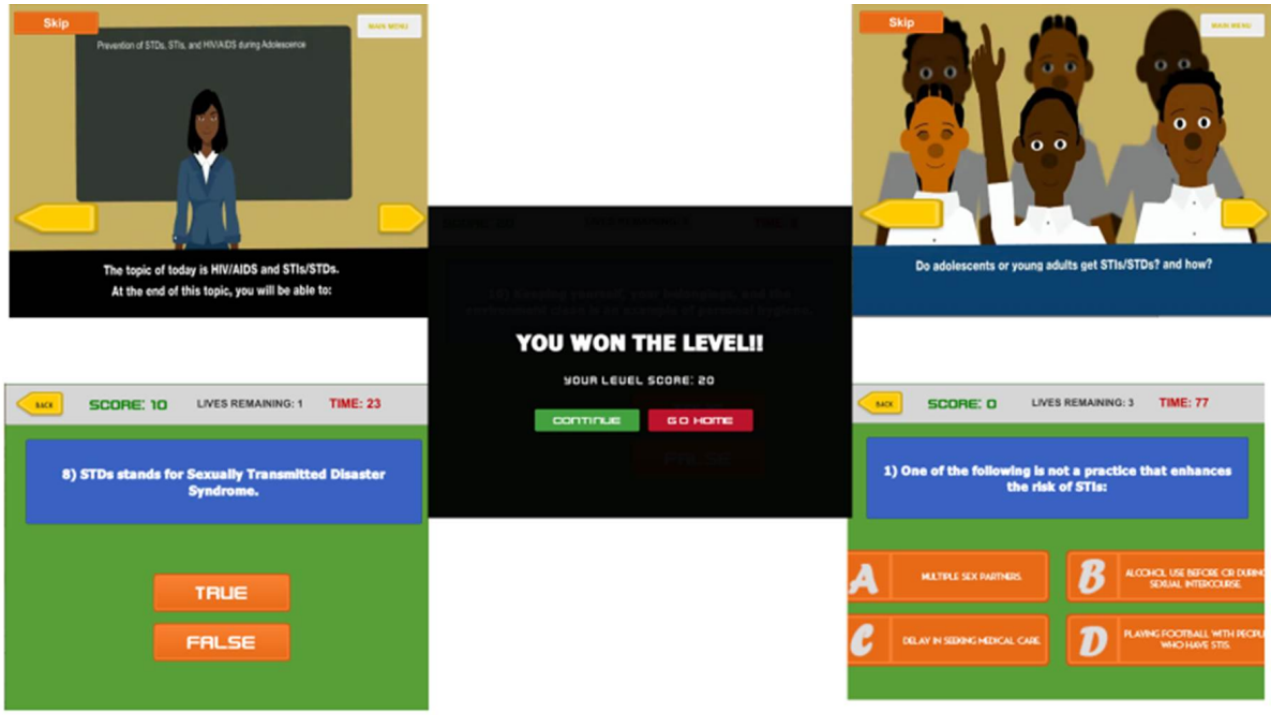
The game has an introduction which presents the general learning objective. Each topic has a game scenario in which there are avatars representing a teacher and students interacting in a classroom. There were conversations between the teacher and the students' avatars. After watching the scenarios, the students were asked to complete the quizzes online. There were 10 quizzes on each topic, to be completed within 90 minutes. Scores were provided for correct answers, and students would lose points for each wrong answer. The students also had an opportunity to repeat the gameplay within the 40-minute margin.

Figure 2. Representation of the third game platform and implementation.



XSL•FO
RenderX

Figure 3. The game structures.



Gamification Teaching Class

GM is the process of giving some of the characteristics of real games to activities that are not games. GM aims to make the learning activities more interactive. This is supposed to motivate students to learn in a way that is more effective because it is fun. GM is an emerging technique within education [3,34,35]. The concept is also a recent development in low-tech settings, especially in SSA countries. With GM, more actual learning tends to take place [36]. There are various types of learning management system platforms with built-in game mechanics [37,38]. This study used "Moodle" to organize and integrate the material we wanted to teach with game elements, such as badges, levels, leader boards, points, scores, competition, and quizzes (Figure 4). As with the SG students, the students who participated in the GM instruction received a 1-week orientation before the classes began.

Figure 4. Types of badges used in the gamification group.

XSL•FO
RenderX

There were 5 topics and for each topic 40 minutes were allocated per week for 5 consecutive weeks. Each student studied individually, but they could interact online via a discussion forum devoted to each topic. Students were asked to read and practice the lesson materials provided online. There were 10 questions under each section, some true/false and some multiple choice. There were 7 types of badges (Figure 4). One was automatically awarded upon completion of a lesson using the award rules as outlined in Table 2. Overall, the GM concept was used to make learning more fun, motivate the students to learn, support a change of attitude, and increase engagement. This study observed that learning in a competitive spirit increased the desire for continuous learning [39,40]. Students were also automatically positioned on the leaderboards and assigned levels based on the points they had gained after completing the learning activities.

Table 2. Criteria and rules to receive badges.

Badge name	Description
1st Level Badge	Rewarded to students who complete the first topic and are moving to the second topic
2nd Level Badge	Rewarded to students who complete the second topic and are moving to the third topic
3rd Level Badge	Rewarded to students who complete the third topic and are moving to the fourth topic
4th Level Badge	Rewarded to students who complete the fourth topic and are moving to the fifth topic
5th Level Badge	Rewarded to students who complete the fifth topic
Exclusive Top Badge	Rewarded to a student with the highest score for a particular topic
Outstanding Achievement Badge	Rewarded to a student with the highest total of points from all topics

Procedure and Data Collection Methods

Sexual Health Literacy Tests

During the design of the interventions the selected research team and participating parties, especially the teachers and sexual/reproductive health specialists, were involved in developing a set of questions covering the 5 topics taught. This test was titled Adolescent Sexual Health Literacy Test (ASHLT). There were 50 questions, 10 per topic in the following format: Section A (multiple choice), Section B (true/false), and Section C (short answer). The ASHLT took up to 45 minutes to complete. Before initiating the actual learning, students were asked to do a pretest (using the ASHLT) to assess their sexual health knowledge (baseline). Within a week following the training, the same ASHLT questions were given to measure their level of understanding.

Students' Perceptions of Teaching Approaches

This study used the MAKE framework [41], according to which a teaching method is regarded as effective if it shows the ability to *m*otivate students, improve their *a*ttitude, increase their *k*nowledge acquisition, and increase their *e*ngagement in learning. Several scholars have employed the method for evaluating the efficacy of the 4 components of MAKE independently [8,15,24,42-44]. This study employed a MAKE evaluation framework to evaluate and compare the efficacy of the 3 instructions by taking into account the 4 different perspectives (motivation, attitude, knowledge, and engagement) at once. The resultant MAKE instrument we used has 46 items, with the motivation construct containing 16 items and the other 3 constructs (attitude, knowledge, and engagement) having 10 each. We measured the students' viewpoints through a self-rating method that had a 5-point Likert scale (5=strongly agree to 1=strongly disagree). The ratings took 10 minutes to complete and were all conducted within a week.

Focus Group Interviews

We conducted focus group interviews (FGIs) to yield more comments on the teaching methods [45]. The FGIs were conducted to corroborate and complement the quantitative data. A total of 21 students were requested to participate in the FGI, 7 students for each of the 3 learning instructions. These are realistic numbers for an FGI [46]. There were 3 focus group discussions, one for each of the instruction categories. A semistructured interview guide/protocol was adopted from the MAKE evaluation framework. Students were asked to share their views on the effectiveness and other aspects of their learning approach, and an audio record of the FGI data was made. Verbatim transcriptions were made using pseudonyms for data analysis. The participants were given equal time (1 hour) to provide their comments.

Quantitative Data Analysis

Overview

The collected data were imported from the Excel (Microsoft) file format to the IBM SPSS software tool for statistics (version 24) to perform the quantitative analytic tests. This was for data generated using ASHLT and the MAKE evaluation framework. A paired *t* test was conducted to compare the pre- and posttest average scores. This was done to determine whether there were changes in the learning scores after a series of sexual health literacy sessions. Besides, a one-way analysis of variance (ANOVA) was performed to analyze the numerical data collected from the pre- and posttest scores. This test compared the variations across the 3 learning approaches. In other words, we performed the pretest comparison across the 3 instructions to establish possible significant differences before the training. This would especially rule out any possible bias in the sexual health knowledge of students collected at baseline. The descriptive statistics was focused on determining the mean, median, and SD on the self-rating scale of the measurement using the MAKE evaluation framework pertaining to the students' perceptions of the 3 instruction approaches. The

XSL•FO
RenderX

self-ratings of the effectiveness of the 3 teaching methods using the MAKE evaluation framework for each component were tested using the one-sample Kolmogorov–Smirnov and Shapiro–Wilk tests for normality. Table 3 presents the results of the normality tests.

Although the results indicated that the data samples violated the assumption of normality ($P<.05$), as the scores are non-normally distributed in the Kolmogorov–Smirnov and Shapiro–Wilk tests (Table 3), a nonparametric Kruskal–Wallis test was consequently used to compare and contrast the responses across the 3 groups for each component. A significant value of $P<.05$ was used to determine the results of the statistical analysis.

Table 3. Normality test results for the MAKE instrument.

Construct	Kolmogorov–Smirnov			Shapiro–Wilk		
	Statistic	*df*	*P* value[a,b]	Statistic	*df*	*P* value[a,b]
Motivation	.085	120	.03	.956	120	<.001
Attitude	.213	120	<.001	.800	120	<.001
Knowledge	.174	120	<.001	.896	120	<.001
Engagement	.103	120	.003	.961	120	.002

[a]All *P* values are <.05, and thus significant.

[b]Lilliefors significance correction.

Measurement Reliability

The validated instruments and reported questionnaires appeared to be satisfactory [41,47] following the factor analysis and reliability checks we conducted and documented for 120 samples. The sample size met the minimum of 100 or larger, or a ten-to-one ratio of observations per domain [48]. The motivation questionnaires showed a Cronbach α of .92. The attitude questionnaire showed a Cronbach α of .90, and the knowledge survey developed from the sexual health education syllabus showed a Cronbach α of .92. The engagement questionnaires developed from many sources with no existing reliability results indicated Cronbach α of .90. The results of the different scale reliability checks are presented in Table 4.

Table 4. Scale reliability for the MAKE evaluation instrument (N=120).

Constructs and components	Number of items	Cronbach α	Standardized α	Kaiser–Meyer–Olkin	*P* value
Motivation			.92	.88	<.001[a]
Attitude	4	.92			
Relevance	4	.93			
Confidence	4	.90			
Satisfaction	4	.85			
Attitude			.90	.88	<.001[a]
Affective	5	.91			
Cognitive	5	.89			
Knowledge			.92	.86	<.001[a]
Importance	4	.93			
Effectiveness	3	.93			
Application	3	.89			
Engagement			.90	.87	<.001[a]
Emotional	6	.91			
Cognitive	4	.88			

[a]The mean difference is significant if *P* value is <.05.

Qualitative Data Reliability and Analysis

The qualitative data collection instrument was developed using the MAKE evaluation. FGI transcriptions and records complemented the quantitative results. Membership checking, conformability, and validation were applied to the collected data to come up with critical comments on the sufficiency of the results for ensuring the reliability of the qualitative data. At the end of the data collection process, students were asked to review the transcripts to determine whether the transcripts

XSL•FO
RenderX

presented incorporated their comments. Thus, based on the MAKE evaluation instrument, 4 themes were developed (ie, motivation, attitude, knowledge, and engagement) to enable an ample analysis of the collected data. Then, a codebook was created using the 4 MAKE constructs. The students' transcripts were merged with the quantitative data (based on the 4 MAKE themes).

Results

Baseline Characteristics of Participants

In all, 120 teenage students participated in testing the interventions. Table 5 presents their demographic features and socioeconomic status, including their access to and use of digital technologies.

Table 5. Descriptive characteristics of participants (N=120).

Characteristics	Value
Sex, n (%)	
Male	69 (57.5)
Female	51 (42.5)
Age, mean (SD)	
Male	13.65 (0.99)
Female	13.65 (1.01)
Living group, n (%)	
With both parents	89 (74.2)
With father only	7 (5.8)
With mother only	16 (13.3)
With guardian only	8 (6.7)
Economic group, n (%)	
High class	14 (11.7)
Middle high class	47 (39.2)
Middle low class	57 (47.5)
Poor	2 (1.7)
Access to a computer at school or home, n (%)	
Yes	118 (98.3)
No	2 (1.7)
Access to smart devices at school or home, n (%)	
Yes	119 (99.2)
No	1 (0.8)
Play of computer or mobile phone games, n (%)	
Yes	117 (97.5)
No	3 (2.5)

ASHLT Test Results

Main Findings

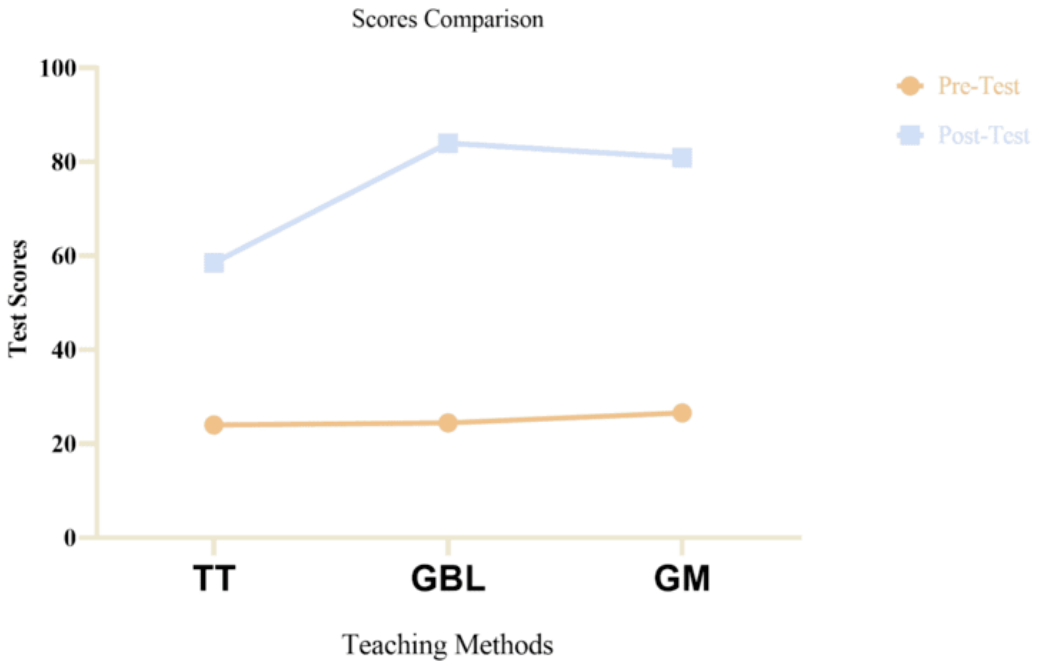
The study carried out a paired sample *t* test to assess the mean differences or effect of the 3 teaching methods based on the students' average scores in the ASHLT. Statistically, there emerged a significant improvement in the knowledge acquisition, as the data demonstrated an increase in mean scores in the ASHLT from pretest mean of 25.10 (SD = 5.50) to posttest mean of 75.86 (SD 13.16; t_{119}=41.252, P<.001; 2-tailed). A one-way ANOVA was then used to compare and contrast pre- and posttest across the 3 instructions. The average of pretest scores indicate that participants were equally distributed in all the 3 teaching methods: $F_{2,117}$=1.048, P=.35. The average posttest scores stemming from the experimental instructions (SG and GM) also indicated an increase—as opposed to their counterparts in the control group (TT): $F_{2,117}$=83.98, P<.001. Figure 5 presents a comparison of the effectiveness of the 3 teaching groups.

Likewise, we conducted post hoc tests (which served as follow-up analysis) to establish the differences in the 3 pairs of teaching groups: 2 experimental and 1 control. Significant divergences emerged for both the traditional and game-based groups (P<.001) and between the TT and GM groups (P<.001). However, there was no significant difference between SG and GM groups (P=.19). These results suggest that students assigned

XSL•FO
RenderX

to the experimental groups achieved a higher score after the lessons than the students in the control group.

Figure 5. Average comparison of three teaching groups.



Comparison of Average Scores Before the Series of Lessons

Table 6 presents the results for all the 5 topics. For example, the descriptive data we provide below demonstrate the average scores for the "personal hygiene and good manner during adolescence" topic as follows:

- TT group mean score of 5.45 (SD 1.73)
- SG group mean score of 5.36 (SD 2.15)
- GM mean score of 5.68 (SD 2.22)

As gathered in Table 6, a one-way ANOVA was conducted to compare each sexual health education topic taught across the 3 instructions (ie, TT, SG, and GM). The results indicated a nonstatistically significant difference in any of the 5 topics detected with $P>.05$. The results suggest that all the 3 teaching methods can be considered similar regarding sexual health knowledge in all 5 topics before the training.

Table 6. Comparison of average score and one-way ANOVA results before the series of lessons.

Topic	Pretest, mean (SD)			One-way ANOVA
	TT	SG	GM	
1	5.45 (1.73)	5.36 (2.15)	5.68 (2.22)	$F_{2,117}$=0.26, P=.76
2	4.92 (1.59)	4.92 (1.45)	5.31 (1.44)	$F_{2,117}$=0.89, P=.41
3	4.85 (1.18)	4.87 (1.65)	5.40 (1.60)	$F_{2,117}$=1.72, P=.18
4	4.66 (1.16)	4.67 (1.57)	4.71 (1.15)	$F_{2,117}$=0.01, P=.98
5	4.76 (1.52)	4.68 (1.53)	5.01 (1.36)	$F_{2,117}$=0.53, P=.58

Comparison of Average Scores After the Series of Lessons

The students in each group also completed the same (ASHLT) quizzes as a posttest. The one-way ANOVAs were performed to evaluate the students' average test scores across the 3 teaching methods. The descriptive data from the one-way ANOVAs and post hoc comparison tests are illustrated in Multimedia Appendix 1. The results indicated that the SG and GM groups had higher averages than the TT group.

The one-way ANOVAs revealed that the students received effective sexual health knowledge and that the acquisition rate increased for each topic in the 3 learning groups. Students from the SG and GM instruction groups showed a significant knowledge acquisition in topic 1, compared with students in the TT group ($F_{2,117}$=19.04, $P<.001$). Constant effects remained for the other 4 topics. The Tukey HSD post hoc multiple comparison tests indicated that the average scores for all the 5 topics significantly varied between the control and experimental groups ($P<.001$). The experimental groups did not differ significantly as P values were over .05 (refer to Multimedia Appendix 1 for details).

XSL•FO
RenderX

Students' Perceptions Toward Instruction Approaches

The study evaluated the students' perceptions of the 3 instructional approaches using a self-rating scale and FGI. The details of both the quantitative and qualitative results are presented below.

Quantitative Component

The averages for the responses to each aspect of the MAKE evaluation framework were compared for the 3 groups. A Kruskal–Wallis test was performed to determine the existence or nonexistence of statistically significant difference among the 3 groups after rating the average scores from each MAKE evaluation framework. Statistically significant dissimilarities between the 3 groups' averages were demonstrated (Multimedia Appendix 2). For instance, the Kruskal–Wallis test showed that there was a significant difference in motivation between the 3 instructions: SG mean of 4.51 (SD 0.25), GM mean of 4.40 (SD 0.38), and TT mean of 4.12 (SD 0.59); $P<.001$. Post hoc tests were also conducted to make pairwise comparisons. In the post hoc tests, we found that TT differed significantly from both GM ($P=.04$) and SG ($P<.001$). By contrast, GM and SG were not significantly different from each other ($P=.79$). Moreover, this effect remained consistent with the other aspects of the MAKE evaluation.

Qualitative Component

Comments were received from both the experimental and the control groups on the 3 instruction methods. Like the results from the quantitative data, students from the experimental groups commented favorably on the SG and GM instructions, whereas those in the control group commented unfavorably on TT. For example, for *motivation*, the students reported:

> *the games were fun* [SG-3]
>
> *...that learning was easy* [SG-5]
>
> *...the learning offered a self-regulatory method that improved my confidence* [GM-1]
>
> *...learning was [done] in a competitive [way], which helped me gain problem solving-skills* [GM-7]

Students also pointed out that

> *...the learning inspired me; hence I focused on learning* [SG-2]
>
> *...I was extremely interested in the learning approach* [GM-4]
>
> *...the availability of badges encouraged [me to learn] the subject* [GM-2]

By contrast, students from the TT group were largely negative about their learning experience:

> *...there were no visuals* [TT-5]
>
> *...[there was] no clarification on many issues* [TT-2]
>
> *...[there were] limited, or no activities for concentration* [TT-6]
>
> *...little or none of the critical thinking strategies were provided, including role play, demos, quizzes, team-work activities, [or] collaboration.* [TT-3]

Regarding the *attitude change*, the FDIs revealed the opinion of the experimental group students about SG and GM:

> *...particularly useful in changing attitudes* [GM-3]
>
> *...a non-embarrassing learning environment* [SG-1]
>
> *...suitable and worthwhile for the delivery of sexual health education* [SG-7]
>
> *I was excited about the activities, competitions, leader-boards, badges, avatars, and scenarios* [GM-4]

By contrast, the control group students commented:

> *I was bored listening to lectures* [TT-1]
>
> *...[it was] hard to understand how the sexual health subject is important for changing my attitude* [TT-7]
>
> *...[it is] unfriendly learning* [TT-6]
>
> *...[it was an] uncomfortable learning environment due to the sensitivity of the topic taught* [TT-5]
>
> *...it hides potential information for changing my attitude* [TT-4]
>
> *...questions were not encouraged or not well clarified; hence I ended up with no clues that could change the myths [that produce] negative sexual health attitudes* [TT-2]

Experimental group students reported having had positive interactions with the SG and GM interventions and a substantial improvement in their sexual health *knowledge*:

> *I acquired the required knowledge for practicing healthy sexual behaviours through this learning approach* [GM-6]
>
> *I acquired potential sexual health knowledge that will help my making informed decisions* [SG-6]
>
> *I will apply the skills and understanding that are essential and applicable for curbing unhealthy sexual behaviours.* [SG-4]
>
> *from today onwards I will not participate in risky sexual activities as I am [now] knowledgeable and will apply the knowledge acquired to make informed decisions for better sexual health outcomes and future goals.* [SG-4]

By contrast, control group students commented on their teaching method as follows:

> *traditional teaching was less informative* [TT-3]
>
> *technical language was used that made it difficult to understand* [TT-7]
>
> *[there was a] lack of vivid examples* [TT-1]
>
> *the learning strategy narrowed the thinking capacity required for applying the knowledge and skills acquired* [TT-4]

Finally, students in the experimental learning groups (SG and GM) reported that the SG and GM components were effective in *engaging* them:

> *the learning activities made our minds active* [GM-6]
>
> *I was connected to the learning process* [SG-4]
>
> *I focused on the learning activities* [SG-4]

XSL•FO
RenderX

the learning made me concentrate on learning all the time [GM-2]

Students in SG-2 and GM-7 reported that

...the learning provided opportunities for hands-on activities that made it easy to learn and remember.

By contrast, the TT students reported that

our learning was indirect [TT-3, TT-5, TT-2]

the learning was passive, as no hands-on activities were provided [TT-1, TT4]

there was little or poor interaction [TT-7]

I lost focus during learning [TT-6]

Discussion

Findings and Interpretation

The study showed that the game elements embedded in SG and GM instruction catalyzed motivation and engagement during learning and that this contributed to attitude change, knowledge acquisition, and ultimately better learning performance.

The study results demonstrated that sexual health education taught using SG and GM approaches works better than TT methods. The SG and GM approaches resulted in higher test scores for knowledge acquisition than the TT control group. This finding conforms with previous research which found gamified learning systems to have a significant impact on sexual health education [25]. In our study, most students acknowledged several factors in their improved learning: The first factor is motivation (to learn), which consists of elements of attention, relevance, confidence, and satisfaction [49]. Indeed, motivation is a significant component for accomplishing or failing a task [50]. Students reported that their interest was caught and improved with the game elements (scenarios, quizzes, competition, challenges, scores) provided during the learning process. The game elements motivated the students to learn [45]. Furthermore, the game elements made the learning more interactive and fun, which increased the students' motivation to learn [50].

This study found that gamified learning contents were experienced as "relevant." For instance, the students realized that there was a common connection between what they were learning and real life. These results were consistent with an earlier study [19], which had indicated that gamifying sexual health learning approaches was promising for adolescents because the role plays and scenarios reflect the actual lifestyle of the current generation. The students felt confident while going through the self-regulated learning material provided through the gamified learning (which stimulated and sustained their learning). Perhaps this means that they would succeed in learning the subject matter to a great extent. As this paper demonstrates, such confidence enabled them to succeed and derive self-esteem from the knowledge they acquired and apply it in real life [49]. Although Keller [49] did not examine the effectiveness of the mediating role of increased knowledge in sexual health literacy, this work showed that sexual health education interventions through gamified learning are effective for the development of self-efficacy. As a result, they encourage healthy sexual practices including the digitally savvy adolescents [24]. Students reported satisfaction with their learning experience. The quality of the gamifying content gave them an experience of fun learning (thus, accomplishing learning goals). The awards and scoring mechanics also inspired them to learn with persistence and intensity [51].

The second factor accounting for the effectiveness of the gamified learning was that it changed the attitude of the students. As a similar study [7] reported, gamified learning induces positive changes in the sexual health attitude of adolescents. Essentially, the My Future Begins Today gamified learning incorporated in its design most of the known relevant features that have proved effective in transforming adolescents' negative sexual health attitudes to positive ones and, as a consequence, curbing risky sexual behavior [10]. It considered specific settings and co-opted various stakeholders, including the targeted users (high-school students and their teachers) in the design.

During the gamified design and development, students provided input on the type of avatars and scenarios they found appealing. Their opinions were based on their different cultural settings, their level of digital literacy, their use of state-of-the-art technology, among others. This study was grounded on the social–cultural theory known as activity theory that encourages participation of different stakeholders in the development of instructional interventions [52-54]. Members of the community participated in the design process by contributing to the design of the knowledge-acquisition components useful in addressing the problems related to acquisition of sexual health knowledge among adolescents in the studied region. The resultant gamified learning elements were found to be relevant in changing the students' attitude toward problems such as negative peer pressure, teen pregnancy, STIs (including HIV/AIDS), and sexual violence. Although this research found no participatory design being applied within the TT environment, the gamified learning instructions invited the targeted users to participate in the design. This allowed us to apprehend the participants' relevant ideas and needs and in turn effectively deliver the sexual health information required to induce a change in attitudes of the participants [15].

The third factor accounting for the efficacy of gamified learning was that it improved the knowledge acquisition among students in the experimental groups. The students indicated that gamified learning helped them to acquire sexual health information and skills that could purportedly help them engage in healthy sexual practices. Notably, the students reported that the sexual health knowledge delivered through gamified instructions were highly effective for their current and future lives and that they now felt knowledgeable and able to resist detrimental sexual health risks or factors. According to Chu et al [15], gamified learning is effective because knowledge is acquired in a safe, nonrisk (simulated) environment. Gamified learning offers students the opportunity to experiment in their learning, play, apply decision-making skills, and test scenarios without negative consequences. Evidence from follow-up studies shows that the effect of sexual health knowledge acquired through gamified learning is compelling and persistent, but no effectiveness has been demonstrated in terms of delay in sexual initiation [7], although the determination of outcomes was based on

XSL•FO
RenderX

self-reporting. An iterative design study with the objective to describe a methodology for developing an SG intervention for improving sexual health education among youth in Boston [8] demonstrated that nonidentified study participants (ie, students and street youths) in underserved communities would have acquired more knowledge on chlamydia because they enjoyed the gameplay and actively participated in acquiring the information. This may explain why the experimental groups in our study showed better results than their counterparts in the control group.

The fourth and final factor that boosted students' learning was engagement. The evaluation of the effectiveness of each of the instructional approaches was based on the students' engagement. Studies have documented the effectiveness of gamified learning in engaging students during learning [8,42,43], with some studies related to the ability of gamified instructions to entertain and reduce stress when learning [43,50]. The My Future Begins Today (GM and SG) platform was designed to trigger students' engagement by having learning tasks performed in a problem-solving way and by having students participate in skills-based challenges that required critical thinking. The gamified learning platforms are useful for the students, especially when it comes to (1) the skills that are needed to thrive or (2) use the latest technologies of this century. The presence of game elements (badges, score, leaderboards, levels, immediate feedback, time pressure, and repetition) may explain why gamified learning increased the engagement of students in the experimental group, which, in turn, bolstered their learning. Besides, game elements were positively commented on by the students who saw increased engagement with the PR:EPARe game [42]. Although Jiang et al [8] did not find a significant correlation between the participants' game engagement and learning, this study shows that the 2 concepts are useful for learning purposes.

Limitations

This study used a participatory design approach. Such an approach is vital in designing instructions that address the needs of the users [15,43] in their social–cultural context. It was informed by design-based research (from a technology perspective) [32] and grounded in sociocultural learning theory [53,54]. As good as the foundations are, we must acknowledge some limitations in our efforts to put them into practice. This study evaluated knowledge acquired by the students and the effectiveness of that knowledge in changing their attitudes toward sexual activities. However, it is would be difficult to determine how much of and for how long the change took place after the study. Hence, would need to know how many students dropped out of school due to pregnancy, were infected with STIs, encountered sexual violence, or were peer pressured into harmful sexual practices. A follow-up study could be conducted when the students are about to finish their ordinary-level studies.

Conclusions

Educational gamified learning (GM and SG) has the potential to significantly increase the sexual health literacy of adolescents. The digital health gamified interventions designed in this study provide a user-friendly learning environment. The designs were influenced by a theory-driven assessment of learning. This assessment is supported by testing the learning of the users. This study treats serious digital health gamified instructions as a brain activator: it keeps students active during the learning process. The students' participation in the learning process is catalyzed by the *motivation* and *engagement* that are enabled by the game elements and mechanics. The My Future Begins Today (the digital health gamified learning interventions using the SG and GM) design increases knowledge acquisition and attitude change. Students reported the learning to be more interactive through participating in the gamified learning activities.

SG and GM methods were found to be effective and efficient in increasing motivation, improving attitudes, increasing knowledge acquisition, and encouraging engagement in the learning process. Future empirical studies may verify the efficacy of the My Future Begins Today learning platforms in improving sexual health literacy acquisition in other countries, especially in SSA, where the TT method is widely practiced and tends to limit the learning with gamified digital technologies and process [19,25]. This paper also addresses more than ever the call from a previous study [25] to evaluate the effect of gamifying sexual health education when different key stakeholders are involved in the design process in low-tech settings. This is due to the fact that in developing countries information and communication technologies and digital literacy are limited.

Conflicts of Interest

Non declared.

Multimedia Appendix 1
Comparison of average scores, one-way ANOVA, and post hoc comparison tests after the lessons.
[DOCX File , 17 KB - games_v9i4e19614_app1.docx]

Multimedia Appendix 2
Evaluation of the efficacy of the instructional approachesa.
[DOCX File , 16 KB - games_v9i4e19614_app2.docx]

References

XSL•FO
RenderX

1. Thompson R, Lewis T, Neilson EC, English DJ, Litrownik AJ, Margolis B, et al. Child Maltreatment and Risky Sexual Behavior. Child Maltreat 2017 Feb 24;22(1):69-78 [FREE Full text] [doi: 10.1177/1077559516674595] [Medline: 27777330]
2. Sani A, Abraham C, Denford S, Mathews C. Design, implementation and evaluation of school-based sexual health education in sub-Saharan Africa: a qualitative study of researchers' perspectives. Sex Education 2017 Dec 21;18(2):172-190. [doi: 10.1080/14681811.2017.1414040]
3. Haruna H, Hu X, Wah Chu SK. Adolescent School-Based Sexual Health Education and Training: A Literature Review on Teaching and Learning Strategies. GJHS 2018 Feb 19;10(3):172. [doi: 10.5539/gjhs.v10n3p172]
4. Zimmerman GM, Farrell C. Parents, Peers, Perceived Risk of Harm, and the Neighborhood: Contextualizing Key Influences on Adolescent Substance Use. J Youth Adolesc 2017 Jan;46(1):228-247. [doi: 10.1007/s10964-016-0475-5] [Medline: 27016218]
5. Sani AS, Abraham C, Denford S, Ball S. School-based sexual health education interventions to prevent STI/HIV in sub-Saharan Africa: a systematic review and meta-analysis. BMC Public Health 2016 Oct 10;16(1):1069 [FREE Full text] [doi: 10.1186/s12889-016-3715-4] [Medline: 27724886]
6. Cates JR, Fuemmeler BF, Diehl SJ, Stockton LL, Porter J, Ihekweazu C, et al. Developing a Serious Videogame for Preteens to Motivate HPV Vaccination Decision Making: Land of Secret Gardens. Games Health J 2018 Feb;7(1):51-66 [FREE Full text] [doi: 10.1089/g4h.2017.0002] [Medline: 29161529]
7. Fiellin LE, Hieftje KD, Pendergrass TM, Kyriakides TC, Duncan LR, Dziura JD, et al. Video Game Intervention for Sexual Risk Reduction in Minority Adolescents: Randomized Controlled Trial. J Med Internet Res 2017 Sep 18;19(9):e314 [FREE Full text] [doi: 10.2196/jmir.8148] [Medline: 28923788]
8. Jiang R, McKanna J, Calabrese S, Seif El-Nasr M. Iterative Design and Testing for the Development of a Game-Based Chlamydia Awareness Intervention: A Pilot Study. Games Health J 2017 Aug;6(4):205-216 [FREE Full text] [doi: 10.1089/g4h.2016.0112] [Medline: 28816513]
9. Temple-smith M, Moore S, Rosenthal D. Sexuality in Adolescence: The digital generation. New York: Taylor & Francis Group; 2016.
10. Goesling B, Colman S, Trenholm C, Terzian M, Moore K. Programs to reduce teen pregnancy, sexually transmitted infections, and associated sexual risk behaviors: a systematic review. J Adolesc Health 2014 May;54(5):499-507. [doi: 10.1016/j.jadohealth.2013.12.004] [Medline: 24525227]
11. Mullinax M, Mathur S, Santelli J. Adolescent sexual healthsexuality education. In: Cherry A, Baltag V, Dillon M, editors. International Handbook on Adolescent Health and Development. Cham, Switzerland: Springer International Publishing; 2017:143-167.
12. Aarø LE, Mathews C, Kaaya S, Katahoire AR, Onya H, Abraham C, et al. Promoting sexual and reproductive health among adolescents in southern and eastern Africa (PREPARE): project design and conceptual framework. BMC Public Health 2014 Jan 18;14:54 [FREE Full text] [doi: 10.1186/1471-2458-14-54] [Medline: 24438582]
13. Pesare E, Roselli T, Corriero N, Rossano V. Game-based learning and Gamification to promote engagement and motivation in medical learning contexts. Smart Learn. Environ 2016 Apr 27;3(1):1-21. [doi: 10.1186/s40561-016-0028-0]
14. Cederbaum JA, Rodriguez AJ, Sullivan K, Gray K. Attitudes, Norms, and the Effect of Social Connectedness on Adolescent Sexual Risk Intention. J Sch Health 2017 Aug 09;87(8):575-583. [doi: 10.1111/josh.12532] [Medline: 28691176]
15. Chu SKW, Kwan ACM, Reynolds R, Mellecker RR, Tam F, Lee G, et al. Promoting Sex Education Among Teenagers Through an Interactive Game: Reasons for Success and Implications. Games Health J 2015 Jun;4(3):168-174. [doi: 10.1089/g4h.2014.0059] [Medline: 26182060]
16. Goldman JDG. An exploration in health education of an integrated theoretical basis for sexuality education pedagogies for young people. Health Educ Res 2011 Jun;26(3):526-541. [doi: 10.1093/her/cyq084] [Medline: 21602573]
17. Kachur R, Mesnick J, Liddon N, Kapsimalis C, Habel M, David-Ferdon C, et al. Adolescents, Technology and Reducing Risk for HIV, STDs and Pregnancy. Atlanta, GA: Centers for Disease Control and Prevention; 2013. URL: https://www.cdc.gov/std/life-stages-populations/adolescents-tech.htm [accessed 2020-06-22]
18. Eleftheriou A, Bullock S, Graham CA, Ingham R. Using Computer Simulations for Investigating a Sex Education Intervention: An Exploratory Study. JMIR Serious Games 2017 May 03;5(2):e9 [FREE Full text] [doi: 10.2196/games.6598] [Medline: 28468747]
19. Patchen L, Ellis L, Ma TX, Ott C, Chang KH, Araya B, et al. Engaging African American Youth in the Development of a Serious Mobile Game for Sexual Health Education: Mixed Methods Study. JMIR Serious Games 2020 Jan 31;8(1):e16254 [FREE Full text] [doi: 10.2196/16254] [Medline: 32012041]
20. Kimberly H, Duncan LR, Pendergrass T, Sawyer B, Fiellin LE. Development of an HIV prevention videogame intervention: lessons learned. IJSG 2016 Apr;3(2):83-90 [FREE Full text] [doi: 10.17083/ijsg.v1i4.47]
21. Lenhart A, Kahne J, Middaugh E E, Rankin Macgill A, Evans C, Vitak J. Teens, Video Games, and Civics. 2008 Sep 16. URL: https://files.eric.ed.gov/fulltext/ED525058.pdf [accessed 2021-09-22]
22. Horn M, Staker H. Trends in digital learning: Empowering innovative classroom models for learning. Proj Tomorrow. URL: http://www.tomorrow.org/speakup/2015_ClassroomModels.html [accessed 2020-07-02]

XSL•FO
RenderX

23. Haruna H, Hu X, Chu S, Mellecker R, Gabriel G, Ndekao P. Improving Sexual Health Education Programs for Adolescent Students through Game-Based Learning and Gamification. Int J Environ Res Public Health 2018 Sep 17;15(9):2027 [FREE Full text] [doi: 10.3390/ijerph15092027] [Medline: 30227642]
24. Chib A. Promoting sexual health education via gaming: Evidence from the Barrios of Lima, Peru. In: Patrick F, editor. Handbook of Research on Improving Learning and Motivation Through Educational Games: Multidisciplinary Approaches. Hershey, PA: IGI Global; 2011:895-912.
25. Sharifzadeh N, Kharrazi H, Nazari E, Tabesh H, Edalati Khodabandeh M, Heidari S, et al. Health Education Serious Games Targeting Health Care Providers, Patients, and Public Health Users: Scoping Review. JMIR Serious Games 2020 Mar 05;8(1):e13459 [FREE Full text] [doi: 10.2196/13459] [Medline: 32134391]
26. Haruna H, Zainuddin Z, Mellecker RR, Chu SK, Hu X. An iterative process for developing digital gamified sexual health education for adolescent students in low-tech settings. ILS 2019 Nov 11;120(11/12):723-742. [doi: 10.1108/ils-07-2019-0066]
27. Feser E. Isserman's Impact. International Regional Science Review 2012 Nov 09;36(1):44-68. [doi: 10.1177/0160017612464051]
28. Thyer B. Quasi-Experimental Research Designs. New York, NY: Oxford University Press; 2012.
29. Creswell J, Creswell D. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: SAGE; 2018.
30. Creswell J. Educational Research : Planning, Conducting, and Evaluating Quantitative and Qualitative Research (4th Edition). New York, NY: Pearson; 2011.
31. Plant T, Mann D. Introduction to new insights into the neurobiology of reproduction and puberty. Brain Res 2010 Dec 10;1364:1-2. [doi: 10.1016/j.brainres.2010.11.002] [Medline: 21109037]
32. Reeves T. Design research from a technology perspective. In: Educational Design Research. London, UK: Routledge; 2006:52-66.
33. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, De Bourdeaudhuij I. Is Participatory Design Associated with the Effectiveness of Serious Digital Games for Healthy Lifestyle Promotion? A Meta-Analysis. J Med Internet Res 2016 Apr 29;18(4):e94 [FREE Full text] [doi: 10.2196/jmir.4444] [Medline: 27129447]
34. Bai S, Hew KF, Huang B. Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. Educational Research Review 2020 Jun;30:100322. [doi: 10.1016/j.edurev.2020.100322]
35. Zainuddin Z, Shujahat M, Haruna H, Chu S. The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. Computers & Education 2020 Feb 7;145:103729. [doi: 10.1016/j.compedu.2019.103729]
36. Kapp K. The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education. Hoboken, NJ: John Wiley and Sons; May 15, 2012.
37. Huang B, Hew KF, Lo CK. Investigating the effects of gamification-enhanced flipped learning on undergraduate students' behavioral and cognitive engagement. Interactive Learning Environments 2018 Jul 20;27(8):1106-1126. [doi: 10.1080/10494820.2018.1495653]
38. Zichermann G. Intrinsic and Extrinsic Motivation in Gamification. 2011. URL: http://www.gamification.co/2011/10/27/intrinsic-and-extrinsic-motivation-in-gamification/ [accessed 2018-01-09]
39. Dicheva D, Dichev C, Agre G, Angelova G. Gamification in Education: A Systematic Mapping Study. J Educ Technol Soc 2015;18(3):88.
40. Sugawara E, Nikaido H. Properties of AdeABC and AdeIJK efflux systems of Acinetobacter baumannii compared with those of the AcrAB-TolC system of Escherichia coli. Antimicrob Agents Chemother 2014 Dec;58(12):7250-7257 [FREE Full text] [doi: 10.1128/AAC.03728-14] [Medline: 25246403]
41. Haruna H, Hu X, Chu S, Mellecker R. Initial Validation of the MAKE Framework: A Comprehensive Instrument for Evaluating the Efficacy of Game-Based Learning and Gamification in Adolescent Sexual Health Literacy. Ann Glob Health 2019 Feb 28;85(1):19 [FREE Full text] [doi: 10.5334/aogh.1110] [Medline: 30873788]
42. Arnab S, Brown K, Clarke S, Dunwell I, Lim T, Suttie N, et al. The development approach of a pedagogically-driven serious game to support Relationship and Sex Education (RSE) within a classroom setting. Computers & Education 2013 Nov;69:15-30. [doi: 10.1016/j.compedu.2013.06.013]
43. Arnab S, Clarke S. Towards a trans-disciplinary methodology for a game-based intervention development process. Br J Educ Technol 2015 Dec 29;48(2):279-312. [doi: 10.1111/bjet.12377]
44. Bowen E, Walker K, Mawer M, Holdsworth E, Sorbring E, Helsing B, et al. "It's like you're actually playing as yourself": Development and preliminary evaluation of 'Green Acres High', a serious game-based primary intervention to combat adolescent dating violence+. Psychosocial Intervention 2014 Apr;23(1):43-55. [doi: 10.5093/in2014a5]
45. Rose J, O'Meara J, Gerhardt T, Williams M. Gamification: using elements of video games to improve engagement in an undergraduate physics class. Phys. Educ 2016 Jul 22;51(5):055007. [doi: 10.1088/0031-9120/51/5/055007]
46. Hyde A, Howlett E, Brady D, Drennan J. The focus group method: insights from focus group interviews on sexual health with adolescents. Soc Sci Med 2005 Dec;61(12):2588-2599. [doi: 10.1016/j.socscimed.2005.04.040] [Medline: 15955604]
47. Vaughn S, Schumm J, Sinagub J. Focus Group Interviews in Education and Psychology. London, UK: SAGE; 1996.

XSL•FO
RenderX

48. Rajalahti T, Kvalheim OM. Multivariate data analysis in pharmaceutics: a tutorial review. Int J Pharm 2011 Sep 30;417(1-2):280-290. [doi: 10.1016/j.ijpharm.2011.02.019] [Medline: 21335075]
49. Keller JM. Motivation, Learning, and Technology: Applying the ARCS-V Motivation Model. Per 2016 Aug 01;3(2):1-15. [doi: 10.17275/per.16.06.3.2]
50. Baskaran G, Rawi MIM. Emotion Digital Game-Based Learning Application: ARCS Motivational Model? International Journal of Advanced Computational Engineering and Networking 2017 Sep;5(9):148 [FREE Full text]
51. Garris R, Ahlers R, Driskell JE. Games, Motivation, and Learning: A Research and Practice Model. Simulation & Gaming 2016 Aug 18;33(4):441-467. [doi: 10.1177/1046878102238607]
52. Hung DWL, Wong AFL. Activity theory as a framework for project work in learning environments. Educ Technol 2000;40(2):33-37.
53. Kaptelinin V. Activity Theory Internet. 2013. URL: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/activity-theory [accessed 2020-05-18]
54. Roth W, Radford L, La CL. Working with cultural-historical activity theory. Forum Qual Sozialforsch 2012;13(2):1-20. [doi: 10.17169/FQS-13.2.1814]

Abbreviations

ASHLT: Adolescent Sexual Health Literacy Test
AVOVA: analysis of variance
FGI: focus group interview
SG: serious game
GM: gamification
MAKE: motivation, attitude, knowledge, and engagement
SSA: sub-Saharan Africa
STI: sexually transmitted infections
TT: traditional teaching

Edited by G Eysenbach; submitted 24.04.20; peer-reviewed by M Tara, S Diehl; comments to author 26.10.20; revised version received 09.01.21; accepted 31.05.21; published 12.10.21.

Please cite as:
Haruna H, Okoye K, Zainuddin Z, Hu X, Chu S, Hosseini S
Gamifying Sexual Education for Adolescents in a Low-Tech Setting: Quasi-Experimental Design Study
JMIR Serious Games 2021;9(4):e19614
URL: https://games.jmir.org/2021/4/e19614
doi:10.2196/19614
PMID:34636739

©Hussein Haruna, Kingsley Okoye, Zamzami Zainuddin, Xiao Hu, Samuel Chu, Samira Hosseini. Originally published in JMIR Serious Games (https://games.jmir.org), 12.10.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evidence of Construct Validity of Computer-Based Tests for Clinical Reasoning: Instrument Validation Study

Tianming Zuo[1], MD, PhD; Baozhi Sun[1], MD, PhD; Xu Guan[2], MD; Bin Zheng[3], MD, PhD; Bo Qu[1], MD, PhD

[1]Institute for International Health Professions Education and Research, China Medical University, Shenyang, China

[2]Education Center for Clinical Skills Practice, China Medical University, Shenyang, China

[3]Surgical Simulation Research Lab, Department of Surgery, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Bo Qu, MD, PhD
Institute for International Health Professions Education and Research
China Medical University
No. 77 Puhe Road
Shenyang, 110122
China
Phone: 86 189 0091 0198
Email: qubo6666@163.com

Abstract

Background: Clinical reasoning (CR) is a fundamental skill for all medical students. In our medical education system, however, there are shortcomings in the conventional methods of teaching CR. New technology is needed to enhance our CR teaching, especially as we are facing an influx of new health trainees. China Medical University (CMU), in response to this need, has developed a computer-based CR training system (CMU-CBCRT).

Objective: We aimed to find evidence of construct validity of the CMU-CBCRT.

Methods: We recruited 385 students from fifth year undergraduates to postgraduate year (PGY) 3 to complete the test on CMU-CBCRT. The known-groups technique was used to evaluate the construct validity of the CBCRT by comparing the test scores among 4 training levels (fifth year MD, PGY-1, PGY-2, and PGY-3).

Results: We found that test scores increased with years of training. Significant differences were found in the test scores on information collection, diagnosis, and treatment and total scores among different training years of participants. However, significant results were not found for treatment errors.

Conclusions: We provided evidence of construct validity of the CMU-CBCRT, which could determine the CR skills of medical students at varying early stage in their careers.

(JMIR Serious Games 2021;9(4):e17670) doi:10.2196/17670

KEYWORDS

medical education; assessment; computer-based test; clinical reasoning; validity

Introduction

Each year, several hundred thousand students enter medical school, all of whom need to equip themselves with the necessary health care skills and knowledge [1]. Since 2014, the vast majority of Chinese medical students attend a 5-year program after high school to earn a bachelor's degree. Then, they work in a 1-year clinical internship before taking the nation's standardized medical licensure exams. If successful, they may register as medical practitioners. Postgraduate training in medical specialties is standardized to 3-year programs with the final credential called Master of Medicine; this is now required of all clinical practitioners. In addition to learning a broad range of medical knowledge and practicing dexterity in hands, practitioners need to learn how to collect information from patients, process this information, and make accurate diagnostic decisions, similar to the expectations from a senior physician [2,3]. Clinical reasoning (CR) is a fundamental skill that separates medical personnel from other professionals. William Osler, a legendary pioneer medical educator, emphasized proper physical examination and diagnostic reasoning while maintaining the intimate physician-patient relationship. His teachings have resonated with generations of physicians [4]. Strictly speaking, CR refers to the procedure of collecting and

XSL•FO
RenderX

integrating patient information from various sources to arrive at a diagnosis and management plan; it is usually case specific [5]. Every medical teaching institute makes a great effort to understand the nature of CR and improve strategies for teaching CR skills to health trainees [6]. However, the conventional methods that are used in our education system today are not optimal [7-11].

Traditionally, CR is taught in the classroom (didactic lecture) and by the patient's side (clinical clerkship) [12-15]. A recent focus of integrating problem-based learning (PBL) has significantly improved the quality of CR education [6,12,16-18]. However, PBL relies heavily on the involvement and commitment of faculty instructors, which may not always be feasible [16,19]. Fidelity of case is also a problem compared to patient-side education [12]. Acquiring patient information by reading PBL cases from charts is quite a different experience than taking information directly from patients. Although instructors are making PBL cases in collaboration with clinicians, students still report a lack of case variety [17,20]. Creating sufficient clinical cases with clinical fidelity for CR training is a difficult task. Due to the above reasons, new technology is needed to improve our CR teaching.

In contrast to a paper- or lecture-based curriculum, computer-based CR training allows trainees to interactively take information from patients in a step-by-step process. There is also the possibility of accumulating a large volume of cases through international collaboration.

Currently, computer-based CR training can have different interfaces such as text, graphics, and animation [21]. The text-based CR training system is most widely used [22]. It is easy to create from clinical cases and deliver in the format of multiple choice questions or direct interface [23]. While medical images (including x-ray films, electrocardiograms, photos of lesions, etc) are required to give students more clinical information, graphic interface is also necessary. In several graphical models, illustrations of patients (in drawing or 2D pictures) can be used to create interactive experience for students when they collect information from patients [24]. Some computer-based CR training includes 3D animation or virtual reality technology to simulate the clinical scenario with high fidelity. However, the cost of creating 3D animation and virtual reality scenarios is much higher than the other computer-based CR models. It is difficult to create virtual patients without a team of technicians and instructional designers (Table 1).

Table 1. Types of computer-based clinical reasoning simulations and comparison.

Media	Advantage	Disadvantage
Text based	Relatively easy and rapid to develop; less expensive	Low level of fidelity
Graphic and animation based	Presents rich clinical evidence; moderate cost with enhanced fidelity	Replicates only part of clinical settings; low level of interactivity
Virtual reality	Combines highly sophisticated, life-like models with computer animations; can provide interactivity and feedback	Challenge to developers; often expensive

Sponsored by the National Medical Examination Center of China, China Medical University (CMU) started to developed computer-based CR training system in 2001. Educators and researchers at the Institute for International Health Professions Education and Research of CMU began to work with clinicians to develop cases for training CR skills and established the computer-based CR testing (CBCRT) system. Since 2002, CBCRT has been used as one part in the final comprehensive examinations of CMU to test the clinical skills of undergraduate students.

The CBCRT is composed of 5 interactive modules that allow students to interact with simulations to complete tasks: (1) history taking and physical examination, (2) writing orders and obtaining lab and medical imaging results, (3) reviewing obtained results, (4) working out diagnosis and differential diagnosis, and (5) observing the patient's condition change at different phases and changing locations for managing different therapies. The main features of the CMU-CBCRT virtual patient are displayed in Figure 1.

To test the face validity of the CMU-CBCRT, we called a series of meetings with physicians and surgeons at which we screened and selected key information on each clinical topic for CR training. When a CR case was developed, our clinical team was surveyed to verify their clinical relevance. They then evaluated the interactive interface and rated their level of satisfaction.

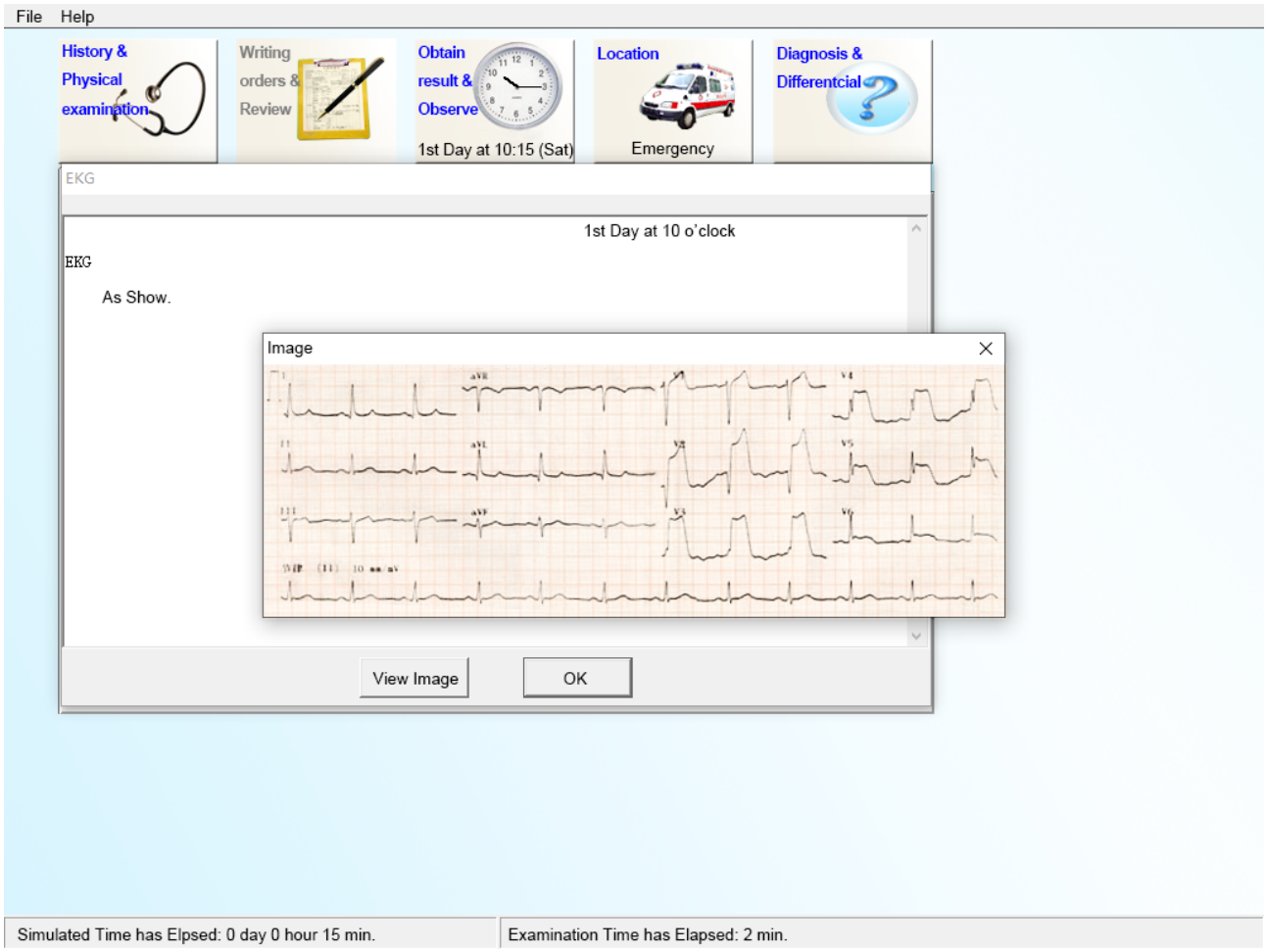
To briefly summarize, the CBCRT provided clinical features of patients including history and physical and laboratory findings and then requires students to make a diagnosis as well as a treatment plan for the simulated patients. The CBCRT has also been welcomed by the examinees based on their positive feedback toward the system. Of 300 students surveyed using the questionnaire, 99.4% enjoyed participating in the CBCRT examination; 95.9% believed that the system accurately represents the real clinical environment; 72.5% agreed that the CBCRT is a better tool for teaching their clinical abilities. We can thus believe that the face validity of the CBCRT is satisfactory.

However, face validity is the weakest form of validity evidence. It can be only used at the primary stage of designing an assessment method [25]. We need to look into the structure of the CBCRT in detail to find more evidence of its validity, especially since there is a paucity of validity evidence for computer-based CR training [22]. In China, this study is the first of its type.

This study investigates construct validity of the CMU-CBCRT in medical trainees over 5 medical school. We hypothesize that the CMU-CBCRT will be able to determine CR level among different years of health trainees; specifically, senior trainees will achieve higher CBCRT scores than juniors.

XSL•FO
RenderX

Figure 1. Screenshot displaying the main features of the China Medical University–computer-based clinical reasoning testing.



Methods

Ethical Statement

Methods used for the project were reviewed and approved by the ethical review boards of the CMU (ERB 2016-027) and the 5 medical schools. Informed consent was obtained from each participant before they started the test with the CMU-CBCRT.

Testing Sites

From November 24 to December 8, 2016, we implemented the CMU-CBCRT system in 5 collaborative medical schools: China Medical University, Fudan University School of Medicine, Sun Yat-sen University School of Medicine, Xuzhou Medical University, and Binzhou Medical College.

Students

In China, medical students start their clerkships on the fifth year of medical training. The clinical training will continue for their 3 postgraduate years (PGYs). PGY 1 to PGY 3 is similar to the residency in North American medical school. We recruited students from fifth year undergraduates to PGY 3. The actual number of participants from each of 5 medical schools and their training years are shown as Table 2.

Table 2. Students from the 5 medical schools and their training years.

Schools	Fifth year medical student	PGY-1[a]	PGY-2	PGY-3	Subtotal
China Medical University	40	18	7	2	67
Fudan University School of Medicine	17	41	16	18	92
Sun Yat-sen University School of Medicine	12	28	20	12	72
Xuzhou Medical University	20	19	20	21	80
Binzhou Medical College	20	19	19	16	74
Total	109	125	82	69	385

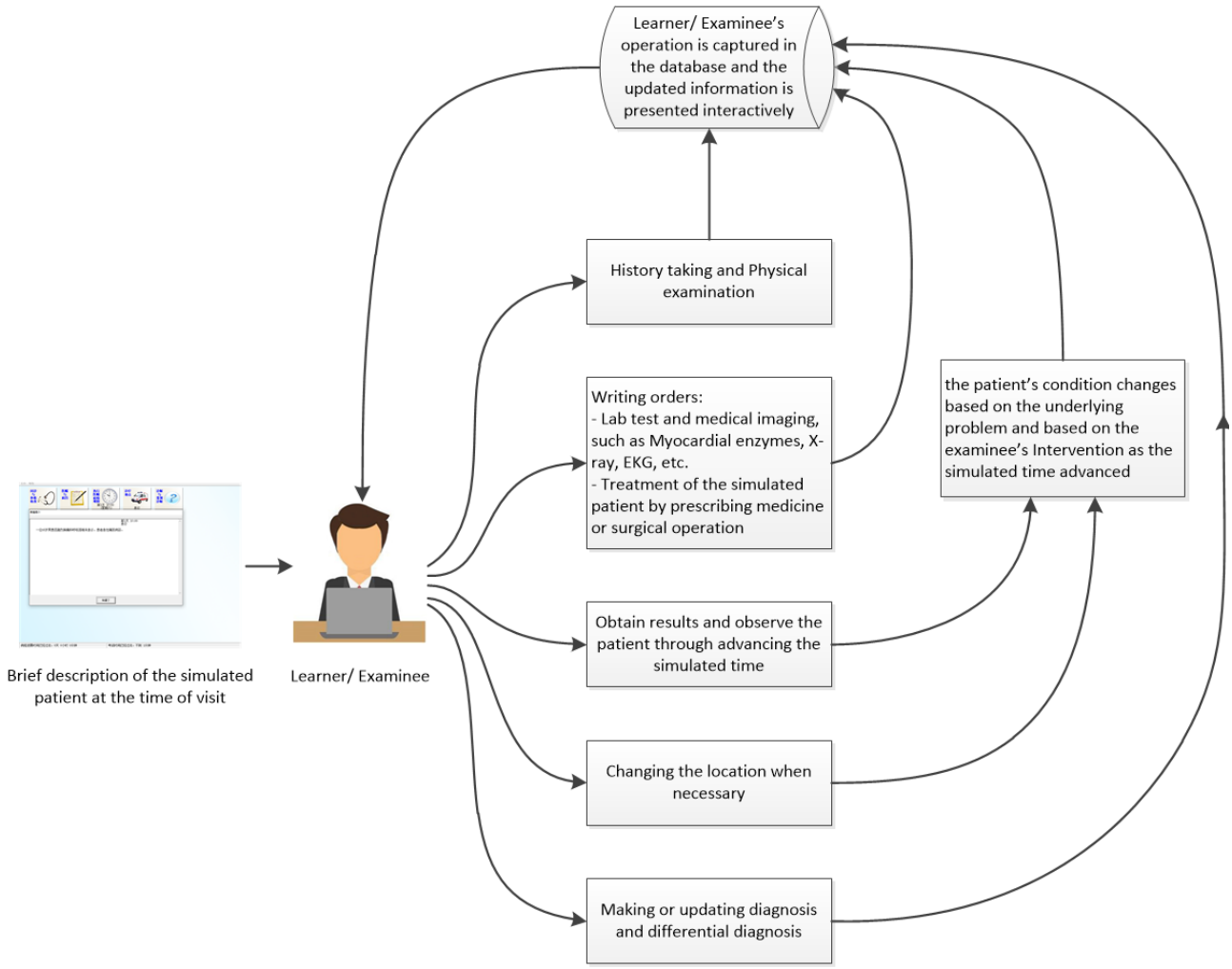
[a]PGY: postgraduate year.

XSL•FO
RenderX

Measures

Before testing, each student was asked to watch a 5-minute presentation and get familiar with the testing interface. Demographics and level of medical training were surveyed and recorded. The computer recorded participants' typing and computer activity, including the typing and performance times. The interaction between a learner and how data are captured is displayed in Figure 2. Once completing the testing on CMU-CBCRT, the system calculated and recorded their total score by comparing the participants' transaction list with the scoring scheme defined by the case developers committee (Multimedia Appendix 1). Subscores on these 4 different areas: information collection, diagnosis, treatment, and treatment error are computed and recorded as well (Multimedia Appendix 2).

Figure 2. Outline of interaction flow through China Medical University–computer-based clinical reasoning testing.



Statistical Model

The known-groups technique was used to evaluate the construct validity of the CBCRT by comparing the scores among the fifth year MD, PGY-1, PGY-2, and PGY-3 participants. Testing scores, including total and subtotal, were compared over the 4 training groups using a 1-way analysis of variance (ANOVA). Results were reported as mean and standard deviation. $P \leq .50$ was considered a significant difference among testing groups.

Results

Total Score

ANOVA revealed a group difference in total score among training levels ($P<.001$). As shown in Table 3 and Figure 3, the score of the fifth year MD students (59.01 [SD 16.68]) was significantly lower than the PGY-2 (68.68 [SD 11.76]) and PGY-3 (68.06 [SD 12.67]) students; the total score of PGY-1 students was also significantly lower than the PGY-2 and PGY-3 students.

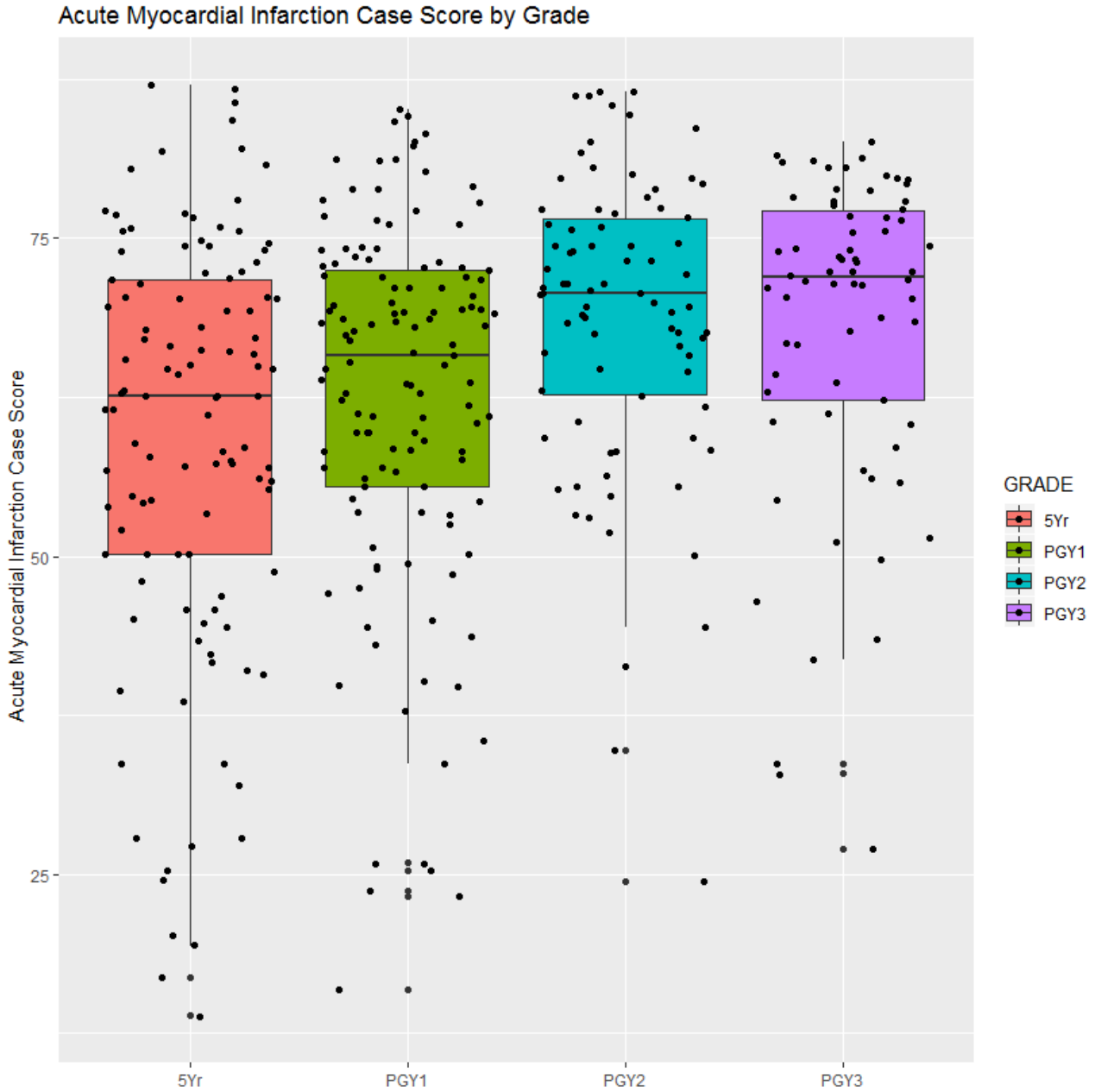
XSL•FO
RenderX

Table 3. Students from the 5 medical schools and their training years.

Scores	Fifth year medical student, mean (SD)	PGY-1[a], mean (SD)	PGY-2, mean (SD)	PGY-3, mean (SD)	*P* value
Information collection	43.42 (12.63)	46.70 (11.48)	49.73 (9.12)	51.38 (9.08)	<.001
Diagnosis	10.90 (4.74)	11.24 (4.97)	12.76 (3.90)	11.25 (4.22)	.034
Treatment	4.79 (3.81)	4.61 (3.36)	6.19 (3.73)	5.45 (3.72)	.013
Treatment error	–0.06 (023)	–0.04 (0.20)	0.00 (0.00)	–0.01 (0.12)	.13
Total	59.01 (16.68)	62.50 (14.45)	68.68 (11.76)	68.06 (12.67)	.001

[a]PGY: postgraduate year.

Figure 3. Total score of students over training years.



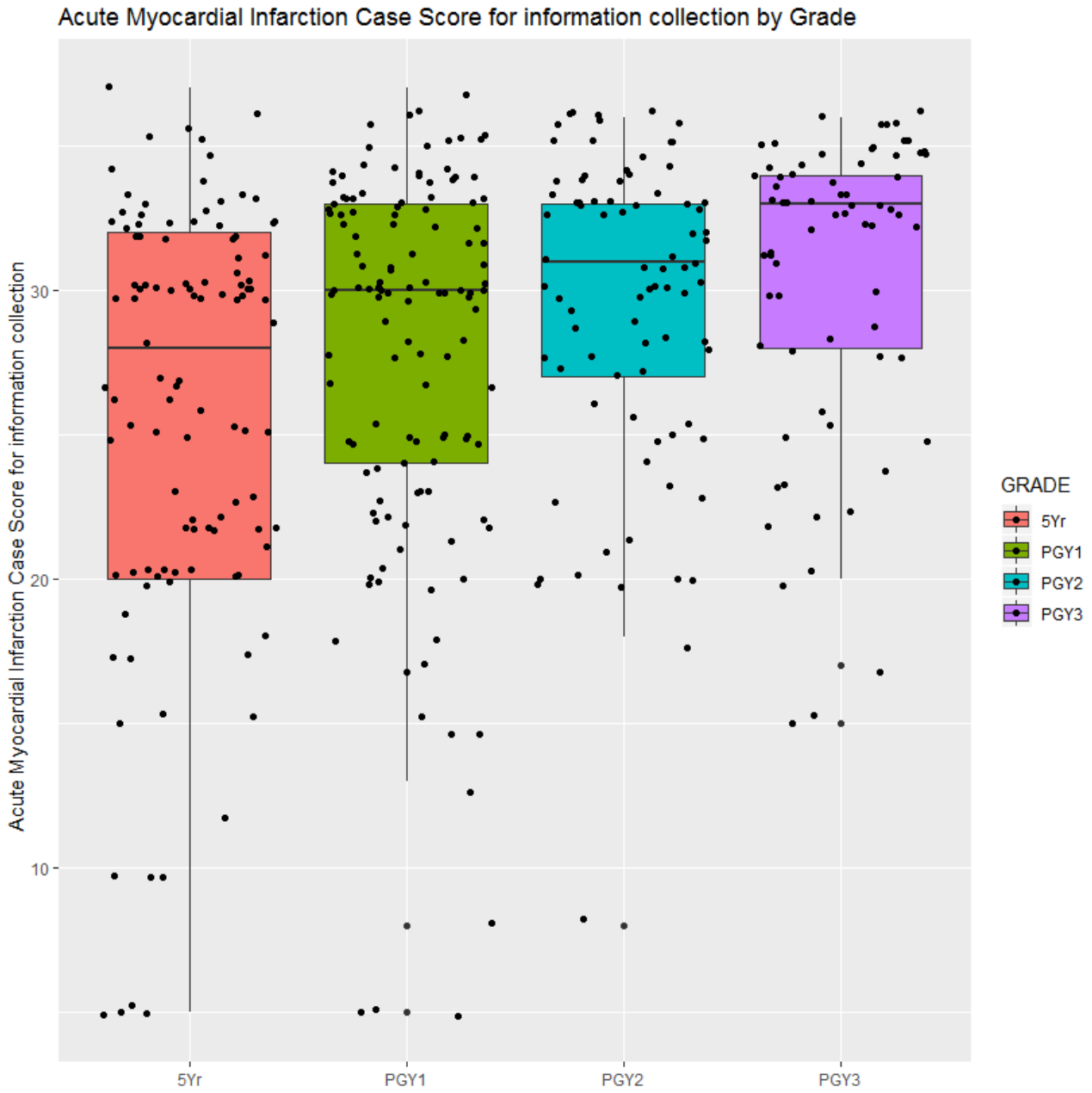
Subscore

ANOVA revealed group differences by training level between information collection (*P*<.001), diagnosis (*P*=.03), and treatment (*P*=.01) scores, but not on treatment error (*P*=.13) score. As shown in Figure 4, the information collection scores of the fifth year MD students (43.42 [SD 12.63]) were significantly lower than the PGY-1 (46.70 [SD 11.48]), PGY-2

XSL•FO
RenderX

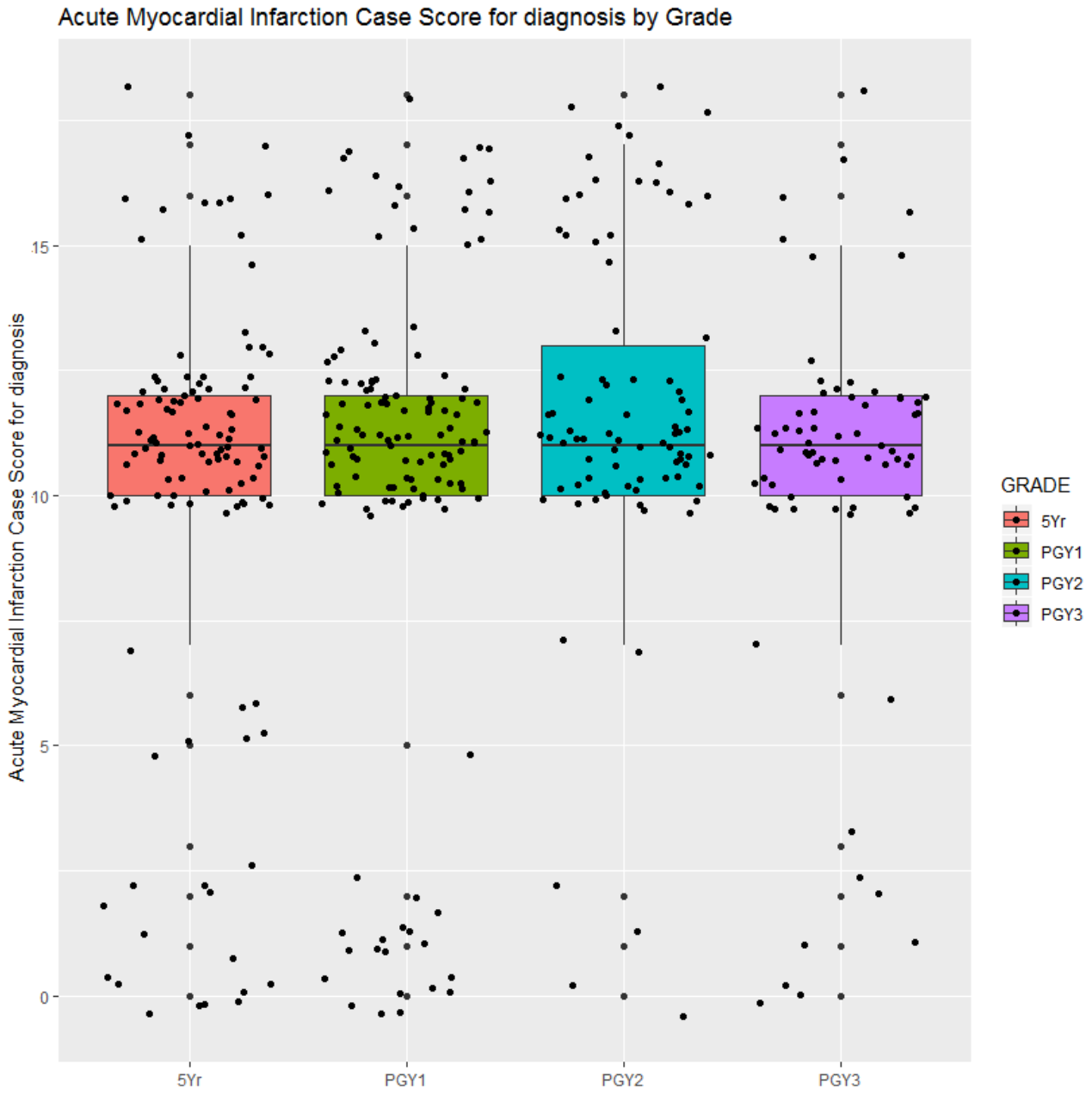
(49.73 [SD 9.12]), and PGY-3 (51.38 [SD 9.08]) students; information collection scores of PGY-1 students were also significantly lower than the PGY-3 students. As shown in Figure 5, the diagnosis scores of the fifth year MD (10.90 [SD 4.74]), PGY-1 (11.24 [SD 4.97]), and PGY-3 (11.25 [SD 4.22]) students were significantly lower than the PGY-2 (12.76 [SD 3.90]) students. As shown in Figure 6, treatment scores of the fifth year MD (4.79 [SD 3.81]) and PGY-1 (4.61 [SD 3.36]) students were significantly lower than the PGY-2 (6.19 [SD 3.73]) students.

Figure 4. Subscore for information collection of students over training years.



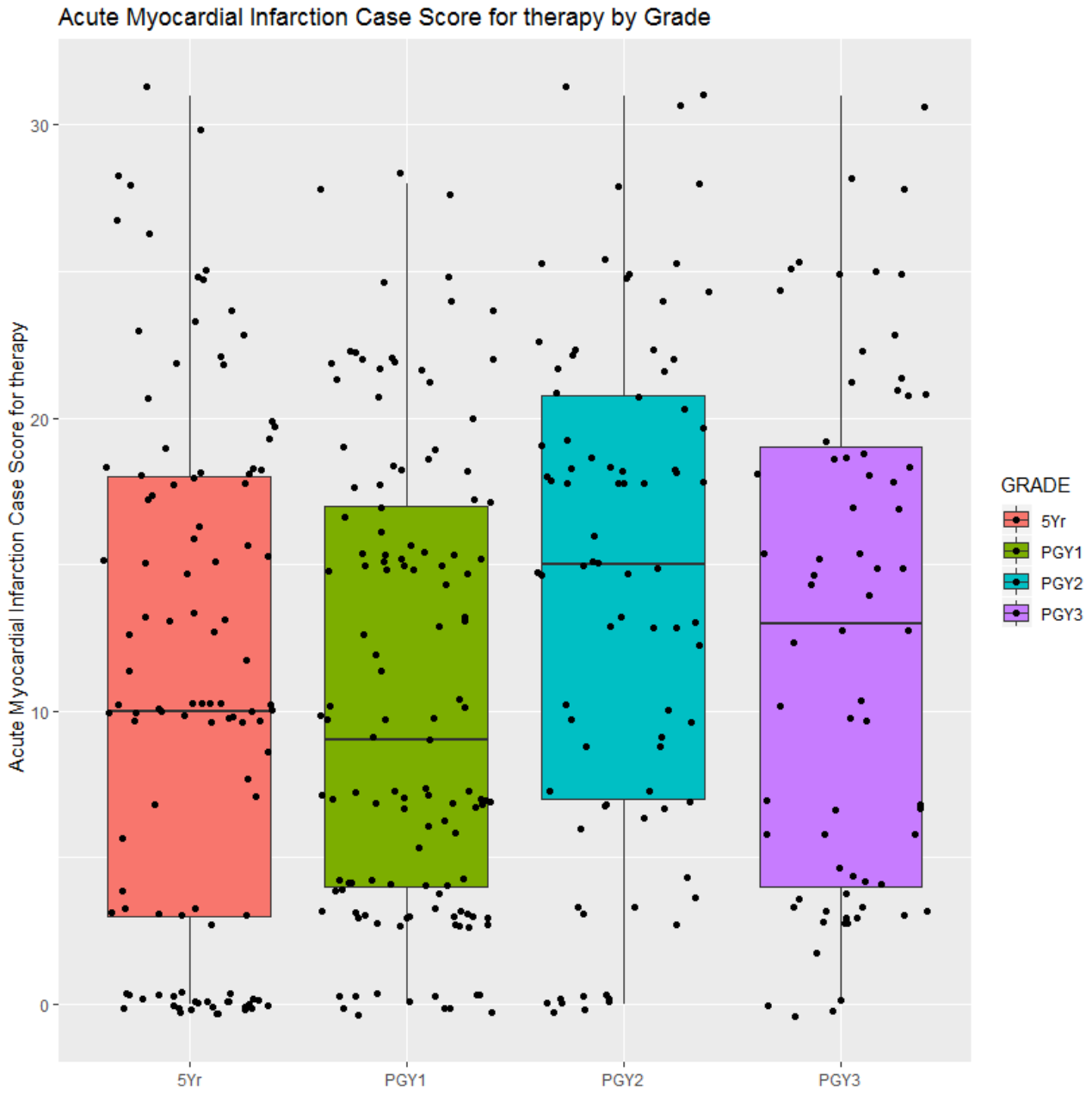
XSL•FO
RenderX

Figure 5. Subscore for diagnosis of students over training years.



XSL•FO
RenderX

Figure 6. Subscore for treatment of students over training years.



Discussion

Principal Findings

Before applying an assessment tool for use with medical students, we must obtain evidence for the instrument's reliability and validity [26-28]. Providing evidence of the validity of CBCRT will help the test management organization understand the effectiveness of the test from a broad and comprehensive perspective, clarify the aspects that the CBCRT can and cannot measure, and hence, allow for its continuous improvement. This is the goal of our current study. Our hypothesis was supported by the results obtained; specifically, senior students displayed higher testing scores than junior students (Table 3, Figure 1). In other words, the CMU-CBCRT is able to determine CR skills over different levels of medical education, especially in the early stage of the students' medical careers.

Looking specifically into the 4 categories of skills that we tested, we found that the most significant differences were revealed in the information collection, diagnosis, and treatment scores among junior and senior medical students. This was as predicted. With years of training, their experience and ability to clinically reason are improving, and as a result, they performed better on the information collection, diagnosis, and treatment, as well as the total CBCRT score. This further suggests that the CMU-CBCRT can determine the CR skills of students at varying levels.

We also carefully studied and analyzed why there were no significant differences in treatment error scores among the 4 training groups. For a simulated case of myocardial infarction, we can observe from the test result the challenge faced by participants who have never experienced this form of examination before. When the passing score was set at 60%,

XSL•FO
RenderX

the average score in this case (59.01 [SD 16.68]) did not pass. The choice of wrong treatment is a negative item in the scoring system, so the item writing expert is very cautious in formulating the scoring standard. Only behavior that caused extreme consequences resulted in points being deducted, and the weight was set at a very low level (ie, –1%). In this test, we observed that treatment error behavior happened more with junior students than senior students, although without statistical significance.

In the absence of an available gold standard for measuring CR, evidence for construct validity is sought after in this area of research. This is an ongoing process, in which the skill measured by the assessment tool is linked to some other attribute by a hypothesis or construct. With the development of validity theory, the validity concept has a new connotation and forms a method based on multilevel evidence [22]. Validity is no longer an attribute of the measurement tool itself but rather the extent to which the evidence collected supports the interpretation, inference, and decision making of the test score [27,29].

With the positive evidence presented, we should still be aware that validity verification is a dynamic process [27] and no education instrument is 100% effective [27]. Even if the evidence indicates that the validity of a course test is significant, the validity study must continue along with the development of the CBCRT system. There are still many problems to be solved, such as the setting of the evidence framework for the specific test validity, determination of the validity criteria, feasibility of the evidence collection method, and quantification of evidence data. This will require in-depth discussion by future researchers. We aim for constant examination of these issues in the process of developing a reliable and valid CR training model. In the future, we would include more simulated cases with a wide range of case difficulties and distribute CMU-CBCRT to more students to increase sample size. We would then carefully collect data on student performance and feedback. We also plan to add graphics and animation to enhance the interface design.

Limitations

However, there were some limitations in our study to its generalizability. First, the respondents of the research were from only 5 medical institutions in China. Second, the findings of our study were limited by the representativeness and scale of the study population.

Conclusions

We provided evidence of construct validity of the CMU-CBCRT. It is able to determine CR skills over different levels of medical education, especially in the early stage of the students' medical careers.

Acknowledgments

Authors thank the National Medical Examination Center of China, which sponsored this research through the funded project Sustainable Development of Computer-based Case Simulations Examination System. We thank educators and collaborators from China Medical University, Fudan University School of Medicine, Sun Yat-sen University School of Medicine, Xuzhou Medical University, and Binzhou Medical College who provided of testing places and expertise that greatly assisted the research. Special thanks to Ms Kritika Taparia for proofreading the manuscript and provide valuable comments.

Authors' Contributions

TZ, XG, and BQ collected the data. TZ, BZ, and BS analyzed the data and wrote the first draft of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Scoring sheet and scoring rubrics.
[XLSX File (Microsoft Excel File), 11 KB - games_v9i4e17670_app1.xlsx]

Multimedia Appendix 2
China Medical University–computer-based clinical reasoning testing data.
[XLSX File (Microsoft Excel File), 113 KB - games_v9i4e17670_app2.xlsx]

References

1. Bendapudi NM, Berry LL, Frey KA, Parish JT, Rayburn WL. Patients' perspectives on ideal physician behaviors. Mayo Clin Proc 2006 Mar;81(3):338-344. [doi: 10.4065/81.3.338] [Medline: 16529138]
2. Ellaway R, Masters K. AMEE Guide 32: e-Learning in medical education Part 1: learning, teaching and assessment. Med Teach 2008 Jun 03;30(5):455-473. [doi: 10.1080/01421590802108331] [Medline: 18576185]
3. Schreuder HWR, Oei G, Maas M, Borleffs JCC, Schijven MP. Implementation of simulation in surgical practice: minimally invasive surgery has taken the lead: the Dutch experience. Med Teach 2011;33(2):105-115. [doi: 10.3109/0142159X.2011.550967] [Medline: 21275542]
4. Bliss M, Osler W. A Life in Medicine. Oxford: Oxford University Press; 1999.

XSL•FO
RenderX

5. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. Med Educ 2009 Apr;43(4):303-311. [doi: 10.1111/j.1365-2923.2008.03286.x] [Medline: 19335571]
6. Kassirer JP. Teaching clinical reasoning: case-based and coached. Acad Med 2010 Jul;85(7):1118-1124. [doi: 10.1097/acm.0b013e3181d5dd0d] [Medline: 20603909]
7. Sevdalis N, McCulloch P. Teaching evidence-based decision-making. Surg Clin North Am 2006 Feb;86(1):59-70. [doi: 10.1016/j.suc.2005.10.008] [Medline: 16442420]
8. Sokol RG, Slawson DC, Shaughnessy AF. Teaching evidence-based medicine application: transformative concepts of information mastery that foster evidence-informed decision-making. BMJ Evid Based Med 2019 Aug;24(4):149-154. [doi: 10.1136/bmjebm-2018-111142] [Medline: 31208984]
9. Mottola MF, Bagnall KM, Belcastro AN. Effects of strenuous maternal exercise on fetal organ weights and skeletal muscle development in rats. J Dev Physiol 1989 Feb;11(2):111-115. [Medline: 2778291]
10. Peile E. Teaching balanced clinical decision-making in primary care: evidence-based and values-based approaches used in conjunction. Educ Prim Care 2014 Mar;25(2):67-70. [doi: 10.1080/14739879.2014.11494248] [Medline: 24593967]
11. Scott PJ, Altenburger PA, Kean J. A collaborative teaching strategy for enhancing learning of evidence-based clinical decision-making. J Allied Health 2011;40(3):120-127. [Medline: 21927777]
12. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. Med Educ 2015 Oct 18;49(10):961-973. [doi: 10.1111/medu.12775] [Medline: 26383068]
13. Furze J, Kenyon LK, Jensen GM. Connecting classroom, clinic, and context: clinical reasoning strategies for clinical instructors and academic faculty. Pediatr Phys Ther 2015;27(4):368-375. [doi: 10.1097/PEP.0000000000000185] [Medline: 26397080]
14. Windish DM. Teaching medical students clinical reasoning skills. Acad Med 2000 Jan;75(1):90. [doi: 10.1097/00001888-200001000-00022] [Medline: 10667884]
15. Custers EJFM. Theories of truth and teaching clinical reasoning and problem solving. Adv Health Sci Educ Theory Pract 2019 Oct;24(4):839-848 [FREE Full text] [doi: 10.1007/s10459-018-09871-4] [Medline: 30671703]
16. Levin M, Cennimo D, Chen S, Lamba S. Teaching clinical reasoning to medical students: a case-based illness script worksheet approach. MedEdPORTAL 2016 Aug 26;12:10445 [FREE Full text] [doi: 10.15766/mep_2374-8265.10445] [Medline: 31008223]
17. van Merrienboer JJG, Sweller J. Cognitive load theory in health professional education: design principles and strategies. Med Educ 2010 Jan;44(1):85-93. [doi: 10.1111/j.1365-2923.2009.03498.x] [Medline: 20078759]
18. Rencic J. Twelve tips for teaching expertise in clinical reasoning. Med Teach 2011;33(11):887-892. [doi: 10.3109/0142159X.2011.558142] [Medline: 21711217]
19. Forrest JL, Miller SA, Newman MG. Teaching evidence-based decision-making versus experience-based dentistry. Alpha Omegan 2004 Jul;97(2):35-41. [Medline: 18712148]
20. Stone J, Moskowitz GB. Non-conscious bias in medical decision making: what can be done to reduce it? Med Educ 2011 Aug;45(8):768-776. [doi: 10.1111/j.1365-2923.2011.04026.x] [Medline: 21752073]
21. Toro-Huamanchumo CJ, Arce-Villalobos LR. The clinical eye: a need to improve the teaching of semiology in undergraduate medical education. Int J Med Students 2014 Oct 25;2(3):144. [doi: 10.5195/ijms.2014.186]
22. Ward RC, Muckle TJ, Kremer MJ, Krogh MA. Computer-based case simulations for assessment in health care: a literature review of validity evidence. Eval Health Prof 2017 Jan 01:163278717718609. [doi: 10.1177/0163278717718609] [Medline: 28727944]
23. Allard M, Lafleur A, Richard E, Lebouthillier A, Vailles C. How medical students edited an OSCE Study Guide and why should you? Int J Med Students 2018 May 09;6(2):78-82. [doi: 10.5195/ijms.2018.248]
24. Nekkanti S, Manjunath S, Mahtani A, Meka A, Rao T. A survey based feedback analysis of the current medical teaching methodology and trends in medical research practice in a South Indian medical institute. Int J Med Students 2018 Apr 30;6(1):6-17. [doi: 10.5195/ijms.2018.3]
25. Liaw SY, Rashasegaran A, Wong LF, Deneen CC, Cooper S, Levett-Jones T, et al. Development and psychometric testing of a Clinical Reasoning Evaluation Simulation Tool (CREST) for assessing nursing students' abilities to recognize and respond to clinical deterioration. Nurse Educ Today 2018 Mar;62:74-79. [doi: 10.1016/j.nedt.2017.12.009] [Medline: 29306102]
26. Feldman LS, Sherman V, Fried GM. Using simulators to assess laparoscopic competence: ready for widespread use? Surgery 2004 Jan;135(1):28-42. [doi: 10.1016/s0039-6060(03)00155-7]
27. Morris SB, Kwaske IH, Daisley RR. The validity of individual psychological assessments. Ind Organ Psychol 2015 Jan 07;4(3):322-326. [doi: 10.1111/j.1754-9434.2011.01347.x]
28. Yu CHY, Straus S, Brydges R. The ABCs of DKA: development and validation of a computer-based simulator and scoring system. J Gen Intern Med 2015 Sep;30(9):1319-1332 [FREE Full text] [doi: 10.1007/s11606-015-3273-y] [Medline: 26173518]
29. Kane MT. Current concerns in validity theory. J Educational Measurement 2001 Dec;38(4):319-342. [doi: 10.1111/j.1745-3984.2001.tb01130.x]

XSL•FO
RenderX

Abbreviations

ANOVA: analysis of variance
CBCRT: computer-based clinical reasoning testing
CMU: China Medical University
CR: clinical reasoning
PBL: problem-based learning
PGY: postgraduate year

Edited by N Zary; submitted 02.01.20; peer-reviewed by M Gaman, R Miclaus; comments to author 26.10.20; revised version received 19.03.21; accepted 19.06.21; published 09.11.21.

Please cite as:
Zuo T, Sun B, Guan X, Zheng B, Qu B
Evidence of Construct Validity of Computer-Based Tests for Clinical Reasoning: Instrument Validation Study
JMIR Serious Games 2021;9(4):e17670
URL: https://games.jmir.org/2021/4/e17670
doi:10.2196/17670
PMID:34751658

©Tianming Zuo, Baozhi Sun, Xu Guan, Bin Zheng, Bo Qu. Originally published in JMIR Serious Games (https://games.jmir.org), 09.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using a 360° Virtual Reality or 2D Video to Learn History Taking and Physical Examination Skills for Undergraduate Medical Students: Pilot Randomized Controlled Trial

Yi-Ping Chao[1,2], PhD; Hai-Hua Chuang[3,4], MD; Li-Jen Hsin[4,5], MD; Chung-Jan Kang[4,5], MD, PhD; Tuan-Jen Fang[4,5], MD; Hsueh-Yu Li[4,5], MD; Chung-Guei Huang[6,7], PhD; Terry B J Kuo[8], MD, PhD; Cheryl C H Yang[8], PhD; Hsin-Yih Shyu[4,9], PhD; Shu-Ling Wang[4,10], PhD; Liang-Yu Shyu[11], PhD; Li-Ang Lee[4,5,8], MD, MSc

[1]Department of Computer Science and Information Engineering, Graduate Institute of Biomedical Engineering, Chang Gung University, Taoyuan, Taiwan

[2]Department of Neurology, Chang Gung Memorial Hospital, Linkou Main Branch, Taoyuan, Taiwan

[3]Department of Family Medicine, Chang Gung Memorial Hospital, Taipei Branch & Linkou Main Branch, Taoyuan, Taiwan

[4]College of Medicine, Chang Gung University, Taoyuan, Taiwan

[5]Department of Otorhinolaryngology-Head and Neck Surgery, Chang Gung Memorial Hospital, Linkou Main Branch, Taoyuan, Taiwan

[6]Department of Laboratory Medicine, Chang Gung Memorial Hospital, Linkou Main Branch, Taoyuan, Taiwan

[7]Department of Medical Biotechnology and Laboratory Science, Graduate Institute of Biomedical Sciences, Chang Gung University, Taoyuan, Taiwan

[8]Institute of Brain Science, National Yang Ming Chiao Tung University, Taipei, Taiwan

[9]Department of Educational Technology, Tamkang University, New Taipei, Taiwan

[10]Center of Teacher Education, National Taiwan University of Science and Technology, Taipei, Taiwan

[11]Department of Biomedical Engineering, Chung Yuan Christian University, Taoyuan, Taiwan

Corresponding Author:
Li-Ang Lee, MD, MSc
Department of Otorhinolaryngology-Head and Neck Surgery
Chang Gung Memorial Hospital
Linkou Main Branch
No 5, Fu-Hsing Street
Guishan District
Taoyuan, 33305
Taiwan
Phone: 886 3 328 1200 ext 3972
Fax: 886 3 397 9361
Email: 5738@cgmh.org.tw

Abstract

Background: Learning through a 360° virtual reality (VR) or 2D video represents an alternative way to learn a complex medical education task. However, there is currently no consensus on how best to assess the effects of different learning materials on cognitive load estimates, heart rate variability (HRV), outcomes, and experience in learning history taking and physical examination (H&P) skills.

Objective: The aim of this study was to investigate how learning materials (ie, VR or 2D video) impact learning outcomes and experience through changes in cognitive load estimates and HRV for learning H&P skills.

Methods: This pilot system–design study included 32 undergraduate medical students at an academic teaching hospital. The students were randomly assigned, with a 1:1 allocation, to a 360° VR video group or a 2D video group, matched by age, sex, and cognitive style. The contents of both videos were different with regard to visual angle and self-determination. Learning outcomes were evaluated using the Milestone reporting form. Subjective and objective cognitive loads were estimated using the Paas Cognitive Load Scale, the National Aeronautics and Space Administration Task Load Index, and secondary-task reaction time. Cardiac autonomic function was assessed using HRV measurements. Learning experience was assessed using the AttrakDiff2 questionnaire and qualitative feedback. Statistical significance was accepted at a two-sided *P* value of <.01.

XSL•FO
RenderX

Results: All 32 participants received the intended intervention. The sample consisted of 20 (63%) males and 12 (38%) females, with a median age of 24 (IQR 23-25) years. The 360° VR video group seemed to have a higher Milestone level than the 2D video group (*P*=.04). The reaction time at the 10th minute in the 360° VR video group was significantly higher than that in the 2D video group (*P*<.001). Multiple logistic regression models of the overall cohort showed that the 360° VR video module was independently and positively associated with a reaction time at the 10th minute of ≥3.6 seconds (exp B=18.8, 95% CI 3.2-110.8; *P*=.001) and a Milestone level of ≥3 (exp B=15.0, 95% CI 2.3-99.6; *P*=.005). However, a reaction time at the 10th minute of ≥3.6 seconds was not related to a Milestone level of ≥3. A low-frequency to high-frequency ratio between the 5th and 10th minute of ≥1.43 seemed to be inversely associated with a hedonic stimulation score of ≥2.0 (exp B=0.14, 95% CI 0.03-0.68; *P*=.015) after adjusting for video module. The main qualitative feedback indicated that the 360° VR video module was fun but caused mild dizziness, whereas the 2D video module was easy to follow but tedious.

Conclusions: Our preliminary results showed that 360° VR video learning may be associated with a better Milestone level than 2D video learning, and that this did not seem to be related to cognitive load estimates or HRV indexes in the novice learners. Of note, an increase in sympathovagal balance may have been associated with a lower hedonic stimulation score, which may have met the learners' needs and prompted learning through the different video modules.

Trial Registration: ClinicalTrials.gov NCT03501641; https://clinicaltrials.gov/ct2/show/NCT03501641

(JMIR Serious Games 2021;9(4):e13124) doi:10.2196/13124

KEYWORDS

cognitive load; heart rate variability; video learning; learning outcome; secondary-task reaction time; virtual reality

Introduction

Competence-Based Medical Education Needs a Multifaceted Assessment System

From premedical students to practicing physicians, competency-based medical education (CBME) has been applied in the design, implementation, and evaluation of medical education programs according to competency-based outcomes [1]. CBME for undergraduate medical students can help to improve task-specific confidence and test performance after the course has been completed, resulting in better performance and patient care before residency training [2]. CBME has been widely employed to promote greater learner-centeredness and curricular outcomes since 2010 [3]. Furthermore, CBME must use multiple assessment tools that meet minimum requirements for quality. Therefore, a robust and multifaceted assessment system, including quantitative and qualitative methods, is essential for evaluating the progress of learners [4].

Simulation Provides Situated Learning and Assessment in a Safe Environment

CBME is used to improve graduates' competency levels, ensuring they are skillful and qualified in all critical areas of their occupation. To this end, CBME emphasizes the use and importance of simulation-based training, which considers patient safety and for which real-life opportunities are limited [5]. This consideration is crucial for medical learners and teachers during the COVID-19 pandemic. History taking and physical examination (H&P) is a principal competency, incorporating knowledge, skills, and behavior to initially approach a patient. Therefore, H&P is a key performance level of otorhinolaryngology–head and neck surgery (ORL-HNS) [6]. To enhance the development of H&P skills, commonly used methods to assess this competency include in-training examinations such as the Milestone [7], Mini Clinical Evaluation Exercise [8], and oral examinations of clinical practice [9]. Recently, simulations such as part-time trainers, integrated simulators, virtual reality (VR), and wearable devices have become increasingly popular for CBME [10]. Simulation-based training, such as VR, has been shown to improve health professionals' knowledge and skills outcomes and be an integrative step toward supervised clinical practice [11]. In particular, VR allowed students to practice rounding skills, facilitate their education during the COVID-19 pandemic, and supplement their in-person clerkship education [12].

360° VR Video Can Provide High Authenticity and Fidelity to Encourage Learners

VR consists of a computer-generated 3D simulation in which the user both explores and manipulates the contents of the environment to learn and assess clinical knowledge and skills [13]. VR provides experiential learning and provides standardized, controlled exposure to situated events, patient-caregiver communication, and teamwork. The use of VR has been shown to be highly acceptable by learners in a wide range of health care settings [14] and to play an essential role in improving performance [15]. As a subtype of image-based VR, 360° VR video represents an immersive 3D medium featuring authenticity and fidelity using a VR head-mounted display. The use of 360° VR video opens up many possibilities in many domains of medical education [16], such as an independent teaching aid or an adjunct to traditional face-to-face teaching [17]. The application of 360° VR video has enhanced the effectiveness of medical education and training, raised the level of diagnosis and treatment, improved the doctor-patient relationship, and boosted medical execution efficiency [18]. Using 360° VR videos to facilitate the acquisition of new clinical skills has been suggested to be a valuable step in developing a clinical teaching curriculum [19].

Incorporating Instructional Design Practices That Address the Elements of Cognitive Load Theory Is Important to Improve Learning Outcomes

However, transferring practical skills to a clinical setting using a criterion-based training program with VR simulators is

XSL•FO
RenderX

difficult [20]. In addition to the poor mechanical performance of the simulated haptic feedback, complex tasks such as H&P and surgical procedures can induce excessive cognitive load during simulation training, which can harm learning, especially for novices [21]. Since the learner's cognitive capacity during learning is limited, improper instructional design may waste precious cognitive resources, impair the essential processing, and cause generative underutilization during learning [22]. Therefore, cognitive overload negatively impacts learning outcomes in simulation training [23,24]. For example, the increased cognitive load was significantly associated with the declined correct identification of a trained murmur during simulation training [23]. A study of laparoscopic training [24] reported that immersive VR simulation training resulted in the cognitive overload that impeded actual learning and skills acquisition compared with conventional VR simulation training. However, structured and distributed VR simulation practices may induce a lower cognitive load when the learning situation is increased in complexity [21]. Recently, VR-based simulation training under cognitive load control has improved performance under similar conditions to an actual surgical task [25]. Therefore, estimating cognitive load during 360° VR and 2D video learning is essential for practical instruction and reform.

Estimating Cognitive Load During 360° VR Video Learning Is a Challenge

Estimations of cognitive load and specific load types are still challenging based on current cognitive load theory [26,27]. For example, subjective cognitive load questionnaires, such as the Paas Cognitive Load Scale (Paas-CLS) [28] and the National Aeronautics and Space Administration Task Load Index (NASA-TLX) [29] have been shown to be good tools to measure intrinsic load, but not extraneous and germane loads [30]. However, a lowering of intrinsic cognitive load can reduce total cognitive load, thus releasing working memory capacity [31]. Furthermore, instructional techniques for reducing cognitive load have been shown to improve learning [32]. In addition to estimating cognitive load, the NASA-TLX has been successfully applied to determine user acceptance when evaluating innovative applications [33]. Nevertheless, a negative carryover effect due to failure in a VR simulation training program could affect the subjective cognitive load estimations [34]. Therefore, objective estimates of cognitive load, including secondary-task performance [35], during the training program have been applied in many medical education studies [34-36]. Secondary-task performance is a task that assesses participants' attention and is limited by the storage capacity of visual short-term memory [37]. Although secondary-task performance is a cognitive function test, it frequently deteriorates from baseline to dual-task among novices and is particularly useful for tracking changes in cognitive load in the early phases of simulation-based skills [35].

Autonomic Function Can Be Altered During VR Immersion

VR immersion, especially containing stressful content, can evoke acute stress reactions accompanied by autonomic dysfunction [38]. Heart rate variability (HRV) is a sensitive indicator of cardiac autonomic modulation [39] and responds to any psychophysical changes immediately. Several physiological or psychological changes, such as stress [40] and anxiety [41], can change HRV. Reduced HRV indexes might reflect loss of cognitive efficiency [36] and might predict increased cognitive load [42] and poor cognitive performance [43]. Nevertheless, immersive VR using the first-person perspective can induce body ownership illusion in an uncomfortable posture, reduced HRV indexes, and more mistakes in a cognitive task [44]. Despite it being challenging to assess associations between HRV and subjective cognitive load due to sizeable interindividual variability [45], measuring HRV during a learning task may help evaluate whether learning materials impact learning outcomes through cardiac changes autonomic function.

Aims and Hypotheses of the Study

This study aimed (1) to evaluate differences between two learning materials (ie, 360° VR video and 2D video) in the subjective and objective cognitive loads, autonomic function, outcome, and learning experience while students were learning H&P skills and (2) to study how learning materials (ie, 360° VR or 2D video) impact learning outcomes through the changes in cognitive load and HRV. The initial hypotheses were that (1) subjective and objective cognitive loads and autonomic function would be changed while using different learning materials and (2) learners with higher learning outcomes and experience would have different cognitive load and autonomic function during the learning tasks than those with lower learning outcomes and experience.

Therefore, we used cognitive questionnaires immediately after completing a video learning module to assess subjective cognitive load. We also used secondary-task reaction time and HRV to objectively estimate the dynamic changes of cognitive load and cardiac autonomic function during a video module for providing ongoing feedback to improve video learning. Subsequently, we used the Milestone report form to evaluate the real-patient H&P. The Milestones have been designed according to the grounding principles of CBME and have become a significant formative component of the current accreditation model of graduate medical education [46]. Although the Milestones were not designed to assess undergraduate medical students, we considered that some descriptors and targets of the Milestones were suitable for evaluating learner performance of H&P as an undergraduate medical student moves from entry into clerkship through graduation. Furthermore, we applied the AttrakDiff2 questionnaire to assess the acceptance of technical innovations [47]. The AttrakDiff2 questionnaire has been used to investigate learners' experience of many educational innovations, such as mobile e-learning [48] and augmented reality [49]. Finally, we used anonymous qualitative feedback to determine the learning experience. We considered that the learning outcomes and the learning experience could meet learners' needs and prompt their use of learning through different video modules.

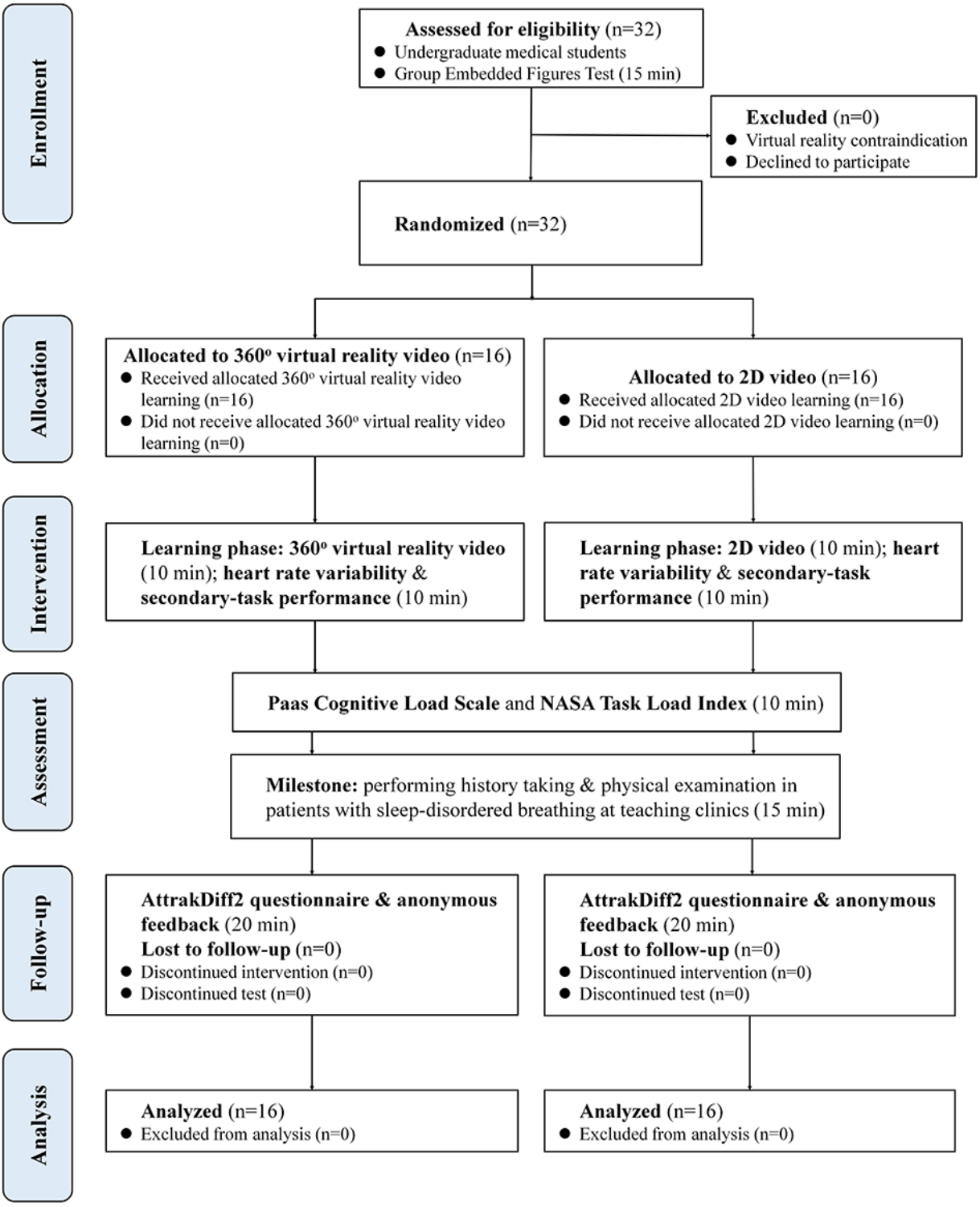
XSL•FO
RenderX

Methods

Study Design

We conducted a prospective, randomized controlled, pilot system–design study from June 1 to October 30, 2018, at an academic teaching hospital (Department of ORL-HNS, Linkou Chang Gung Memorial Hospital, Taoyuan, Taiwan). The Institutional Review Board of the Chang Gung Medical Foundation approved this study (No. 201601821B0), and we conducted all procedures in compliance with the Declaration of Helsinki 1975. We informed the participants about the aims of the study and then obtained written informed consent from them. We registered the entire study proposal at ClinicalTrials.gov (NCT03501641). Figure 1 shows the study flowchart following the CONSORT 2010 guidelines [50].

Figure 1. The CONSORT flow diagram. NASA: National Aeronautics and Space Administration.



XSL•FO
RenderX

Setting

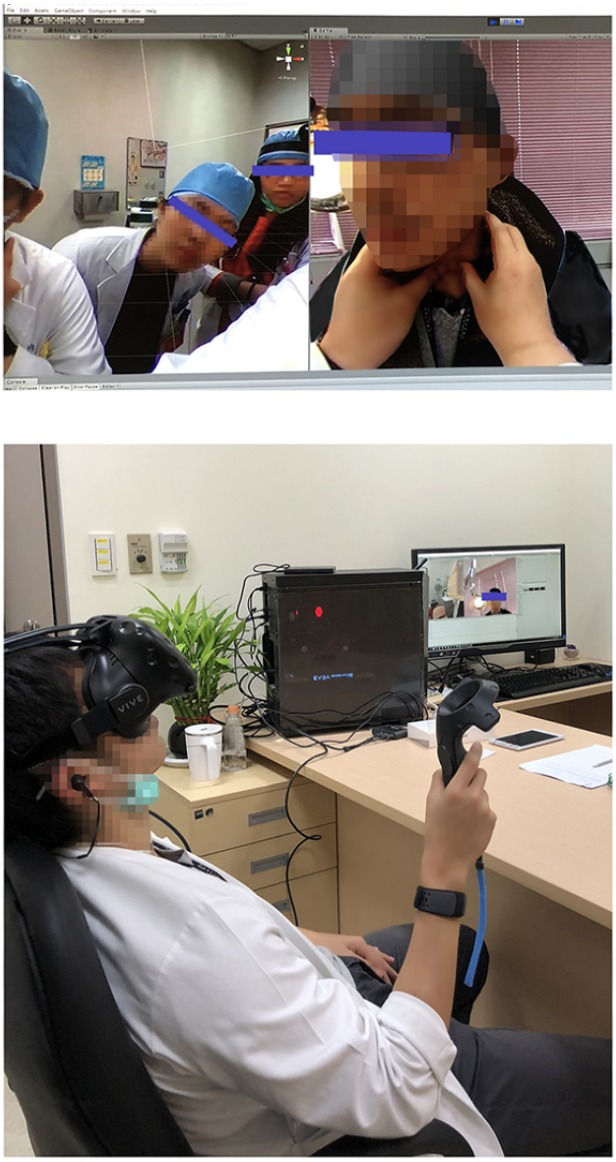
Overview

We used analysis, design, development, implementation, and evaluation models [51] to design an effective instruction module for H&P, including essential knowledge and competence according to the guidelines of the American Board of Otolaryngology [7]. We used different working samples [52], including instructions for how to formulate problems, solution steps, and final solutions, to demonstrate, step by step, how to perform an H&P task in an outpatient setting. We also used self-explanation prompts [53] to encourage the learner to recognize links between the knowledge and skills they learned. We recorded a 10-minute 360° video (4K resolution, 30 frames/s) with in-camera stitching, capturing 360° audio, and spherical stabilization using a 360° camera (Garmin VIRB 360; Garmin Ltd). We constructed the contents and scenario of this video according to a real clinical setting. The first portion of the video demonstrated skills of history taking under normal conditions, and the second portion demonstrated skills of how to quickly perform a physical examination (Table 1). Subsequently, we produced two videos with different visual angles (ie, 360° and 120°) using PowerDirector software (version 16; CyberLink Corp). Two senior investigators evaluated the videos and validated the learning materials. We then developed courseware with the same user interface for the 360° VR and 2D videos using Unity Editor (version 2017.3.1; Unity Technologies).

The 360° VR Video Module

This module was developed to arbitrarily review the immersive 3D 360° VR video through a head-mounted display (Figure 2). The users were immersed in the 360° experience to learn the H&P skills.

Figure 2. Example of 360° virtual reality video learning. Screenshot of the 360° virtual reality video (upper) demonstrating the learners watching a highly immersive 360° video. They arbitrarily changed their field of view to watch the skills of history taking and physical examination from a first-person perspective (lower).



XSL•FO
RenderX

The 2D Video Module

The 2D video was played in a fixed 120° focused field of view through the same head-mounted display (Figure 3). The users reviewed the instructional video as in a theater environment.

Figure 3. Example of 2D video learning. Screenshot of the 2D video (upper) showing the learners watching a 2D video in a theater environment. They watched the skills of history taking and physical examination in a fixed focus of view from a third-person perspective (lower).

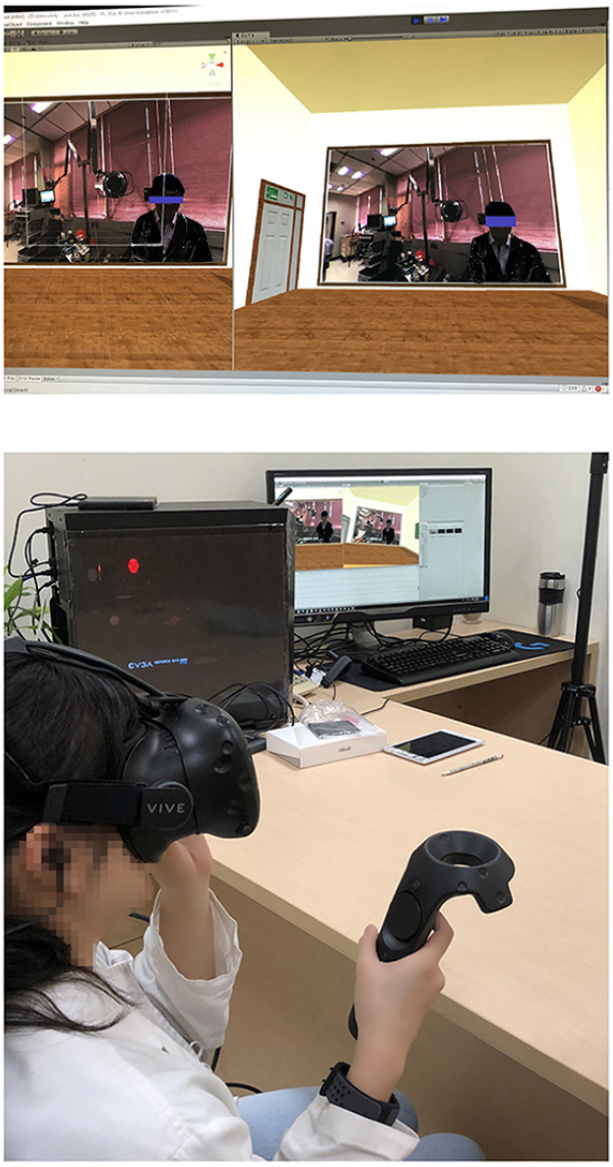


Table 1. Summary of the models of 360° virtual reality video and 2D video.

Feature	360° virtual reality video module	2D video module
Head-mounted display	Yes	Yes
Part I: history taking	Statically filmed	Statically filmed
Part II: physical examination	Dynamically filmed	Dynamically filmed
Visual angle	360°	120°
Immersion	Yes	No
Perspective	First person	Third person

Selection of Participants

A convenience sampling approach was used, and 32 consecutive volunteers were recruited during the study period. The inclusion criteria were as follows: (1) aged >20 years and (2) novices in ORL-HNS (ie, undergraduate medical students). The exclusion criteria were as follows: (1) contraindications for using VR, such as pregnancy, hypertension, motion sickness, inner ear infections, claustrophobia, recent surgery, pre-existing binocular vision abnormalities, heart disorders, or epilepsy, and (2)

XSL•FO
RenderX

declining to participate. All of the volunteers had at least a basic level of computer literacy and could use VR headsets and controllers after instruction. We used the 25-item Group Embedded Figures Test (GEFT) (score range 0-18) to assess the participants' cognitive style [54]. The GEFT has been shown to have high reliability in medical education [55], and we have previously validated its effectiveness of classifying learning preference in millennial undergraduate medical students [56]. Field-independent learners prefer and have better performance in computer-assisted learning. We stratified the students into two subgroups: "field-dependent" (GEFT score ≤12) and "field-independent" (GEFT score >12) [57].

Randomization and Blinding

The participants were blinded to the purpose of the study during recruitment to minimize preparation bias. After the participants had provided consent and completed the GEFT, we randomly assigned them, with a 1:1 allocation, to the 360° VR video group and the 2D video group, matched by age, sex, and cognitive style (Figure 1). The Random Number Generators tool in SPSS software (version 24; IBM Corp) was used to create a list of random numbers for allocating the students, who were stratified by center with a 1:1 allocation using a fixed block size of 8 in both parallel subgroups. We concealed the allocation sequence before implementing the video module, and the module adhered to our computer-generated randomization protocol.

Intervention

After randomization, the participants reviewed their allocated video through a head-mounted display in the same space for 10 minutes. To reduce the effect of the head-mounted display on learning experience, both groups used the same VR device (VIVE VR headset; HTC Corp). We explained the functionality of the VR device to the participants before the intervention. In the 360° VR video group, the learners arbitrarily reviewed the instructor's demonstrations and responses from standard patients and other medical staff from a first-person perspective in an immersive 360° environment (Figure 2). In the 2D video group, the learners simply watched the instructor's demonstrations from a third-person perspective in a theater environment (Figure 3). During the learning course, the participants watched the movie by themselves at a time of their choosing.

Methods of Measurement

Overview

We used five different face-to-face assessments, including the Paas-CLS and the NASA-TLX for subjectively estimating the cognitive load, the Milestone for assessing the learning outcomes, and the AttrakDiff2 questionnaire and anonymous qualitative feedback for determining the learning experience. There was one objective estimate of cognitive load using the secondary-task performance and one measurement of cardiac autonomic function using HRV monitoring.

Paas Cognitive Load Scale

We used the Paas-CLS [28] to estimate the total cognitive load of the learning task immediately after the intervention. The Paas-CLS questionnaire is a single-item measure to rate the perceived intensity of mental effort along a 9-point scale, ranging from 1 (very, very low mental effort) to 9 (very, very high mental effort). The Paas-CLS questionnaire has good reliability (Cronbach α=.82-.90) in instructional research [58].

NASA Task Load Index

The NASA-TLX questionnaire is a subjective assessment of cognitive load [29]. This instrument consists of six subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration. Participants rate the level of each dimension by making a mark on a visual analog scale (range 0-20) immediately after the intervention. The NASA-TLX questionnaire has good reliability (Cronbach α≥.80) to assess cognitive load [59].

Secondary-Task Performance

Secondary-task performance has been shown to be sensitive in estimating intrinsic cognitive load among novices engaged in simulation-based learning [35]. As the primary task in this study, the participants reviewed the video for 10 minutes. To ensure the use of similar perceptual-cognitive resources for the primary task [60], the participants were asked to respond to a visual cue by pressing a button on a controller as soon as possible. The visual cue appeared in their field of view and lasted for 10 seconds. For exploring the importance of monitoring secondary-task reaction time, we measured the reaction time, defined as the time from visual cue presentation to when the button was pressed, at 0 minutes, 5 minutes, and 10 minutes during the reviewing period. When the participants missed a response, we considered the reaction time to be "11 seconds" to avoid gaps in the performance results and inaccuracies in the estimates of the reaction time.

Heart Rate Variability

HRV has been shown to be an objective estimation of learners' cognitive load in a learning environment [61,62]. For controlling preinterventional stress, the participants sat quietly for 20 minutes. We recorded 5-minute electrocardiogram signals of a single lead (lead I) using a Holter-like NeXus-4 amplifier and recording system (Mind Media BV) when the participant wore a head-mounted display and breathed normally as baseline data. During the 10-minute intervention, we recorded heart rates simultaneously and continuously. Electrocardiogram data were acquired with a sample rate of 1024 Hz, and the raw data were saved. The power spectrum was quantified by a fast Fourier transform [39]. In our preliminary results, the energy-frequency-time distributions of the electrocardiogram data were nonlinear and nonstationary. Therefore, we used the empirical mode decomposition (EMD) method to decompose complicated data into a finite number of intrinsic mode functions that admitted well-behaved Hilbert transforms [63]. The EMD method is an adaptive preprocessing technique for overcoming the limitations of HRV spectral analysis when assessing nonlinear and nonstationary system data [64]. We performed HRV analysis using custom-developed MATLAB (version 7; The MathWorks, Inc) codes that allowed us to determine HRV parameters from sequences of 5-minute consecutive epochs of electrocardiogram signals. The requirements for the quality of HRV analysis were based on the European Society of Cardiology and the North American Society of Pacing and

XSL•FO
RenderX

Electrophysiology guidelines [65]. For the analysis, sequences of R wave to R wave (RR) intervals were selected without artifacts, ventricular excitations, or supraventricular excitations [66]. We then calculated the RR interval, the SD of normal-to-normal RR intervals (SDNN), and the root mean square of successive heartbeat interval difference (RMSSD) using time-domain analysis. HRV in the frequency domain was described using spectral power values in the low-frequency (LF) band (0.04-0.15 Hz), the high-frequency (HF) band (0.15-0.40 Hz), and the LF/HF ratio. To reflect the different instructional content, we analyzed three time intervals: baseline, 0 to 5 minutes, and 5 to 10 minutes. These HRV indexes were chosen because low SDNN was associated with higher intrinsic and germane cognitive loads [67] and poor performance [68], low RMSSD was related to higher intrinsic cognitive load [67] and worse performance on executive tasks [69], and low LF/HF ratio was associated with high cognitive load [70]. The HRV variables were measured before and after the intervention.

Milestones

The Milestones have been used to assess the development of resident physicians in key dimensions of the elements of physician competency in otolaryngology since 2004 [7]. In this study, we prospectively recruited participants to perform H&P in real patients with sleep-disordered breathing (SDB) at our teaching clinics. For evaluating learner performance of H&P as an undergraduate medical student moves from entry into clerkship through graduation, we selected the level that best described that learner's performance with the Milestone for care for patients with SDB [6]. We used a brief 5-level Milestone report form to rate learner performance of H&P, including (1) obtaining general H&P, (2) recognizing signs and symptoms of SDB and the differences between children and adults, (3) performing detailed examinations with evaluations of upper airway anatomy, (4) interpreting the examinations and advanced diagnostic testing, and (5) teaching-focused H&P.

AttrakDiff2 Questionnaire

The AttrakDiff2 questionnaire was developed to reliably evaluate the acceptance of technical innovations [47]. It assesses qualities of pragmatism, hedonic stimulation, hedonic identification, and attractiveness using 28 questions. The participants were asked to respond to each question by making a mark on a 7-point Likert-like scale, ranging from –3 to 3, with a semantic differential design. Subsequently, the mean value of each quality created a scale value for pragmatism, hedonic stimulation, hedonic identification, and attractiveness. The AttrakDiff2 questionnaire has been applied to evaluate learners' experience [48,71].

Anonymous Qualitative Feedback

Each participant in this study provided anonymous feedback about their experience of the module used.

Outcome Measures

The primary outcome measure of this study was the Milestone level after completing the video learning module. The secondary outcomes were the AttrakDiff2 questionnaire scales.

Sample Size

The sample size was estimated using primary outcome effects (the Milestone) based on a priori study (360° VR video: mean 3.1, SD 0.7; 2D video: mean 2.3, SD 0.6). A two-tailed Wilcoxon-Mann-Whitney *U* test was used to calculate a sample size of 16 in each group (normal parent distribution; calculated effect size, 1.23; type I error, 0.01; power, 70%). For a block size of 8, we decided to enroll a total of 32 students to show the difference in the Milestone level.

Statistical Analysis

The D'Agostino-Pearson omnibus normality test showed that most of the continuous variables were nonnormally distributed, and they were presented as median and IQR. Differences between groups were analyzed using the Wilcoxon signed-rank test, the Mann-Whitney *U* test, or the Fisher exact test as appropriate. Effect sizes were calculated using the Hodges-Lehmann method for the Mann-Whiney *U* test and the Wilcoxon signed-rank test, or odds ratios for the Fisher exact test. The Spearman correlation test was used to analyze relationships between variables of interest. The Bonferroni correction was used to adjust *P* values because of the increased risk of a type I error when conducting multiple statistical tests at the same time [72]. Continuous variables were dichotomized using the median split. Variables of interest were analyzed for multivariate logistic regression models. All *P* values were two - sided, and statistical significance was accepted at $P<.01$. Statistical analyses were performed using G*Power software (version 3.1.9.2; Heinrich-Heine-Universität Düsseldorf), Prism for Windows (version 7.0; GraphPad Software Inc), and SPSS Statistics for Windows (version 25; IBM Corp).

Results

Study Participants

A total of 32 volunteers were recruited, of which 20 (63%) were males and 12 (38%) were females. The median age was 24 (IQR 23-25) years. There were 3 (9%) field-dependent and 29 (91%) field-independent participants. Table 2 summarizes the variables of interest for the overall study cohort. As expected, there were no significant differences in age, sex, or cognitive style between the 360° VR video and 2D video groups at baseline. After randomization, all participants received the intended intervention. There were no protocol deviations in this study.

XSL•FO
RenderX

Table 2. Demographics and cognitive style.

Variables	Overall (N=32)	360° virtual reality video group (n=16)	2D video group (n=16)	Effect size[a] (95% CI)	*P* value[b]
Demographics					
Age (years), median (IQR)	24 (23-25)	24 (23-25)	24 (23-25)	0 (–1 to 0)	.29
Sex (male), n (%)	20 (63)	10 (63)	10 (63)	1.0 (0.2 to 4.2)	>.99
Cognitive style: field-dependent, n (%)	3 (9)	1 (6)	2 (13)	2.1 (0.2 to 26.3)	>.99

[a]Effect sizes were calculated using the Hodges-Lehmann method for the Mann-Whiney *U* test and Wilcoxon signed-rank test, or odds ratios for the Fisher exact test.

[b]*P* values were calculated based on the Mann-Whiney *U* test for continuous variables (two-tailed) or the Fisher exact test for categorical variables (two-tailed).

Estimates of Subjective and Objective Cognitive Load

Paas Cognitive Load Scale

The Paas-CLS showed that the participants had a significantly higher total cognitive load than the reference value of "5" in the overall cohort (*P*=.001) after 10 minutes of video instruction (Table 3). Furthermore, the Paas-CLS score of the 360° VR video group was comparable to that of the 2D video group.

NASA Task Load Index

The overall cohort had a significantly higher score for mental demand (corrected *P*<.001) than the reference value of "10" before and after the Bonferroni correction (Table 3). The physical demand score of the 360° VR video group was higher than that of the 2D video group; however, this difference did not reach statistical significance after the Bonferroni correction (Figure 4, upper).

Table 3. Subjective measures of cognitive load.

Variables	Overall (N=32), median (IQR)	360° virtual reality video group (n=16), median (IQR)	2D video group (n=16), median (IQR)	Effect size[a] (95% CI)	*P* value[b]
Subjective measurement: Paas Cognitive Load Scale	6 (5-7)[c]	6 (5-7)	5 (5-7)	0 (–1 to 1)	.78
National Aeronautics and Space Administration Task Load Index					
Mental demand	14 (11-15)[c]	14 (12-15)	12 (10-16)	1 (–2 to 3)	.45
Physical demand	10 (7-14)	12 (9-14)	9 (4-12)	4 (0 to 7)	.047
Temporal demand	10 (7-11)	10 (8-13)	9 (6-10)	1 (–1 to 4)	.18
Performance	12 (6-15)	11 (6-15)	13 (6-15)	0 (–4 to 4)	.93
Effort	12 (10-15)	13 (11-15)	12 (7-15)	1 (–2 to 5)	.45
Frustration	8 (4-13)	10 (5-13)	7 (3-13)	2 (–2 to 6)	.29

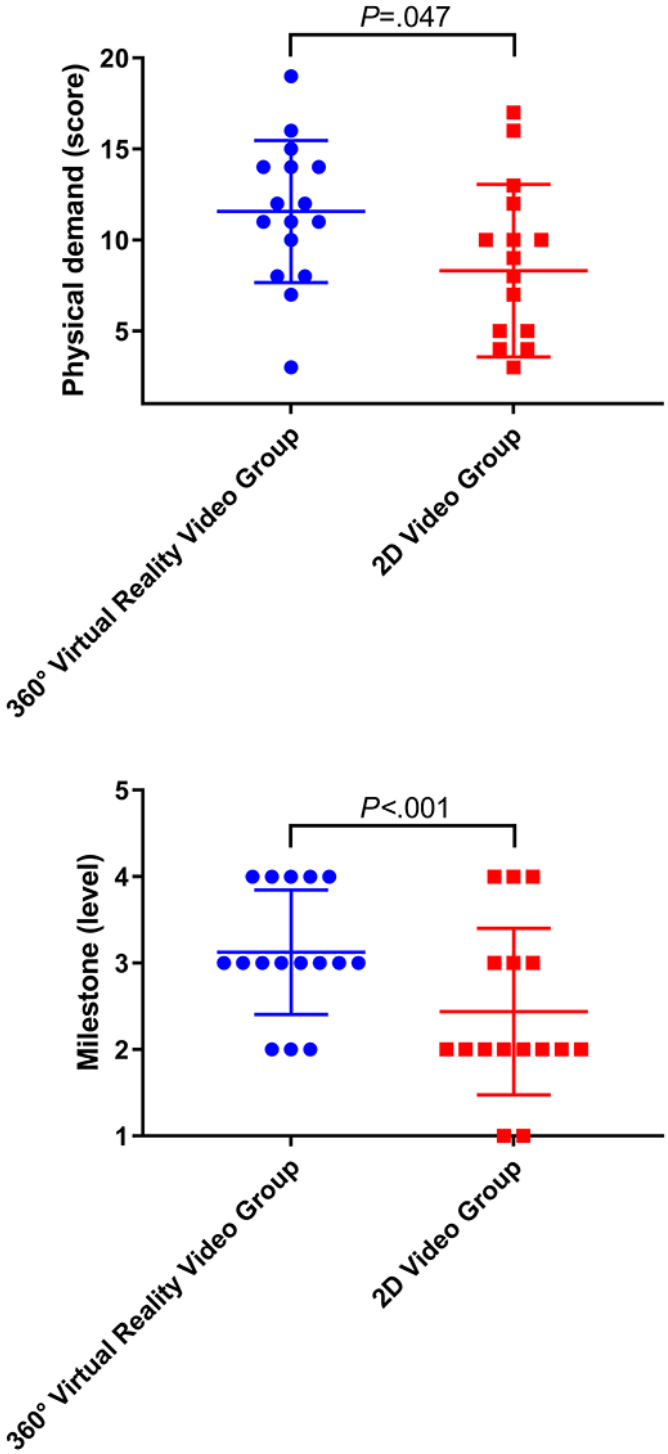
[a]Effect sizes were calculated using the Hodges-Lehmann method for the Mann-Whiney *U* test and the Wilcoxon signed-rank test, or odds ratios for the Fisher exact test.

[b]*P* values were calculated based on the Mann-Whiney *U* test (two-tailed).

[c]*P*<.01, compared with a reference value (score of "5" for the Paas Cognitive Load Scale or score of "10" for the National Aeronautics and Space Administration Task Load Index subscale) and based on the Wilcoxon signed-rank test (two-tailed). *P* values were significant after the Bonferroni correction.

XSL•FO
RenderX

Figure 4. Comparisons of the physical demands and Milestone levels between the 360° virtual reality video group and the 2D video group. The 360° virtual reality video group had a higher physical demand score (upper) and Milestone level (lower) than the 2D video group. However, these differences did not reach statistical significance. *P* values were calculated using the Mann-Whiney *U* test (two-tailed).

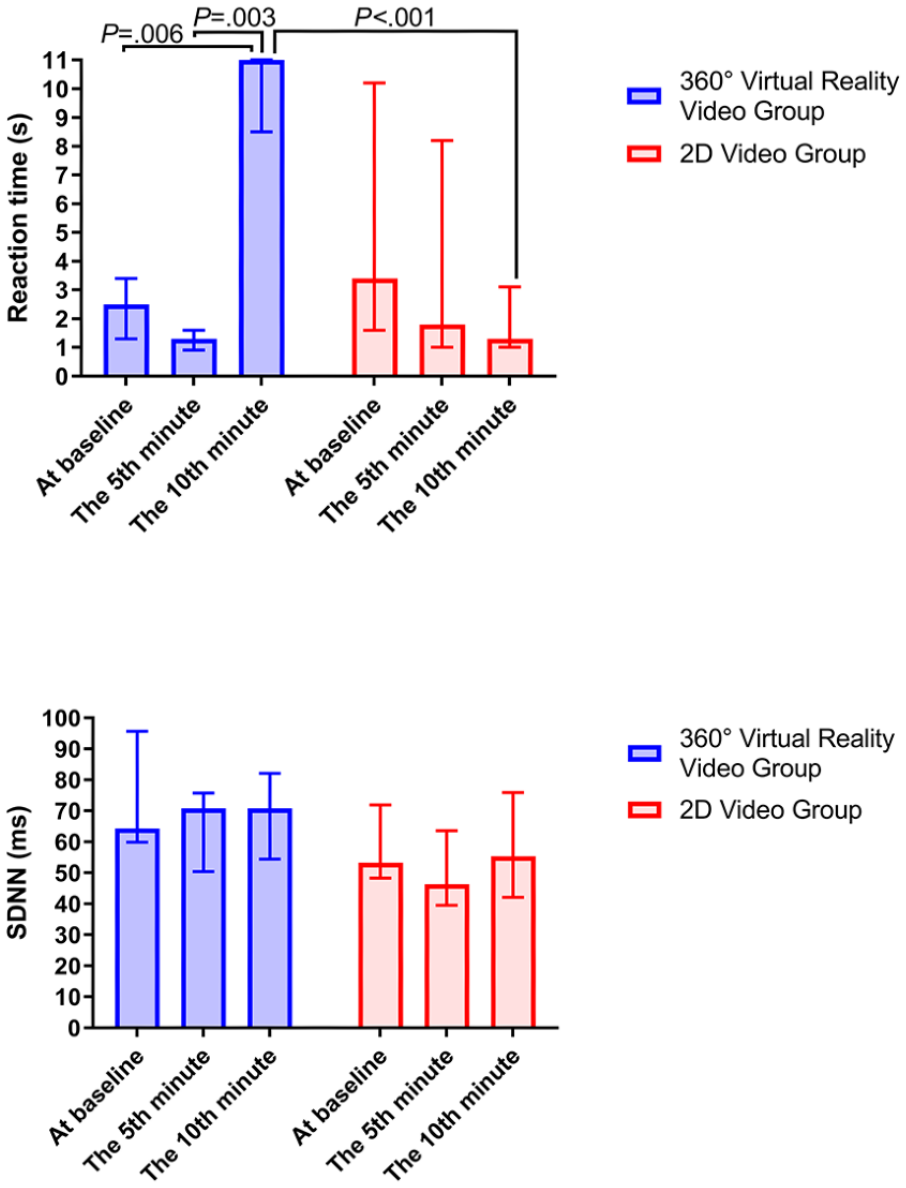


Secondary-Task Performance

In the overall cohort (Table 4), the reaction time at the 10th minute was significantly higher than the reaction time at the 5th minute (corrected *P*=.003) but comparable to the reaction time at baseline. In contrast, the reaction time at the 5th minute was equal to the reaction time at baseline after the Bonferroni correction. In the 360° VR video group, the reaction time at the 10th minute was significantly higher than the reaction time at the 5th minute (corrected *P*=.003) and the reaction time at baseline (corrected *P*=.006). Still, the reaction time at the 5th minute and the reaction time at baseline were comparable (Figure 5, upper). In the 2D video group, differences in the reaction time across various time points were not statistically significant. Furthermore, the reaction time at the 10th minute of the 360° VR video group was significantly higher than that of the 2D video group (corrected *P*<.001), even though differences in the reaction time at the 5th minute and reaction time at baseline were not statistically significant.

XSL•FO
RenderX

Figure 5. Comparisons of the secondary-task reaction time and SD of normal-to-normal RR intervals (SDNN) between the 360° virtual reality video group and the 2D video group. Notably, the reaction time in the 360° virtual reality video group was significantly higher at the 10th minute compared to the reaction times at baseline and at the 5th minute, whereas the reaction time in the 2D video group was lower at the 10th minute compared to that at baseline. The 360° virtual reality video group had a significantly higher reaction time at the 10th minute (upper) than the 2D video group (upper). The SDNN at 5 to 10 minutes in the 360° virtual reality video group was lower than the SDNN at baseline (lower). *P* values were calculated using the Wilcoxon signed-rank test or Mann-Whitney *U* test as appropriate after the Bonferroni correction.



Heart Rate Variability

Differences across the RR at 5 to 10 minutes, RR at 0 to 5 minutes, and RR at baseline were not statistically significant in the overall cohort, the 360° VR video group, and the 2D video group, respectively. Furthermore, the RR at 5 to 10 minutes, RR at 0 to 5 minutes, and RR at baseline were comparable between both groups (Table 4). Differences in the SDNN index were not statistically significant for either intragroup comparisons or intergroup comparisons. Differences in the RMSSD index were not statistically significant for either intragroup comparisons or intergroup comparisons. Differences in the LF/HF index were not statistically significant for either intragroup comparisons or intergroup comparisons.

XSL•FO
RenderX

Table 4. Objective estimates of cognitive load.

Variables	Overall (N=32), median (IQR)	360° virtual reality video group (n=16), median (IQR)	2D video group (n=16), median (IQR)	Effect size[a] (95% CI)	*P* value[b]
Secondary-task performance					
Reaction time–baseline (s)	2.6 (1.4-7.4)	2.5 (1.3-3.4)[c,d]	3.4 (1.6-10.2)	–0.6 (–5.0 to 0.7)	.34
Reaction time–5th min (s)	1.6 (0.9-3.6)[d,e]	1.3 (0.9-1.6)[d,e]	1.8 (1.0-8.2)	–0.4 (–2.0 to 0.13)	.18
Reaction time–10th min (s)	3.6 (1.3-11.0)[d,e]	11.0 (7.5-11.0)[c,d,e]	1.3 (1.0-3.1)[d]	8.3 (3.6 to 9.8)[d]	<.001
Heart rate variability					
RR–baseline (ms)	810 (741-918)	825 (742-937)	802 (721-891)	25 (–75 to 122)	.62
RR–0 to 5 min (ms)	810 (730-908)	833 (762-925)	772 (724-884)	45 (–42 to 143)	.31
RR–5 to 10 min (ms)	779 (722-889)	794 (736-911)	779 (709-866)	20 (–64 to 117)	.75
SDNN–baseline (ms)	69.9 (48.7-135.1)	88.3 (65.7-139.0)	57.3 (42.7-135.6)	25.3 (–6.1 to 65.6)	.10
SDNN–0 to 5 min (ms)	59.8 (45.5-151.3)	72.0 (58.6-184.4)	51.7 (39.2-107.7)	20.1 (–6.4 to 95.3)	.11
SDNN–5 to 10 min (ms)	69.8 (42.3-121.9)	71.2 (51.7-128.2)	58.9 (35.8-91.7)	16.3 (–20.9 to 52.2)	.27
RMSSD–baseline (ms)	60.7 (31.7-179.4)	106.3 (39.7-189.1)	34.3 (28.6-175.1)	29.8 (–6.7 to 107.2)	.13
RMSSD–0 to 5 min (ms)	63.0 (30.4-203.8)	78.0 (34.5-252.4)	37.4 (23.8-133.7)	24.8 (–12.9 to 129.0)	.29
RMSSD–5 to 10 min (ms)	54.9 (27.8-145.0)	72.1 (31.9-158.5)	39.4 (27.5-96.8)	8.7 (–17.9 to 71.5)	.51
LF/HF ratio–baseline	1.02 (0.78-1.96)	0.86 (0.77-1.96)	1.21 (0.79-1.96)	–0.13 (–0.63 to 0.39)	.59
LF/HF ratio–0 to 5 min	1.19 (0.72-2.83)	0.93 (0.67-3.35)	1.59 (0.77-2.34)	–0.11 (–1.09 to 1.11)	.91
LF/HF ratio–5 to 10 min	1.43 (0.86-2.80)	1.43 (0.74-2.89)	1.43 (0.89-2.86)	–0.04 (–0.84 to 0.79)	.81

[a]Effect sizes were calculated using the Hodges-Lehmann method for the Mann-Whiney *U* test.

[b]*P* values were calculated based on the Mann-Whiney *U* test (continuous variables).

[c]*P*<.01, compared with the baseline value, using the Wilcoxon signed-rank test (two-tailed).

[d]*P* values were significant after the Bonferroni correction.

[e]*P*<.01, compared with the 5th-minute value of reaction time or the 0-to-5–minute values of R wave to R wave (RR), standard deviation of normal-to-normal RR intervals (SDNN), root mean square of successive heartbeat interval difference (RMSSD), low frequency (LF), high frequency (HF), or LF/HF ratio, using the Wilcoxon signed-rank test (two-tailed).

Primary Outcome: Milestone Level

Overall, the participants had a significantly higher Milestone level (median 3, IQR 2-4) than the reference value of "1" (*P*<.001) after 10 minutes of video instruction. Although the Milestone level of the 360° VR video group (median 3, IQR 3-4) was higher than that of the 2D video group (median 2, IQR 2-3), the difference did not reach statistical significance (effect size=1, 95% CI 0-1; *P*=.02) (Figure 4, lower).

Secondary Outcomes: AttrakDiff2 Questionnaire Scales

The overall cohort and the 360° VR and 2D video groups had significantly positive learning experiences in terms of pragmatic quality, hedonic stimulation, hedonic identification, and attractiveness, compared with the reference value of "0" with all corrected *P*<.01, after 10 minutes of video instruction (Table 5). There were no statistically significant differences in pragmatic quality, hedonic stimulation, hedonic identification, or attractiveness between the two groups.

XSL•FO
RenderX

Table 5. Acceptance of technical innovations assessed by the AttrakDiff2 questionnaire.

Variables	Overall (N=32), median (IQR)[a]	360° virtual reality video group (n=16), median (IQR)[a]	2D video group (n=16), median (IQR)[a]	Effect size[b] (95% CI)	*P* value[c]
Pragmatic quality	1.9 (1.2-2.3)	2.0 (1.2-2.4)	1.8 (1.2-2.3)	0.1 (–0.4 to 0.7)	.59
Hedonic stimulation	2.0 (1.0-2.5)	2.1 (1.2-2.6)	2.0 (0.9-2.4)	0.1 (–0.4 to 0.7)	.42
Hedonic identification	1.9 (1.2-2.4)	1.8 (1.5-2.4)	2.1 (1.5-2.9)	0.1 (–0.6 to 0.9)	.90
Attractiveness	1.3 (0.9-2.1)	1.2 (1.1-1.9)	1.6 (0.6-2.3)	–0.1 (–0.9 to 0.7)	.84

[a]$P<.01$ for all variables, compared with a reference value of "0" for the AttrakDiff2 questionnaire using the Wilcoxon signed-rank test (two-tailed). *P* values were significant after the Bonferroni correction.

[b]Effect sizes were calculated using the Hodges-Lehmann method for the Mann-Whiney *U* test.

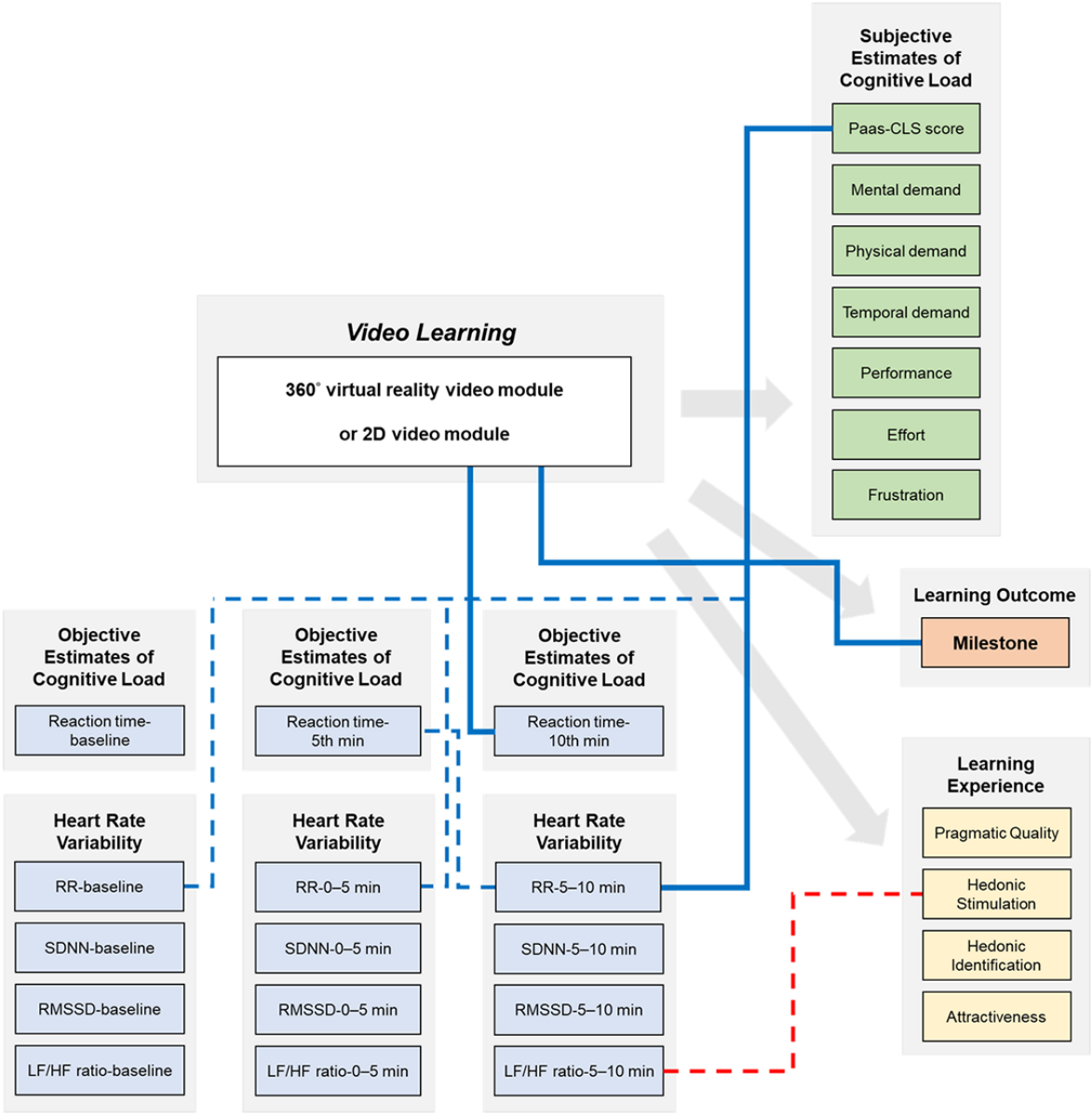
[c]*P* values were calculated based on the Mann-Whiney *U* test for continuous variables (two-tailed).

Correlations Between the Learning Modules, Cognitive Load Estimates, Learning Outcomes, and Technical Acceptance

To investigate how learning materials impact learning outcomes through the changes in cognitive load and HRV, we dichotomized the Milestone level, AttrakDiff2 questionnaire scales, cognitive load variables, and HRV indexes. Figure 6 demonstrates the associations between variables of interest.

XSL•FO
RenderX

Figure 6. Associations between the video learning modules, cognitive load measures, heart rate variability, learning outcome, and learning experience. Solid blue lines indicate an independent positive association between two variables after adjustment for video module using multivariate logistic regression models. In contrast, dashed blue lines indicate a positive association, and a dashed red line shows an inverse association without statistical significance after adjustment for the video module. HF: high frequency; LF: low frequency; Paas-CLS: Paas Cognitive Load Scale; RMSSD: root mean square of successive heartbeat interval difference; RR: R wave to R wave; SDNN: SD of normal-to-normal RR intervals.



Qualitative Evaluation

Benefits

The 360° VR video group reported that the 360° VR video module was “fun to learn” (3/16, 19%) and “good for physical examinations” (1/16, 6%). Qualitative feedback from the 2D video group emphasized that they found the 2D video module “easy to follow” (2/16, 13%) and “highly efficient” (1/16, 6%).

Faults

Out of 16 participants in the 360° VR video group, 3 (19%) reported that it caused “mild VR dizziness.” Notably, the 2D video learners said that the module was “tedious” (3/16, 19%) and that there was “no interaction” (1/16, 6%).

Discussion

Principal Findings

This study’s main findings highlighted the complexities of cognitive load estimations and HRV indexes inherent in developing and evaluating 360° VR and 2D video learning. Our results showed that video learning resulted in a higher total cognitive load and mental demand and prolonged the reaction time at the 10th minute; however, video learning also enhanced a Milestone level of H&P skills with positive learning experiences. Using the same head-mounted display, both the 360° VR and 2D videos were efficient learning methods with positive learning experiences for novice learners. Notably, the 360° VR video learners demonstrated a higher Milestone level of H&P in actual patients with SDB than the 2D video learners.

XSL•FO
RenderX

Although both video modules produced comparable total cognitive load and subscales, the 360° VR video learners had a more prolonged reaction time at the 10th minute than the 2D video learners.

Interestingly, in the second portion of the video, the 360° VR video learners had a longer reaction time than the baseline data. These findings suggested that the immersive 360° VR video could reduce secondary-task performance without increasing subjective cognitive loads. Furthermore, the elevated LF/HF ratio at 0 to 5 minutes correlated with a reduced hedonic stimulation scale. Combined with qualitative feedback, we found that the students seemed to consume more cognitive resources to fit the immersive 360° VR video than the 2D video learning. This preliminary study suggests that the immersive 360° VR video might lead to a better Milestone level than the 2D video, which seemed to be unrelated to cognitive load estimates and HRV indexes in novice learners.

Limitations

This study used quantitative and qualitative measures and methods to evaluate different video learning materials and showed its potential clinical applicability for learning H&P skills. However, several limitations should be addressed. First, this study needs external validation due to the small sample size. Although H&P is one of the essential training topics for undergraduate medical students, our outcomes may not be generalizable to junior residents who already have basic knowledge, skills, and attitudes. Second, assessments of the learner's performance in the workplace, such as the Mini Clinical Evaluation Exercise, were not included in this study. Third, there were several significant associations between variables of interest in this study. However, the potential relationships between video module, cognitive load, and HRV might be underestimated using such stringent criteria for multiple tests [73]. Accordingly, further studies with a larger sample size with the extended competence spectrum and learning outcome assessments are warranted to confirm our results and inferences.

Comparison With Prior Work

Our results indicate that both 360° VR and 2D video learning modules can increase H&P competency. Traditionally, students learn clinical skills, such as H&P, from clinical teachers. Instructional videos offer an excellent pedagogical approach to enhance medical students' clinical competencies and self-confidence levels [74,75]. Moreover, using videos as a delivery format can improve the effects and attractiveness of text-based e-learning [76]. Therefore, clinical teachers can use videos as learning resources for students' independent learning [77]. Recently, many universities and teaching institutes have promoted the use of instructional videos as a resource for self-directed education and student-centered programs, especially during the COVID-19 pandemic.

Immersive 360° VR videos allow for a virtual experience where learners can improve their competencies by experiencing all of the sights and sounds of teaching clinics and responses from the patients, families, and medical staff from a first-person perspective. When working memory and information processing capacities are limited, 360° VR video learning can trigger a higher cognitive load or stress than conventional 2D video learning. In this study, 360° VR video learning might reduce participants' attention toward the secondary task without changing subjective cognitive load estimates and HRV indexes regardless of the learning outcome. Therefore, the quality of 360° VR video learning modules still needs to be improved concerning reducing objective cognitive load.

Approximately 25% of our volunteers reported high vibration, out-of-focus blur, eye fatigue, and motion sickness in the preparation phase of the 360° VR video production. A literature review revealed that immersive 360° VR video might induce 3D visual fatigue [78] and visually induced motion sickness [79,80] and may interrupt autonomic balance [81]. Therefore, we used spherical stabilization to smooth quick movements and vibrations and dynamic depth of field to reduce visual discomfort [82]. In addition, we chose a 10-minute instructional course to reduce these known side effects. Although levels of cardiac autonomic balance, as reflected by the LF/HF ratio, were similar during the 360° VR video learning and equivalent to those of the 2D video learning, our 360° VR video learners still experienced a more prolonged secondary-task reaction time, mainly in the second portion of the video.

In contrast, the 2D video learners did not report these problems while using the same head-mounted display. Although the participants might have an adequate cognitive function to overcome the impaired attention for obtaining the H&P competence, the instructors need to design a 360° VR video with less distraction and lower cognitive load. A motion capture unit to synchronize vestibular response and visual information may reduce motion sickness [83], and further studies are warranted to investigate this issue.

Interestingly, the elevated LF/HF ratio at 5 to 10 minutes seemed to be associated with lower hedonic stimulation. Since these cognitive load estimates were not related to Milestone level, they could be used to improve the quality of video learning. For example, the virtual environment sustained the increased LF/HF ratio and acted as a stimulatory driver for cardiac autonomic activity after mental stress [84]. Therefore, limited time to immerse into the virtual environment could improve the learners' experience.

This study applied the adaptive EMD method and the well-behaved Hilbert transform technique to process the nonstationary electrocardiogram signals [63,64]. Importantly, we found that all RR intervals at different time points significantly and positively correlated with the total cognitive load. Furthermore, the significant association between the RR at 5 to 10 minutes and the Paas-CLS score was independent of the video module. Using similar analytic technology, Ghaderyan and Abbasi also observed that an increased RR interval was associated with higher cognitive load estimations [85]. Moreover, prolonged reaction time at the 5th minute seemed to be related to the RR at 5 to 10 minutes. These observations reflected that novice learners need to improve their cognitive efficiency when learning from dynamically filmed video from a first-person perspective for dealing with elevated cognitive load. These results also indicated that subjective and objective

XSL•FO
RenderX

estimates of cognitive load and HRV monitoring have considerable potential to help us better understand a learner's processing strategy to video learning modules, enhance our comprehension of contextual information, and improve video learning. Future studies are warranted to confirm the benefits of cognitive load and HRV measures in video learning.

Conclusions

Our preliminary results suggested that 360° VR video learning may result in a better H&P performance despite causing a more prolonged secondary-task reaction time than 2D video learning. Furthermore, the increased LF/HF ratio was associated with lower acceptance of video learning. Qualitative evaluation reflected varying benefits and faults between the 360° VR and 2D videos. Consequently, the students seemed to consume more cognitive resources to fit the immersive 360° VR video than 2D video learning. Without cognitive overload, the undergraduate medical students could adjust their mental efficiency to handle the decreased cognitive reserve when learning through the 360° VR video module. This study indicates that the ubiquitous and diverse roles of multimodal cognitive load estimations and HRV measures in video learning may function as part of an integrated CBME curriculum. Further research is necessary to assess their effects on other performance outcomes.

Acknowledgments

We are indebted to our trained study staff, especially Ping-Yun Lin, Yun-Kai Chang, Ming-Yen Chung, and Chung-Fang Hsiao. This work was financially supported by grants from the Ministry of Science and Technology, Taiwan, Republic of China (grants 106-2511-S-182A-003-MY2 and 108-2511-H-182A-001), and grants from the Chang Gung Medical Foundation, Taiwan, Republic of China (grants CMRPG3G1381–3 and CMRPG3G1931). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

Authors' Contributions

YPC, HHC, HYS, SLW, and LAL participated in the conception and design of this work. YPC, LJH, LYS, and LAL collected the data. YPC, HHC, TBJK, CCHY, HYS, SLW, and LAL analyzed and interpreted the data. YPC, LJH, CJK, TJF, HYL, TBJK, CCHY, HYS, LYS, and LAL carried out the development of the project software. All authors participated in the writing of the manuscript and take public responsibility for it. YPC, HHC, LJH, CGH, TBJK, CCHY, HYS, SLW, and LAL critically revised the manuscript for important intellectual content. All authors reviewed the final version of the manuscript and approved it for publication. All authors attested to the validity and legitimacy of the data in the manuscript and agree to be named as author of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT-EHEALTH checklist V1.6.1.
[PDF File (Adobe PDF File), 2510 KB - games_v9i4e13124_app1.pdf]

References

1. Carraccio C, Wolfsthal SD, Englander R, Ferentz K, Martin C. Shifting paradigms: From Flexner to competencies. Acad Med 2002 May;77(5):361-367. [doi: 10.1097/00001888-200205000-00003] [Medline: 12010689]
2. Antonoff MB, Swanson JA, Green CA, Mann BD, Maddaus MA, D'Cunha J. The significant impact of a competency-based preparatory course for senior medical students entering surgical residency. Acad Med 2012 Mar;87(3):308-319. [doi: 10.1097/ACM.0b013e318244bc71] [Medline: 22373623]
3. Abe T, Dar F, Amnattrakul P, Aydin A, Raison N, Shinohara N, et al. The effect of repeated full immersion simulation training in ureterorenoscopy on mental workload of novice operators. BMC Med Educ 2019 Aug 22;19(1):318 [FREE Full text] [doi: 10.1186/s12909-019-1752-2] [Medline: 31438934]
4. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The role of assessment in competency-based medical education. Med Teach 2010;32(8):676-682. [doi: 10.3109/0142159X.2010.500704] [Medline: 20662580]
5. Liaw SY, Wong LF, Lim EYP, Ang SBL, Mujumdar S, Ho JTY, et al. Effectiveness of a web-based simulation in improving nurses' workplace practice with deteriorating ward patients: A pre- and postintervention study. J Med Internet Res 2016 Feb 19;18(2):e37 [FREE Full text] [doi: 10.2196/jmir.5294] [Medline: 26895723]
6. Accreditation Council for Graduate Medical Education. The Otolaryngology Milestone Project. J Grad Med Educ 2014 Mar;6(1 Suppl 1):166-179 [FREE Full text] [doi: 10.4300/JGME-06-01s1-22] [Medline: 24701280]
7. Tsue TT. Developing the otolaryngology milestones. J Grad Med Educ 2014 Mar;6(1 Suppl 1):162-165 [FREE Full text] [doi: 10.4300/JGME-06-01s1-21] [Medline: 24701279]

XSL•FO
RenderX

8. Boulet JR, McKinley DW, Norcini JJ, Whelan GP. Assessing the comparability of standardized patient and physician evaluations of clinical skills. Adv Health Sci Educ Theory Pract 2002;7(2):85-97. [doi: 10.1023/a:1015750009235] [Medline: 12075142]
9. Stewart CM, Masood H, Pandian V, Laeeq K, Akst L, Francis HW, et al. Development and pilot testing of an objective structured clinical examination (OSCE) on hoarseness. Laryngoscope 2010 Nov;120(11):2177-2182. [doi: 10.1002/lary.21095] [Medline: 20824785]
10. Guze PA. Using technology to meet the challenges of medical education. Trans Am Clin Climatol Assoc 2015;126:260-270 [FREE Full text] [Medline: 26330687]
11. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual reality for health professions education: Systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
12. Sukumar S, Zakaria A, Lai CJ, Sakumoto M, Khanna R, Choi N. Designing and implementing a novel virtual rounds curriculum for medical students' internal medicine clerkship during the COVID-19 pandemic. MedEdPORTAL 2021 Mar 02;17:11106 [FREE Full text] [doi: 10.15766/mep_2374-8265.11106] [Medline: 33768143]
13. Kaufman DM, Bell W. Teaching and assessing clinical skills using virtual reality. Stud Health Technol Inform 1997;39:467-472. [Medline: 10168941]
14. Keller MS, Park HJ, Cunningham ME, Fouladian JE, Chen M, Spiegel BMR. Public perceptions regarding use of virtual reality in health care: A social media content analysis using Facebook. J Med Internet Res 2017 Dec 19;19(12):e419 [FREE Full text] [doi: 10.2196/jmir.7467] [Medline: 29258975]
15. Samadbeik M, Yaaghobi D, Bastani P, Abhari S, Rezaee R, Garavand A. The applications of virtual reality technology in medical groups teaching. J Adv Med Educ Prof 2018 Jul;6(3):123-129 [FREE Full text] [Medline: 30013996]
16. Evans CH, Schenarts KD. Evolving educational techniques in surgical training. Surg Clin North Am 2016 Feb;96(1):71-88. [doi: 10.1016/j.suc.2015.09.005] [Medline: 26612021]
17. Yoganathan S, Finch DA, Parkin E, Pollard J. 360° virtual reality video for the acquisition of knot tying skills: A randomised controlled trial. Int J Surg 2018 Jun;54(Pt A):24-27 [FREE Full text] [doi: 10.1016/j.ijsu.2018.04.002] [Medline: 29649669]
18. Hsieh MC, Lin YH. VR and AR applications in medical practice and education [Article in Chinese]. Hu Li Za Zhi 2017 Dec;64(6):12-18. [doi: 10.6224/JN.000078] [Medline: 29164542]
19. Hovgaard LH, Andersen SAW, Konge L, Dalsgaard T, Larsen CR. Validity evidence for procedural competency in virtual reality robotic simulation, establishing a credible pass/fail standard for the vaginal cuff closure procedure. Surg Endosc 2018 Oct;32(10):4200-4208. [doi: 10.1007/s00464-018-6165-5] [Medline: 29603003]
20. Våpenstad C, Hofstad EF, Bø LE, Kuhry E, Johnsen G, Mårvik R, et al. Lack of transfer of skills after virtual reality simulator training with haptic feedback. Minim Invasive Ther Allied Technol 2017 Dec;26(6):346-354. [doi: 10.1080/13645706.2017.1319866] [Medline: 28486087]
21. Andersen SAW, Konge L, Sørensen MS. The effect of distributed virtual reality simulation training on cognitive load during subsequent dissection training. Med Teach 2018 Jul;40(7):684-689. [doi: 10.1080/0142159X.2018.1465182] [Medline: 29730952]
22. Mayer RE, Moreno R. Nine ways to reduce cognitive load in multimedia learning. Educ Psychol 2003 Mar;38(1):43-52. [doi: 10.1207/S15326985EP3801_6]
23. Fraser K, Ma I, Teteris E, Baxter H, Wright B, McLaughlin K. Emotion, cognitive load and learning outcomes during simulation training. Med Educ 2012 Nov;46(11):1055-1062. [doi: 10.1111/j.1365-2923.2012.04355.x] [Medline: 23078682]
24. Frederiksen JG, Sørensen SMD, Konge L, Svendsen MBS, Nobel-Jørgensen M, Bjerrum F, et al. Cognitive load and performance in immersive virtual reality versus conventional virtual reality simulation training of laparoscopic surgery: A randomized trial. Surg Endosc 2020 Mar;34(3):1244-1252. [doi: 10.1007/s00464-019-06887-8] [Medline: 31172325]
25. Sankaranarayanan G, Odlozil CA, Wells KO, Leeds SG, Chauhan S, Fleshman JW, et al. Training with cognitive load improves performance under similar conditions in a real surgical task. Am J Surg 2020 Sep;220(3):620-629 [FREE Full text] [doi: 10.1016/j.amjsurg.2020.02.002] [Medline: 32107012]
26. van Merriënboer JJG, Sweller J. Cognitive load theory in health professional education: Design principles and strategies. Med Educ 2010 Jan;44(1):85-93. [doi: 10.1111/j.1365-2923.2009.03498.x] [Medline: 20078759]
27. Young JQ, Sewell JL. Applying cognitive load theory to medical education: Construct and measurement challenges. Perspect Med Educ 2015 Jun;4(3):107-109 [FREE Full text] [doi: 10.1007/s40037-015-0193-9] [Medline: 26016430]
28. Paas FG. Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. J Educ Psychol 1992 Dec;84(4):429-434. [doi: 10.1037/0022-0663.84.4.429]
29. Leppink J, Paas F, van Gog T, van der Vleuten CP, van Merriënboer JJG. Effects of pairs of problems and examples on task performance and different types of cognitive load. Learn Instr 2014 Apr;30:32-42. [doi: 10.1016/j.learninstruc.2013.12.001]
30. Naismith LM, Cheung JJH, Ringsted C, Cavalcanti RB. Limitations of subjective cognitive load measures in simulation-based procedural training. Med Educ 2015 Aug;49(8):805-814. [doi: 10.1111/medu.12732] [Medline: 26152492]
31. Paas F, Renkl A, Sweller J. Cognitive load theory and instructional design: Recent developments. Educ Psychol 2003 Mar;38(1):1-4. [doi: 10.1207/s15326985ep3801_1]

XSL•FO
RenderX

32. Ginns P. Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. Learn Instr 2006 Dec;16(6):511-525. [doi: 10.1016/j.learninstruc.2006.10.001]
33. Georgsson M. NASA RTLX as a novel assessment for determining cognitive load and user acceptance of expert and user-based evaluation methods exemplified through a mHealth diabetes self-management application evaluation. Stud Health Technol Inform 2019;261:185-190. [Medline: 31156113]
34. Rölfing JD, Nørskov JK, Paltved C, Konge L, Andersen SAW. Failure affects subjective estimates of cognitive load through a negative carry-over effect in virtual reality simulation of hip fracture surgery. Adv Simul (Lond) 2019;4:26 [FREE Full text] [doi: 10.1186/s41077-019-0114-9] [Medline: 31832243]
35. Haji FA, Khan R, Regehr G, Drake J, de Ribaupierre S, Dubrowski A. Measuring cognitive load during simulation-based psychomotor skills training: Sensitivity of secondary-task performance and subjective ratings. Adv Health Sci Educ Theory Pract 2015 Dec;20(5):1237-1253. [doi: 10.1007/s10459-015-9599-8] [Medline: 25761454]
36. Mandrick K, Peysakhovich V, Rémy F, Lepron E, Causse M. Neural and psychophysiological correlates of human performance under stress and high mental workload. Biol Psychol 2016 Dec;121(Pt A):62-73. [doi: 10.1016/j.biopsycho.2016.10.002] [Medline: 27725244]
37. Künstler ECS, Penning MD, Napiórkowski N, Klingner CM, Witte OW, Müller HJ, et al. Dual task effects on visual attention capacity in normal aging. Front Psychol 2018;9:1564 [FREE Full text] [doi: 10.3389/fpsyg.2018.01564] [Medline: 30233452]
38. Fadeev KA, Smirnov AS, Zhigalova OP, Bazhina PS, Tumialis AV, Golokhvast KS. Too real to be virtual: Autonomic and EEG responses to extreme stress scenarios in virtual reality. Behav Neurol 2020;2020:1-11 [FREE Full text] [doi: 10.1155/2020/5758038] [Medline: 32256856]
39. 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. Circulation 1996 Mar 01;93(5):1043-1065. [Medline: 8598068]
40. Kim HG, Cheon EJ, Bai DS, Lee YH, Koo BH. Stress and heart rate variability: A meta-analysis and review of the literature. Psychiatry Investig 2018 Mar;15(3):235-245 [FREE Full text] [doi: 10.30773/pi.2017.08.17] [Medline: 29486547]
41. Ham J, Cho D, Oh J, Lee B. Discrimination of multiple stress levels in virtual reality environments using heart rate variability. In: Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2017 Presented at: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; July 11-15, 2017; Jeju, South Korea p. 3989-3992. [doi: 10.1109/EMBC.2017.8037730]
42. McDuff D, Gontarek S, Picard R. Remote measurement of cognitive stress via heart rate variability. In: Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2014 Presented at: 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 26-30, 2014; Chicago, IL p. 2957-2960. [doi: 10.1109/EMBC.2014.6944243]
43. Fuentes-García JP, Villafaina S, Collado-Mateo D, de la Vega R, Olivares PR, Clemente-Suárez VJ. Differences between high vs low performance chess players in heart rate variability during chess problems. Front Psychol 2019;10:409 [FREE Full text] [doi: 10.3389/fpsyg.2019.00409] [Medline: 30863351]
44. Bergström I, Kilteni K, Slater M. First-person perspective virtual body posture influences stress: A virtual reality body ownership study. PLoS One 2016;11(2):e0148060 [FREE Full text] [doi: 10.1371/journal.pone.0148060] [Medline: 26828365]
45. Ohyama S, Nishiike S, Watanabe H, Matsuoka K, Akizuki H, Takeda N, et al. Autonomic responses during motion sickness induced by virtual reality. Auris Nasus Larynx 2007 Sep;34(3):303-306. [doi: 10.1016/j.anl.2007.01.002] [Medline: 17336009]
46. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system--Rationale and benefits. N Engl J Med 2012 Mar 15;366(11):1051-1056. [doi: 10.1056/NEJMsr1200117] [Medline: 22356262]
47. Hassenzahl M, Burmester M, Koller F. AttrakDiffin fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität. In: Ziegler J, Szwillus G, editors. Mensch & Computer. Weisbaden, German: Vieweg+Teubner Verlag; 2003:187-196.
48. Lee LA, Wang SL, Chao YP, Tsai MS, Hsin LJ, Kang CJ, et al. Mobile technology in e-learning for undergraduate medical education on emergent otorhinolaryngology-head and neck surgery disorders: Pilot randomized controlled trial. JMIR Med Educ 2018 Mar 08;4(1):e8 [FREE Full text] [doi: 10.2196/mededu.9237] [Medline: 29519776]
49. Noll C, von Jan U, Raap U, Albrecht UV. Mobile augmented reality as a feature for self-oriented, blended learning in medicine: Randomized controlled trial. JMIR Mhealth Uhealth 2017 Sep 14;5(9):e139 [FREE Full text] [doi: 10.2196/mhealth.7943] [Medline: 28912113]
50. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: Improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
51. Morrison GR, Ross SM, Kalman HK, Kemp JE. Designing Effective Instruction. 7th edition. Hoboken, NJ: John Wiley & Sons, Inc; 2013.

XSL•FO
RenderX

52. Renkl A, Stark R, Gruber H, Mandl H. Learning from worked-out examples: The effects of example variability and elicited self-explanations. Contemp Educ Psychol 1998 Jan;23(1):90-108. [doi: 10.1006/ceps.1997.0959] [Medline: 9514690]
53. Chi MTH, Bassok M, Lewis MW, Reimann P, Glaser R. Self-explanations: How students study and use examples in learning to solve problems. Cogn Sci 1989 Jun;13(2):145-182. [doi: 10.1016/0364-0213(89)90002-5]
54. Scott N, Smith DU, Rosenberg IK. Cognitive style and instructional materials for medical students. J Med Educ 1981 Jul;56(7):565-571. [doi: 10.1097/00001888-198107000-00004] [Medline: 7241554]
55. Chapman DM, Calhoun JG. Validation of learning style measures: Implications for medical education practice. Med Educ 2006 Jun;40(6):576-583. [doi: 10.1111/j.1365-2929.2006.02476.x] [Medline: 16700774]
56. Lee LA, Chao YP, Huang CG, Fang JT, Wang SL, Chuang CK, et al. Cognitive style and mobile e-learning in emergent otorhinolaryngology-head and neck surgery disorders for millennial undergraduate medical students: Randomized controlled trial. J Med Internet Res 2018 Feb 13;20(2):e56 [FREE Full text] [doi: 10.2196/jmir.8987] [Medline: 29439943]
57. Witkin HA, Oltman PK, Raskin E, Karp SA. A Manual for the Embedded Figures Tests. Palo Alto, CA: Consulting Psychologists Press; 1971.
58. Paas FG, Van Merriënboer JJ, Adam JJ. Measurement of cognitive load in instructional research. Percept Mot Skills 1994 Aug;79(1 Pt 2):419-430. [doi: 10.2466/pms.1994.79.1.419] [Medline: 7808878]
59. Xiao YM, Wang ZM, Wang MZ, Lan YJ. The appraisal of reliability and validity of subjective workload assessment technique and NASA-Task Load Index [Article in Chinese]. Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi 2005 Jun;23(3):178-181. [Medline: 16124892]
60. Carswell CM, Clarke D, Seales WB. Assessing mental workload during laparoscopic surgery. Surg Innov 2005 Mar;12(1):80-90. [doi: 10.1177/155335060501200112] [Medline: 15846451]
61. Guo J, Dai Y, Wang C, Wu H, Xu T, Lin K. A physiological data‐driven model for learners' cognitive load detection using HRV‐PRV feature fusion and optimized XGBoost classification. J Softw 2019 Jul 30;50(11):2046-2064. [doi: 10.1002/spe.2730]
62. Fuentes-García JP, Pereira T, Castro MA, Carvalho Santos A, Villafaina S. Psychophysiological stress response of adolescent chess players during problem-solving tasks. Physiol Behav 2019 Oct 01;209:112609. [doi: 10.1016/j.physbeh.2019.112609] [Medline: 31295452]
63. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond A Math Phys Sci 1998 Mar 08;454(1971):903-995. [doi: 10.1098/rspa.1998.0193]
64. Souza Neto EP, Custaud MA, Cejka JC, Abry P, Frutoso J, Gharib C, et al. Assessment of cardiovascular autonomic control by the empirical mode decomposition. Methods Inf Med 2004;43(1):60-65. [Medline: 15026839]
65. 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;17(3):354-381. [Medline: 8737210]
66. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health 2017;5:258 [FREE Full text] [doi: 10.3389/fpubh.2017.00258] [Medline: 29034226]
67. Solhjoo S, Haigney MC, McBee E, van Merrienboer JJG, Schuwirth L, Artino AR, et al. Heart rate and heart rate variability correlate with clinical reasoning performance and self-reported measures of cognitive load. Sci Rep 2019 Oct 11;9(1):14668 [FREE Full text] [doi: 10.1038/s41598-019-50280-3] [Medline: 31604964]
68. Suriya-Prakash M, John-Preetham G, Sharma R. Does heart rate variability predict human cognitive performance at higher memory loads? Indian J Physiol Pharmacol 2017;61(1):14-22 [FREE Full text]
69. Hansen AL, Johnsen BH, Sollers JJ3, Stenvik K, Thayer JF. Heart rate variability and its relation to prefrontal cognitive function: The effects of training and detraining. Eur J Appl Physiol 2004 Dec;93(3):263-272. [doi: 10.1007/s00421-004-1208-0] [Medline: 15338220]
70. Durantin G, Gagnon JF, Tremblay S, Dehais F. Using near infrared spectroscopy and heart rate variability to detect mental overload. Behav Brain Res 2014 Feb 01;259:16-23. [doi: 10.1016/j.bbr.2013.10.042] [Medline: 24184083]
71. Ingadottir B, Blondal K, Thue D, Zoega S, Thylen I, Jaarsma T. Development, usability, and efficacy of a serious game to help patients learn about pain management after surgery: An evaluation study. JMIR Serious Games 2017 May 10;5(2):e10 [FREE Full text] [doi: 10.2196/games.6894] [Medline: 28490419]
72. Armstrong RA. When to use the Bonferroni correction. Ophthalmic Physiol Opt 2014 Sep;34(5):502-508. [doi: 10.1111/opo.12131] [Medline: 24697967]
73. Perneger TV. What's wrong with Bonferroni adjustments. BMJ 1998 Apr 18;316(7139):1236-1238 [FREE Full text] [doi: 10.1136/bmj.316.7139.1236] [Medline: 9553006]
74. Hansen M, Oosthuizen G, Windsor J, Doherty I, Greig S, McHardy K, et al. Enhancement of medical interns' levels of clinical skills competence and self-confidence levels via video iPods: Pilot randomized controlled trial. J Med Internet Res 2011 Mar 01;13(1):e29 [FREE Full text] [doi: 10.2196/jmir.1596] [Medline: 21447471]
75. Bochenska K, Milad MP, DeLancey JO, Lewicky-Gaupp C. Instructional video and medical student surgical knot-tying proficiency: Randomized controlled trial. JMIR Med Educ 2018 Apr 12;4(1):e9 [FREE Full text] [doi: 10.2196/mededu.9068] [Medline: 29650503]

XSL•FO
RenderX

76. Walthouwer MJ, Oenema A, Lechner L, de Vries H. Comparing a video and text version of a web-based computer-tailored intervention for obesity prevention: A randomized controlled trial. J Med Internet Res 2015 Oct 19;17(10):e236 [FREE Full text] [doi: 10.2196/jmir.4083] [Medline: 26481772]
77. Azer SA, Algrain HA, AlKhelaif RA, AlEshaiwi SM. Evaluation of the educational value of YouTube videos about physical examination of the cardiovascular and respiratory systems. J Med Internet Res 2013 Nov 13;15(11):e241 [FREE Full text] [doi: 10.2196/jmir.2728] [Medline: 24225171]
78. Lambooij M, IJsselsteijn W, Fortuin M, Heynderickx I. Visual discomfort and visual fatigue of stereoscopic displays: A review. J Imaging Sci Technol 2009;53(3):030201. [doi: 10.2352/J.ImagingSci.Technol.2009.53.3.030201]
79. Moss JD, Muth ER. Characteristics of head-mounted displays and their effects on simulator sickness. Hum Factors 2011 Jun;53(3):308-319. [doi: 10.1177/0018720811405196] [Medline: 21830515]
80. Xu W, Liang HN, Zhang Z, Baghaei N. Studying the effect of display type and viewing perspective on user experience in virtual reality exergames. Games Health J 2020 Dec;9(6):405-414. [doi: 10.1089/g4h.2019.0102] [Medline: 32074463]
81. Park S, Won MJ, Mun S, Lee EC, Whang M. Does visual fatigue from 3D displays affect autonomic regulation and heart rhythm? Int J Psychophysiol 2014 Apr;92(1):42-48. [doi: 10.1016/j.ijpsycho.2014.02.003] [Medline: 24534823]
82. Carnegie K, Rhee T. Reducing visual discomfort with HMDs using dynamic depth of field. IEEE Comput Graph Appl 2015;35(5):34-41. [doi: 10.1109/MCG.2015.98] [Medline: 26416360]
83. Byagowi A, Mohaddes D, Moussavi Z. Design and application of a novel virtual reality navigational technology (VRNChair). J Exp Neurosci 2014;8:7-14 [FREE Full text] [doi: 10.4137/JEN.S13448] [Medline: 25161366]
84. Kazzi C, Blackmore C, Shirbani F, Tan I, Butlin M, Avolio AP, et al. Effects of instructed meditation augmented by computer-rendered artificial virtual environment on heart rate variability. In: Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2018 Presented at: 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; July 18-21, 2018; Honolulu, HI p. 2768-2771. [doi: 10.1109/EMBC.2018.8512816]
85. Ghaderyan P, Abbasi A. Dynamic Hilbert warping, a new measure of RR-interval signals evaluated in the cognitive load estimation. Med Eng Phys 2017 Feb;40:103-109. [doi: 10.1016/j.medengphy.2016.12.008] [Medline: 28100405]

Abbreviations

CBME: competency-based medical education
EMD: empirical mode decomposition
GEFT: Group Embedded Figures Test
H&P: history taking and physical examination
HF: high frequency
HRV: heart rate variability
LF: low frequency
NASA-TLX: National Aeronautics and Space Administration Task Load Index
ORL-HNS: otorhinolaryngology–head and neck surgery
Paas-CLS: Paas Cognitive Load Scale
RMSSD: root mean square of successive heartbeat interval difference
RR: R wave to R wave
SDB: sleep-disordered breathing
SDNN: SD of normal-to-normal RR intervals
VR: virtual reality

Edited by N Zary; submitted 12.12.18; peer-reviewed by B Birckhead, S Gallagher, T Clavier, YH Lin, S Jonas; comments to author 29.04.19; revised version received 02.04.20; accepted 10.09.21; published 22.11.21.

Please cite as:
Chao YP, Chuang HH, Hsin LJ, Kang CJ, Fang TJ, Li HY, Huang CG, Kuo TBJ, Yang CCH, Shyu HY, Wang SL, Shyu LY, Lee LA Using a 360° Virtual Reality or 2D Video to Learn History Taking and Physical Examination Skills for Undergraduate Medical Students: Pilot Randomized Controlled Trial
JMIR Serious Games 2021;9(4):e13124
URL: https://games.jmir.org/2021/4/e13124
doi:10.2196/13124
PMID:34813485

XSL•FO
RenderX

©Yi-Ping Chao, Hai-Hua Chuang, Li-Jen Hsin, Chung-Jan Kang, Tuan-Jen Fang, Hsueh-Yu Li, Chung-Guei Huang, Terry B J Kuo, Cheryl C H Yang, Hsin-Yih Shyu, Shu-Ling Wang, Liang-Yu Shyu, Li-Ang Lee. Originally published in JMIR Serious Games (https://games.jmir.org), 22.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Effects of Active Video Games on Health-Related Physical Fitness and Motor Competence in Children and Adolescents With Overweight or Obesity: Systematic Review and Meta-Analysis

Cristina Comeras-Chueca[1,2,3], MSc[‡]; Jorge Marin-Puyalto[2,3,4,5], PhD; Angel Matute-Llorente[2,3,4,5,6], PhD; German Vicente-Rodriguez[2,3,4,5,6], PhD; Jose Antonio Casajus[1,2,3,5,6], PhD; Alex Gonzalez-Aguero[2,3,4,5,6], PhD

[1]Department of Physiatry and Nursing, Faculty of Health Science, University of Zaragoza, Zaragoza, Spain
[2]GENUD Research Group (Growth, Exercise, NUtrition and Development), University of Zaragoza, Zaragoza, Spain
[3]EXERNET, Red de Investigación en Ejercicio Físico y Salud para Poblaciones Especiales, Zaragoza, Spain
[4]Department of Physiatry and Nursing, Faculty of Health and Sport Science, University of Zaragoza, Huesca, Spain
[5]Instituto Agroalimentario de Aragón -IA2-, CITA, Universidad de Zaragoza, Zaragoza, Spain
[6]Centro de Investigación Biomédica en Red de Fisiopatología de la Obesidad y Nutrición (CIBERObn), Madrid, Spain
[‡]GENUD Research Group

Corresponding Author:
Alex Gonzalez-Aguero, PhD
Department of Physiatry and Nursing
Faculty of Health and Sport Science
University of Zaragoza
Despacho nº 8
Pabellón polideportivo río Isuela, Ronda de la misericordia 5
Huesca, 22001
Spain
Phone: 34 876553755
Email: alexgonz@unizar.es

Abstract

Background: Childhood obesity is one of the most important public health problems. Active video games (AVGs) have been proposed as an attractive alternative to increase energy expenditure and are being investigated to determine their effectiveness against childhood obesity.

Objective: The aim of this study is to summarize the existing research and draw conclusions about the effects of AVGs on health-related physical fitness and motor competence in children and adolescents with overweight and obesity.

Methods: The search strategy was applied to PubMed, MEDLINE, Web of Science, and SPORTDiscus, including randomized and nonrandomized controlled trials investigating the effects of AVG programs on health-related physical fitness and motor competence in children and adolescents with overweight and obesity. To measure the risk of bias in randomized and nonrandomized controlled trials, 2 different quality assessment tools were used. In total, 15 articles met the inclusion criteria, and the variables of interest were BMI, body fat percentage, cardiorespiratory fitness (CRF), waist circumference, fat-free mass, muscular fitness, and motor competence. A meta-analysis was performed.

Results: Positive effects were found for BMI and body fat percentage, favoring the AVG group compared with a control group with no intervention (mean difference −0.209; 95% CI −0.388 to −0.031 vs mean difference −0.879; 95% CI −1.138 to −0.602). Positive effects seem to be observed for CRF. The effects of AVG interventions on muscular fitness, fat-free mass, waist circumference, and motor competence are unclear.

Conclusions: AVG programs showed positive effects on BMI, body fat percentage, and CRF. AVG could be a good strategy to combat childhood obesity.

(JMIR Serious Games 2021;9(4):e29981) doi:10.2196/29981

KEYWORDS

active videogames; exergaming; BMI; body fat; motor skills; cardiorespiratory fitness; muscle

XSL•FO
RenderX

Introduction

Background

Childhood obesity is one of the most important public health problems in the 21st century in high-income societies [1]. The prevalence of overweight and obesity in childhood has acquired the status of an epidemic. The global prevalence of overweight and obesity among children and adolescents (aged 5-19 years) has risen dramatically from 4% in 1975 to over 18% in 2016. For instance, the prevalence of overweight was over 30%, and the prevalence of obesity was over 10% in European children and adolescents in 2016 [2]. Obesity has become a pandemic owing to an obesogenic environment that causes cardiovascular and cardiometabolic diseases and psychosocial problems [3,4]. Children with overweight and obesity are likely to remain obese during adulthood and are more likely to develop many other types of cardiovascular and metabolic pathologies [1]. Evidence shows that cardiovascular risk is inversely related to physical fitness [5] and the amount of physical activity (PA) [6] performed by youth. The components of health-related physical fitness are cardiorespiratory fitness (CRF), body composition, muscular strength, muscular endurance, and flexibility [7]. Childhood obesity is related to poor health-related physical fitness, such as CRF and muscular strength [8].

On the other hand, the recommendation of the World Health Organization indicates that a daily average of 60 minutes of moderate-to-vigorous PA provides any of the health benefits in young people, although daily average of beyond 60 minutes of moderate-to-vigorous PA provides additional benefits [9]. In 2016, a study including 1.6 million students aged 11-17 years showed that 81% of them did not meet this recommendation [10]. PA, especially at moderate-to-vigorous intensity, is associated with better physical fitness, independent of sedentary time [11,12].

In addition, one of the main sedentary behaviors of this population is playing electronic games, such as computer or console games [13]. The World Health Organization reported that 40.2% of children and adolescents spend at least 2 hours per day watching television or using electronic devices on weekdays, and this percentage rises to 75.8% during weekends, going further than the recommendations of maximum screen time [14]. This inactivity and excessive sedentary screen time are catastrophic for motor development in children and adolescents [15]. A recent systematic review performed by Han et al [16] showed that children and adolescents with overweight and obesity have a lower motor competence level than children and adolescents with healthy weight; therefore, low motor competence needs to be taken into consideration in children with overweight or obesity. Moreover, children with high actual and perceived motor competence will probably show higher PA and lower BMI status [17]. An improvement in motor competence may promote better perceived motor competence, which entails higher motivation and participation in extracurricular PA and sports [18,19]. In addition, evidence shows a relationship between motor competence and health-related physical fitness during childhood and adolescence [20]. Thus, improving motor skills in children with overweight and obesity is one of the main objectives.

It is well known that exercise is an effective tool to fight obesity [21], with all its associated benefits, such as improvements in BMI status or adiposity, cardiorespiratory and muscular fitness, or bone health [22]. However, the main challenge is to ensure adherence to exercise in children with overweight and obesity [23]. Therefore, the implementation of new types of exercise that are more attractive and motivational to this population is needed.

Active video games (AVGs) have been proposed as a suitable alternative to exercise and are being investigated to determine their effectiveness against childhood obesity. AVGs generally require full-body movement and therefore increase energy expenditure [24]. A systematic review showed that structured AVG sessions had the potential to increase PA in children, but there was no evidence of the benefits of conducting them in the home setting [25]. PA and energy expenditure during AVGs are a well-studied topic showing that AVGs elicit light-to-moderate PA, and also elevate energy expenditure to moderate-to-vigorous intensity, thus having a favorable influence on energy balance [26-28]. Nevertheless, energy expenditure has been found to be higher in structured programs [29]. AVGs seem to be an interesting strategic tool to encourage an active and healthy lifestyle as an alternative to sedentary behaviors [30-32]. However, according to the overview performed by Kari [33], additional high-quality research and systematic reviews concerning exergaming are needed. In addition, AVGs seem to be an effective tool for improving self-concept, self-efficacy, situational interest and motivation, enjoyment, and psychological and social well-being [34,35]. Specifically, AVGs may have positive effects on the psychological aspects and mental health of children and adolescents with overweight or obese [36,37].

Finally, AVGs may have a positive effect on motor competence and health-related physical fitness. Some studies have shown enhancements in children's motor competence and perceived competence [38-40] or improvements in health-related physical fitness, such as cardiorespiratory and muscular fitness [41-43] and body composition [44], after an AVG intervention.

Objective

Therefore, the main aim of this systematic review is to summarize and critically appraise the existing research on the effects of AVGs on health-related physical fitness and motor competence in children and adolescents with overweight and obesity and to extract conclusions from a fair comparison of the studies included.

Methods

Data Sources and Search Strategy

This review was performed following the criteria and methodology established by the Cochrane Handbook for Systematic Reviews of Interventions (version 5.1.0) [45]. This review was performed according to the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) 2020 statement [46]. The PRISMA checklist is shown in

XSL•FO
RenderX

Multimedia Appendix 1. The protocol was registered in the International Prospective Register of Systematic Reviews, PROSPERO (CRD42020189138).

Journal articles were identified by searching electronic databases, scanning reference lists of articles, and examining tables from previous systematic reviews. The search strategy was applied to PubMed, MEDLINE, Web of Science, and SPORTDiscus up to and including March 2021.

The search strategy used to identify the articles in PubMed and MEDLINE was as follows: exergam* OR *active video gam** OR *active videogam** OR *active gam** OR *interactive video gam** OR *interactive videogam** OR *Wii* OR *Xbox* OR *Kinect* OR *PlayStation*, and *Species: Humans* and *Language: English* filters were applied, along with *Journal Article* for MEDLINE. The search strategy applied in SPORTDiscus was as follows: TX=(exergam* OR *active gam** OR *active video gam** OR *active videogam** OR *interactive video gam** OR *interactive videogam** OR *Wii* OR *Xbox* OR *Kinect* OR *PlayStation*) and *document type: article* and *language: English* filters were applied. The search strategy used in Web of Science was as follows: TS=(exergam* OR *active gam** OR *active video gam** OR *active videogam** OR *interactive video gam** OR *interactive videogam** OR *Wii* OR *Xbox* OR *Kinect* OR *PlayStation*) and *document type: article* and *language: English* filters were applied.

Two reviewers (CCC and AGA) independently evaluated all studies. Titles and abstracts were examined, and relevant articles were obtained and assessed using the inclusion and exclusion criteria presented in Textbox 1. The inclusion criteria were used following the PICOS (Population, Intervention, Comparison, Outcomes and Study) format [47]. Interreviewer disagreements were resolved by consensus. A third reviewer (JAC) resolved these disagreements.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Types of participants were children and adolescents with overweight and obesity
- Trials studying the effects of active video game programs on health-related physical fitness and motor competence
- Control group with no intervention or with traditional exercise intervention
- Types of outcome measures included variables of health-related physical fitness, such as cardiorespiratory fitness, musculoskeletal fitness (muscular strength and muscular endurance), and body composition and variables related to motor competence
- Types of studies were randomized and nonrandomized controlled trials

Exclusion criteria

- Studies were conducted in languages other than English or Spanish
- Data were unpublished
- Studies were conducted with animals
- Studies included participants aged ≥18 years
- Studies included participants with disabilities, diseases, or disorders other than obesity
- Studies were conducted without pre- and postassessments of the variables of interest
- Studies were dissertations or abstracts from society proceedings or congresses
- Studies included participants with normal weight
- Noncontrolled trials were considered in the discussion of the article with the consideration of the great limitation of the lack of a control group in interpretation of the results
- Noncontrolled trials were not included in the risk of bias assessments or meta-analysis
- All the noncontrolled trials concerning the effects of active video games on motor competence and health-related physical fitness in children and adolescents with overweight and obesity are summarized in Multimedia Appendix 2

Risk of Bias

For assessing risk of bias proposed in the PRISMA 2020 statement, 2 risk of bias assessment tools were used—the Risk of Bias 2 in randomized controlled trials (RCTs) updated by Sterne et al [48] and the ROBINS-I (Risk of Bias in Nonrandomised Studies of Interventions) in nonrandomized controlled trials developed by Sterne et al [49].

Data Extraction

The following information was extracted from each included trial: name of first author, year of publication, sample size, participant characteristics including number of participants, age and sex, type of study, type of intervention, training characteristics including intervention length and frequency, variables and data sources, and outcomes. The reported variables were weight, BMI, z-score of BMI, fat mass, body fat percentage, CRF, waist circumference, fat-free mass, muscular fitness, and motor competence.

XSL•FO
RenderX

Meta-Analyses

Children and adolescents with overweight and obesity who underwent an AVG intervention were compared with a control group (ie, group with participants performing a PA intervention and with nonintervention participants). Effect sizes were calculated for each outcome (BMI, body fat percentage, CRF, waist circumference, fat-free mass, muscular fitness, and motor competence). Different meta-analyses were performed by stratifying the studies by type of control group (no intervention or exercise intervention without AVG). When the number of articles made it possible, analyses by subgroups were performed by dividing the studies by the length of the intervention. The free cross-platform software OpenMeta[Analyst] for advanced meta-analysis was used for data processing.

Mean differences (MD) between participants in AVG interventions and controls and their 95% CIs were calculated using a continuous random-effects model (DerSimonian-Laird method). The heterogeneity of the studies was tested using the I^2 statistic [48]. This statistic describes the variance between studies as a proportion of the total variance and was interpreted as follows: I^2=0%-25% no heterogeneity; I^2=25%-50% moderate heterogeneity; I^2=50%-75% high heterogeneity; and I^2=75%-100% very high heterogeneity. All analyses were performed using the OpenMeta[Analyst] software.

Results

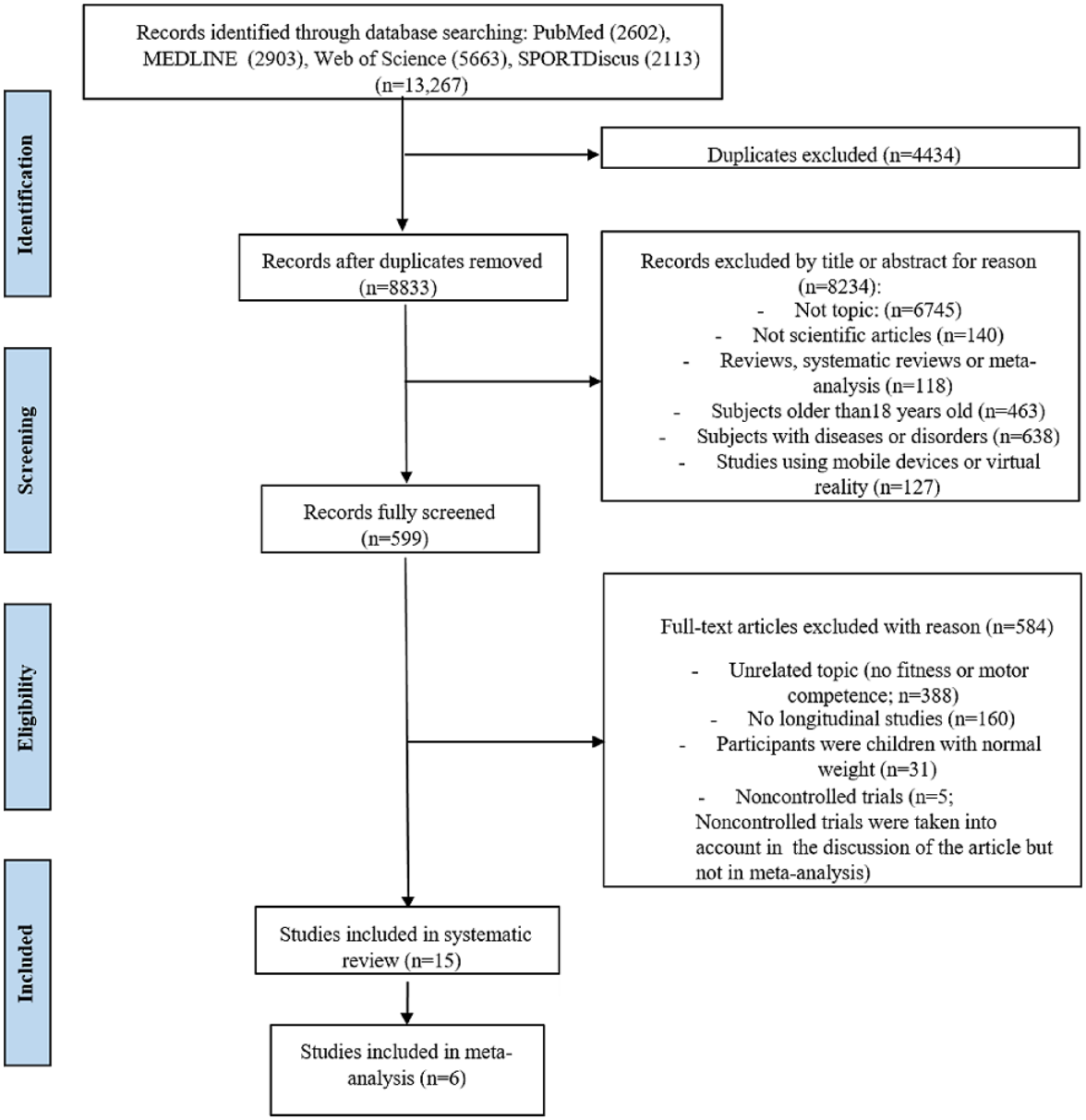
Search Summary

A total of 13,267 relevant articles were identified using the abovementioned search strategies. Following a review of titles and abstracts and excluding duplicates, the total number of articles was reduced to 599. Of them, 15 articles met the inclusion criteria and were selected for this review. Articles were excluded for the following reasons: studies were cross-sectional (n=160); only psychological, cognitive, nutritional, and balance variables, PA, or energy expenditure were measured (n=388); participants were children with normal weight (n=31); and studies were noncontrolled trials (n=5; Figure 1).

The characteristics of each study included in this systematic review were summarized in different sections following the PICOS format [47].

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of articles that were selected.



Methodological Quality

Multimedia Appendix 3 summarizes the methodological quality assessment of RCTs. The risk of bias in the RCT was low.

The quality assessment for the only nonrandomized controlled trial shows a low risk of bias on preintervention, intervention and postintervention, and therefore, a low overall risk of bias [50].

AVG Interventions

There was a great deal of variety across the AVGs used. Interventions mostly ran during physical education lessons, during playtime or lunch time as extracurricular activities after school or at home. The most commonly used devices in AVG interventions were gaming consoles, such as Xbox 360 with Kinect, Nintendo Wii, Sony PlayStation 2, dance mats, and interactive video game cycling. Games included Just Dance, Wii Fit, Wii Sports, Kinect Adventures, Kinect Sport, Dance Central, Dance Dance Revolution (DDR), and EyeToy.

The length of AVG interventions ranged from 8 weeks to 6 months (mean 16.3, SD 6.7 weeks). The frequency of AVG sessions ranged from 1 day to 5 days per week (mean 126.3, SD 55.8 minutes per week). Sessions typically lasted between 30 and 90 minutes (mean 52.0, SD 11.1 minutes) and were delivered by teachers and research assistants. It is therefore complicated to establish a standard length, intensity, and duration of sessions or type of the AVG intervention.

The different control groups either performed another intervention without AVGs, such as physical education or exercise sessions, access to sedentary video games, and learning sessions or were only asked to continue their normal activities

XSL•FO
RenderX

of daily life, the latter being the most used option for the control group.

AVG Effects

All the studies concerning the effects of AVG on motor competence and health-related physical fitness in children and adolescents with overweight and obesity are summarized in Multimedia Appendix 4.

A total of 15 randomized and nonrandomized controlled trials showed effects of AVGs on health-related physical fitness [42,44,51-62] and motor competence [50,61] in children and adolescents with overweight and obesity.

BMI, fat mass, or body fat percentage were measured in 14 studies using dual-energy x-ray absorptiometry [52,54] or bioelectrical impedance [51,58-60] to measure body fat. Waist circumference was measured in 4 studies [42,51,52,58]. CRF was evaluated in 8 studies using different tests, such as the 20-m shuttle run test [42,51,59,61], the 3-minute step test [57], and a submaximal test with a cycle ergometer [58,60]. Motor competence was only measured by Van Biljon et al [50] using the Bruininks-Oseretsky Test and by Bonney et al [61] using the Movement Assessment Battery for Children Test-Second Edition.

A quantitative analysis was performed for BMI, BMI z-score, body fat percentage, fat mass, fat-free mass, and waist circumference. Individual study results and global effects are presented in Figures 2-6, whereas a summary of the global results is presented in Table 1. Data from 2 studies [51,54] were included, taking into consideration when interpreting the data that the results obtained by these studies were adjusted for baseline outcome measures, age, and sex. Before including them, it was ascertained that the studies did not change the trend of the results without them.

Figure 2. BMI effect sizes for active video games compared with those for control group. AVG: active video game.

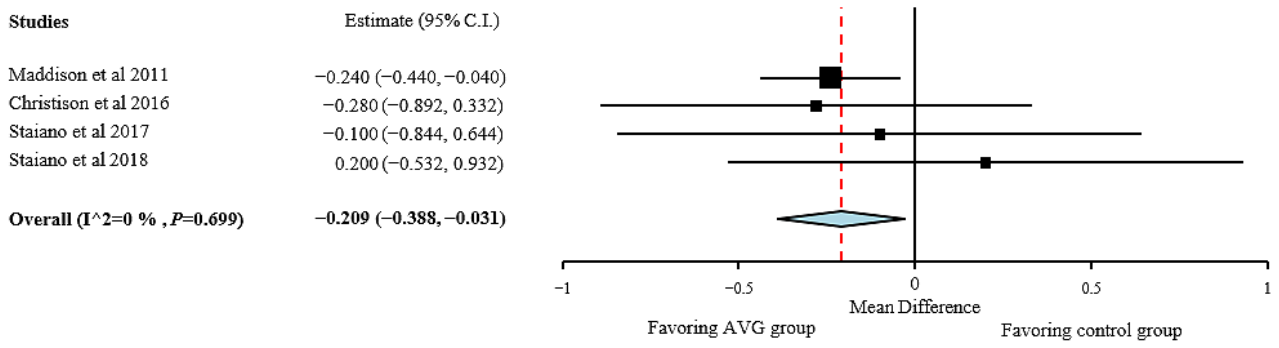
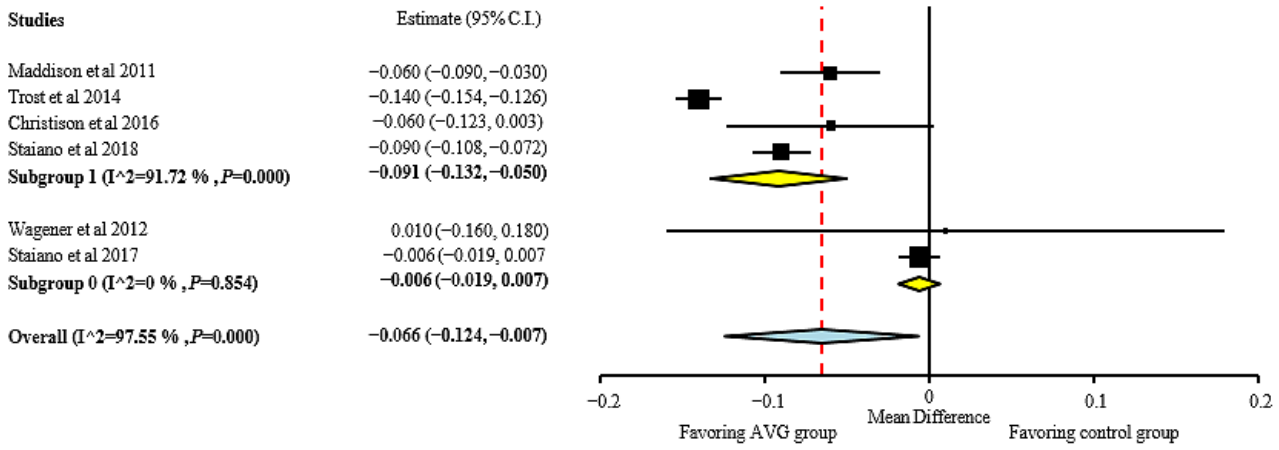


Figure 3. BMI z-score effect sizes for active video games compared with those for control group. Analysis by length of the intervention: subgroup 1: interventions lasting more than 12 weeks; subgroup 0: interventions lasting 12 weeks or less. AVG: active video game.



XSL•FO
RenderX

Figure 4. Body fat percentage effect sizes for active video games compared with those for control group. AVG: active video game.

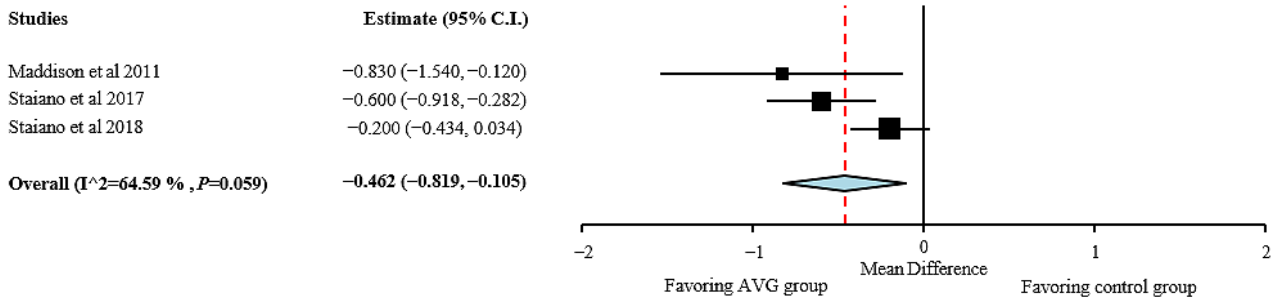


Figure 5. Waist circumference effect sizes for active video games compared with those for control group. AVG: active video game.

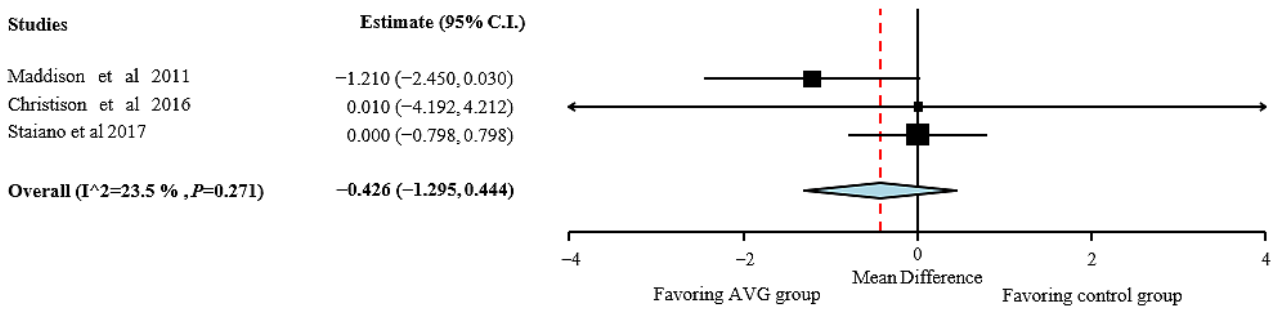


Figure 6. Fat-free mass effect sizes for active video games compared with those for control group. AVG: active video game.

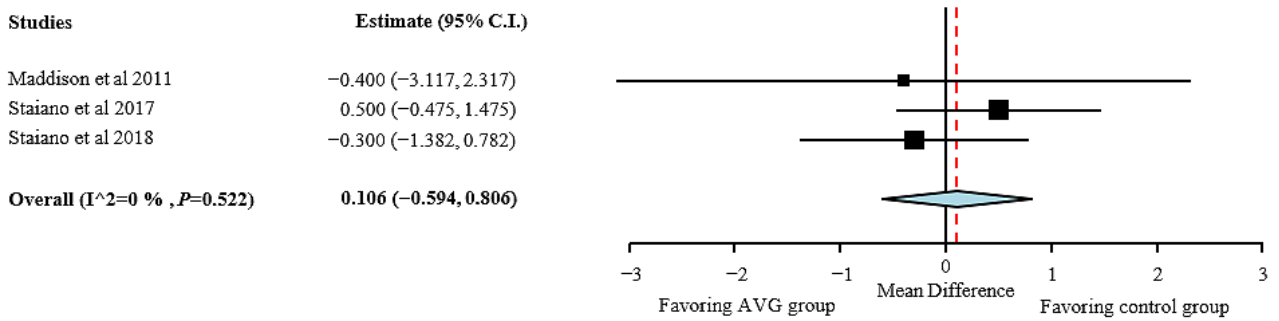


Table 1. Effect sizes and heterogeneity of findings for studies comparing active video game intervention versus control group (N=15).

Measures	Studies, n (%)	Hedge *g* effect size	Value, 95% CI	*P* value	I² (%)
BMI	4 (27)	−0.209	−0.388 to −0.031	.70	0
BMI z-score	6 (40)	−0.066	−0.124 to −0.007	<.001	97.55
Body fat percentage	3 (20)	−0.462	−0.819 to −0.105	.06	64.59
Fat-free mass	3 (20)	0.106	−0.594 to 0.806	.52	0
Waist circumference	3 (20)	−0.426	−1.295 to 0.444	.27	23.5

A quantitative analysis was not performed for CRF because of different measurement methods and articles with the same sample and for motor competence or muscular fitness because of the lack of articles. Some articles were excluded from the quantitative analyses given that the effect sizes could not be calculated from the information available in the papers [56,57], the sample was the same between studies [44,59], or the number of studies found was insufficient [50]. Another article was excluded from the quantitative analyses because of the control group exercised [58,60,61]. Noncontrolled trials [63-67] were

XSL•FO
RenderX

considered in the discussion and are summarized in Multimedia Appendix 1.

Weight, BMI, and Body Fat

A total of 13 studies evaluated changes in weight, BMI, fat mass, or body fat percentage measured. Of the 13 studies, 9 reported positive effects of an AVG intervention on body weight, BMI, or body fat in children with overweight and obesity. The first study that investigated the effect of AVG on BMI status and body composition in adolescents with overweight and obesity was performed by Adamo et al [58], and they compared a 10-week AVG cycling intervention with a stationary bike music intervention. There were no significant group or group by time effects on body weight, BMI, fat mass, or fat-free mass, but a reduction in body fat percentage was found when groups were combined and compared with baseline. Nonsignificant results could be explained by the small sample size of the study and the shortness of the intervention period. The difference in effect sizes produced by the 2 types of training on BMI or body fat could be explained by the different energy expenditure, because AVG cycling intervention spent 576.2 kcal, whereas stationary bike music group spent 554.6 kcal.

Maddison et al [51] investigated the effect of the EyeToy of PlayStation on the body composition of 322 children. Participants in the intervention group were encouraged to meet the recommendations of 60 minutes per day and to substitute periods of traditional inactive video games, and they received a PlayStation 2 and the EyeToy to play at home. Differences between control and intervention groups were found for BMI (0.24 kg/m²; 95% CI −0.44 to −0.04; *P*=.02), BMI z-score (0.06; 95% CI −0.12 to −0.03; *P*=.03), body weight (0.72 kg; 95% CI −1.33 to −0.10; *P*=.02), fat mass (0.80 kg; 95% CI −1.36 to −0.24; *P*=.005), and body fat percentage (0.83%; 95% CI −1.54 to −0.12; *P*=.02), favoring the AVG group. A year later, derived from the previous study, the authors studied the mediating effect of CRF on body composition and concluded that an AVG intervention with EyeToy can have a positive effect on body composition in children with overweight or obesity and that this effect is most likely mediated through an improvement in CRF [59]. Foley et al [44] divided the analyses of the previous study by subgroups such as ethnicity, sex, and CRF level, and the results showed that AVG can be used to improve body composition regardless of ethnicity, sex, and CRF level.

Staiano et al [56] studied the effect of playing Nintendo Wii on school days on the weight of adolescents and compared the effects between co-operative and competitive AVGs versus a control group. The results showed that the co-operative AVG group lost more weight than the control group, whereas the competitive AVG group did not differ from the co-operative AVG group and control group. The authors studied the effect of psychological variables on weight loss and, as expected, those who had higher peer support at baseline lost marginally more weight over time, but, unexpectedly, initial self-efficacy did not affect weight change over time, nor did higher self-esteem cause more weight loss over time. The 2 studies by Staiano et al [52,54] investigated the effects of AVG interventions using Xbox Kinect. In the first study [52], a dancing AVG had no effects on or differences between the AVG and control groups. The small sample size and the short length of the intervention could explain the nonsignificant effects of AVG. In the other study by Staiano et al [54], a home-based AVG intervention led to a reduction in the BMI z-score (mean −0.06, SD 0.03) and the weight z-score (mean −0.09, SD 0.05) in the AVG group in comparison with the control group (mean 0.03, SD 0.03 and mean 0.07, SD 0.04 for BMI z-score and weight z-score, respectively) when one control outlier was excluded. There was a nonsignificant intervention effect for fat mass or body fat percentage. This could be influenced by the small sample size, in addition to the fact that the performance in the sessions with AVG at home showed lower or no benefits.

Trost et al [55] compared the effects of a 16-week weight management program with family-based theoretical sessions focused on lifestyle and the effects of the same program with AVG intervention. The overweight rate was reduced by 5.4% in the weight management program without PA and 10.9% in the AVG group, with significant pre-post and between-group differences. Both groups exhibited reductions in BMI z-score, but the AVG group showed greater reductions (mean −0.25, SD 0.03 vs mean −0.11, SD 0.03).

The most recent study was conducted by Irandoust et al [62], and the results showed reductions in weight and BMI in the AVG and exercise groups from pretest to posttest, resulting in lower weight and BMI at posttest measures in these groups in comparison with the control group after a 6-week intervention using Xbox Kinect and Nintendo Wii.

Of the 13 studies, 4 reported no positive effects of an AVG intervention on BMI or body fat in children with overweight and obesity [42,52,53,57]. Apart from Staiano et al [52], no other authors found effects after an AVG intervention. Christison et al [42] compared the effects of weight management didactic sessions with an AVG intervention with Nintendo Wii and DDR. The results showed a trend of reduction in the BMI z-score in the AVG group, whereas the BMI z-score among the control group was essentially unchanged (*P*=.07). Wagener et al [53] showed no pre-post differences in BMI within or between conditions after 10 weeks of an AVG intervention by playing dance-based AVGs. Probably, no benefits of AVGs and even nonsignificant worse results on BMI for the AVG group were found because of the small sample size and the short length of intervention. There were same limitations in the study performed by Maloney et al [57], which showed that there were no changes between pretest and posttest participant weight in the AVG group or control group after 12 weeks of intervention with AVGs.

Furthermore, 5 noncontrolled studies observed positive effects of an AVG intervention on BMI or body fat in children with overweight and obesity. The most recent study was performed by Argarini et al [67], and the results showed a significant decrease in weight, BMI, and body fat percentage after a 4-week AVG intervention with Xbox Kinect. Christison et al [66] evaluated the efficacy of an AVG intervention and the results showed a significant decrease in BMI (mean −0.48, SD 0.93 kg/m²) and BMI z-score (mean −0.07, SD 0.14) after 10 weeks of training (*P*=.002 and *P*<.001, respectively). Duman et al [64] investigated the effects of a combination of music-accompanied

XSL•FO
RenderX

aerobics, callisthenic exercises, and AVGs, and the results showed decrease in BMI and triceps skinfold thickness. The percentage of obese children decreased from 72% to 40%; those children who were obese became children with overweight, so the percentage of overweight children increased from 28% to 46%; the percentage of children with normal weight increased from 0% to 14%. Calcaterra et al [65] demonstrated the effectiveness of a combination of circuit-based aerobics and strength and resistance exercises with AVGs, showing a significant decrease in BMI (from 32.9 to 31.9 kg/m²; *P*=.002) and body fat percentage (from 39.3% to 36.0%; *P*=.001). A very interesting result of this study was that 27.2% of the participants reported a previous negative experience with exercise, so a reduced drop-out rate during activity may be achieved with a playful aspect and adapted activities such as AVGs. Finally, Huang et al [63] investigated the effect of AVGs using Nintendo Wii and Xbox Kinect, with no effects on the percentage of body fat, probably because of the short length of the intervention and the reduced number of participants.

Systematic reviews have been performed on the effects of AVG on BMI or body fat, but they are mostly not focused on children or adolescents with overweight or obesity and including studies with children with normal weight; some limitations can be found in these studies, such as the inclusion of noncontrolled trials. The results of these studies are in line with the results of this study. The latest systematic review was performed by Gao et al [68], who included noncontrolled trials and studies with children with normal weight. Reduction in BMI after AVG interventions was found in children and adolescents. Hernández-Jimenez et al [69] performed a meta-analysis that showed a significant effect in favor of AVGs on BMI in children and adolescents, with better results achieved when the AVG intervention was applied to children with overweight or obesity. Another systematic review [70] included 4 RCTs, which are also included in this systematic review, which reported decreases in BMI or body fat after an AVG intervention. A previous systematic review performed by Gao et al [31] concluded that AVGs were a promising tool to promote PA and health as long as the AVG intervention is not home based, but this review did not focus on children with overweight or obesity. Two systematic reviews [30,71] supported the findings, although being among the first reviews on the effects of AVGs, quantitative analyses were not conducted because of a lack of articles. Lamboglia et al [30] found that AVG led to increased PA and CRF and decreased body fat, with considerable potential to fight obesity. Leblanc et al [71] found that AVG attenuated weight gain in participants with overweight and obesity, including 3 articles that are included in this systematic review. The improvement of cardiometabolic health through AVG was inconclusive because of the small number of articles at the time.

Quality Assessment of BMI, Body Fat Percentage, and Fat Mass

As shown in Figure 2, positive effects of the interventions were found for BMI, favoring the AVG group compared with the control group with no intervention (MD −0.209; 95% CI −0.388 to −0.031). Heterogeneity among studies for BMI was low (I^2=0%; *P*=.70). AVG showed more positive effects on BMI z-score than on BMI (MD −0.066; 95% CI −0.124 to −0.01), but it also showed a very high heterogeneity (I^2=97.55%; *P*<.001; Figure 3). The results of the subgroup analysis by the length of the intervention showed that the decrease in BMI z-score was higher in the AVG interventions longer than 12 weeks.

As shown in Figure 4, positive effects of AVG interventions were found for body fat percentage, favoring the AVG group compared with the control group with no intervention (MD −0.462; 95% CI −0.819 to −0.105). Heterogeneity among studies for BMI was high (I^2=64.59%; *P*=.06).

These results clearly showed the influence of AVG intervention length on weight, BMI, and body fat percentage. Positive effects were observed for AVG interventions longer than 12 weeks. It seems that a combination of AVG with multicomponent exercises could have more benefits on BMI and body fat percentage in children and adolescents with overweight and obesity, but RCTs are needed to confirm this.

Waist Circumference

Changes in waist circumference were evaluated by 4 RCTs [42,51,52,58], and no effects were found. The first study was performed by Adamo et al [58] and showed no effects or differences between groups for waist circumference. Maddison et al [51] reported no changes in waist circumference after a 24-week AVG intervention with EyeToy performed at the participants' homes. Christison et al [42] performed a 6-month AVG intervention with Nintendo Wii and DDR; they also did not report positive effects or differences between groups. The most recent study on waist circumference in children and adolescents with overweight and obesity was performed by Staiano et al [52], who investigated the effects of Xbox Kinect and found no effect on waist circumference or differences between AVG and control groups.

The noncontrolled trial performed by Calcaterra et al [65] demonstrated a decrease in waist circumference (-5.9 cm) and waist circumference to height ratio (-0.08) after a 12-week training program combining traditional exercise with AVGs.

Quality Assessment of Waist Circumference

As shown in Figure 5, no overall effects were found on waist circumference after the AVG interventions (MD −0.426; 95% CI −1.295 to 0.444). Heterogeneity among studies for waist circumference was moderate (I^2=23.5%; *P*=.27). AVGs seem not to be effective in decreasing waist circumference in children and adolescents with overweight and obesity. It is necessary to look for a way to increase the demands of the activity.

Interventions with AVG do not seem to be effective in decreasing waist circumference in children and adolescents with overweight and obesity. This result may be because of the length of the interventions in the included articles. A reduction in the waist circumference in children with obesity seems to be possible with a combination of AVG with multicomponent exercise instead of AVG exclusively, but RCTs are needed to confirm this. Waist circumference is as important as BMI or body fat percentage because they are good predictors of cardiovascular disease risk factors in children and adolescents, even better than BMI [72]. Therefore, the main aim is to

XSL•FO
RenderX

decrease waist circumference of children and adolescents with overweight and obesity. However, the results suggest that AVG interventions do not seem to be effective in decreasing waist circumference in children and adolescents with overweight and obesity.

Fat-Free Mass

The effects of AVG on fat-free mass were reported in 4 articles [51,52,54,58]. To measure fat-free mass, bioelectrical impedance [51,58] or dual-energy x-ray absorptiometry [52,54] were used. Maddison et al [51] did not find any effects or differences between groups for fat-free mass. Similar results were reported by Adamo et al [58], with no changes or differences between groups. In contrast, Staiano et al [52,54] reported no effects on lean mass after an AVG intervention. Evidence on the effect of AVGs on fat-free mass is limited, and no effects have been shown.

Quality Assessment of Fat-Free Mass

As shown in Figure 6, no overall effects were found for fat-free mass after the AVG interventions (MD 0.106; 95% CI −0.594 to 0.806). Heterogeneity among studies for waist circumference was low (I^2=0%; *P*=.52).

Cardiorespiratory Fitness

CRF assessments following their AVG interventions were included by 5 studies, and 4 of them managed to find positive results. In general, CRF was improved after an intervention with AVG, such as Nintendo Wii, DDR, EyeToy, or a cycling AVG.

The first study that reported the effects of AVG on CRF in adolescents with overweight and obesity was the study by Adamo et al [58], who observed a significant training effect over time in both AVG cycling and stationary bike interventions. Both interventions produced significant improvements in peak heart rate, peak workload, and time to exhaustion, but no significant differences were found between the exercise groups. With this same intervention, Goldfield et al [60] observed that the psychological benefits of aerobic exercises were related to improved aerobic fitness. The abovementioned study by Maddison et al [51] did not find significant increases in CRF in the AVG group, but the positive effect of AVG on body composition in children with overweight or obesity is most likely mediated through improved aerobic fitness [59]. Maloney et al [57] showed no improvements in CRF in either the AVG or control group after playing DDR for 12 weeks. Christison et al [42] showed that the number of shuttle runs did not change after a 6-month AVG intervention. The most recent study was conducted by Bonney et al [61], who investigated the effect of Wii Fit in comparison with a task-oriented functional training on the performance of the shuttle run test and positive effects on CRF in both groups, but no differences were found between the AVG and control groups performing the task-oriented functional training.

The effects of AVG on CRF in children with overweight and obesity have been studied by 2 noncontrolled trials [63,65]. Calcaterra et al [65] demonstrated in their study an improvement in CRF (3.8 mL/kg/min; *P*<.001) measured by a walking test on a treadmill reaching 85% of the maximal heart rate. Huang et al [63] showed no effects of AVG using Nintendo Wii and Xbox Kinect on CRF, but the heart rate demonstrated that most participants were able to achieve moderate or vigorous intensity of exercise during most AVG sessions.

The effect of AVG interventions on CRF remains unclear. Probably, the limited effects of AVG interventions on CRF of children with overweight and obesity might be because of insufficient training volume in terms of either weekly frequency or overall duration of the interventions. As mentioned earlier, interventions performed at home could be ineffective for improving health-related physical fitness, such as CRF. As it occurs in children with normal weight, only Calcaterra et al [65] used a submaximal or maximal incremental cardiopulmonary exercise test with a gas analyzer, which is widely recognized as the best single index of aerobic fitness [73,74]. Once again, science-based evidence shows that a combination of AVG with multicomponent exercise could produce more benefits on CRF than AVG exclusively, probably because of a higher volume of training. Therefore, these results must be interpreted with caution because the studies that report results from interventions using AVGs with multicomponent exercise are noncontrolled trials. RCTs are needed to confirm this finding.

A systematic review performed by Zeng and Gao [70] included only 1 RCT, also included in this systematic review, which reported positive effects of an AVG intervention in comparison with the effects of an exercise group, but these results were unclear because of the inclusion of only 1 article.

Muscular Fitness

Only 1 RCT [61] and 3 noncontrolled trials [63-65] investigated the effects of an AVG intervention on the muscular fitness of children and adolescents with overweight and obesity. The only RCT about the effects of AVGs on muscle fitness showed that both the AVG group trained with Wii Fit and the control group that performed a task-oriented functional training for 14 weeks improved knee extensors and ankle plantar flexors to maximal isometric strength assessed with a handheld dynamometer.

Calcaterra et al [65] demonstrated an increase in muscular strength, improving from a mean of 29.6 kg (SD 9.3 kg) to 32.3 kg (SD 9.8 kg; *P*=.003) in a handgrip test. Duman et al [64] investigated the effects of an AVG intervention combined with traditional exercise on several physical performance tests that require muscle strength and endurance, such as time to ascend and descend 20 stairs, number of squats they can perform in 120 seconds, time to run 50 m, and rope jumps in 30 seconds. The results showed enhancements in all test performances. Finally, Huang et al [63] investigated the effect of AVG using Nintendo Wii and Xbox Kinect on the muscular strength of the quadriceps and hamstrings that were assessed using a handheld dynamometer and muscular endurance that were assessed by a 1-minute half-sit-up test consisting of completing as many half-sit-ups as possible within a minute. Nonsignificant changes in the muscular strength of the quadriceps or muscular endurance were observed. Low frequency and program duration may explain the lack of significant changes.

XSL•FO
RenderX

A combination of AVG with multicomponent exercise could enhance muscular fitness in children with overweight and obesity, but RCTs are required to confirm these results. A systematic review and meta-analysis [5] showed the importance of muscular fitness for children and adolescents and found associations between muscle fitness and bone health, total and central adiposity, and cardiovascular diseases, metabolic risk factors and self-esteem. In addition, according to Tomlinson et al [75], children and adolescents with overweight and obesity have a greater absolute maximum muscle strength than nonobese persons because increased adiposity induces a chronic overload stimulus on the antigravity muscles; however, when maximum muscular strength is normalized to body mass, individuals with obesity appear weaker, probably because of reduced mobility, neural adaptations, and changes in muscle morphology. Therefore, it is important to include exercises aimed at improving muscular fitness in programs for children and adolescents with overweight and obesity.

Motor Competence

Motor competence after an AVG intervention was reported in 2 articles. The first was performed by Van Biljon et al [50], and motor competence was evaluated in 30 individuals with overweight and obesity using the shorter version of the Bruininks-Oseretsky test for motor proficiency. The intervention group performed a 6-week AVG intervention for 3 days per week and 30 minutes per session using Wii; there were 2 control groups, one with access to traditional video games and the other continued with their everyday life activities with no intervention. The AVG group showed improvements in motor competence compared with both control groups, specifically in terms of agility and speed, co-ordination, and reaction time. Another more recent study by Bonney et al [61] showed that both the AVG group that trained with Wii Fit and the control group that performed a task-oriented functional training for 14 weeks improved motor co-ordination, as measured by the Movement Assessment Battery for Children Test-Second Edition. A notable difference between the controlled trials investigating the effect of AVG on motor competence is the length of the intervention; therefore, the study with the longest duration of the intervention [61] (14 weeks) showed positive effects, whereas the one with the shortest duration [50] (6 weeks) showed no effect after the AVG intervention. Another study with no control group showed an improvement in motor competence after 4 weeks of training with Xbox Kinect [67]. This scarcity of studies investigating the effect of AVG on the motor competence of children and adolescents with overweight and obesity is of great importance, given that evidence shows a relationship between motor competence and health-related physical fitness during childhood and adolescence [20], as mentioned earlier. Furthermore, low motor competence, denominated as *physical illiteracy*, is a component of the pediatric inactivity triad observed by Faigenbaum et al [76], together with exercise deficit disorder and pediatric dynapenia. These 3 components are closely related to each other. Children and adolescents who perceive themselves as having poor motor competence might feel less inclined to participate in PA or sports, which in turn will reduce their ability to improve their muscular fitness or motor competence. Motor competence is particularly important in children and adolescents with overweight and obesity, as they have a lower motor competence level than children and adolescents with a healthy weight [16]; this has health consequences because it is directly related to PA. Improving motor competence could increase PA, reduce sedentary behaviors, and positively impact health-related physical fitness [16,77]. Improvements in motor competence entail higher motivation and participation in extracurricular PA and sports [18,19]. AVGs could be a tool for improving motor competence, but randomized clinical trials are needed to corroborate this.

Interventions with AVG appear to be more effective in decreasing BMI and body fat in children and adolescents with overweight or obesity than in children with normal weight, but effectiveness in other health-related physical fitness parameters such as waist circumference or CRF is still unclear. Children and adolescents with overweight or obesity could benefit from AVGs to improve their motor competence, which seems to be a variable more susceptible to enhancement as improvements have been observed with shorter interventions, but further research is needed to confirm this hypothesis. Therefore, AVG programs could be a good strategy to combat childhood obesity.

Discussion

Principal Findings

This paper provides knowledge about the effects of AVG on health-related physical fitness and motor competence in children and adolescents by gathering previous scientific evidence, and it also provides the prospects for future studies.

Mental health can be considered as a major influencing variable and can undoubtedly influence the effects of AVG interventions on childhood obesity; however, this systematic review has focused on the effects of AVG interventions on health-related physical fitness and motor competence in children and adolescents with overweight and obesity, discarding the psychological aspects. However, a very recent overview [37] and a systematic review [36] found a positive effect of AVG on mental health, but it also showed the need for increasing scientific research in this area.

In contrast, the participants in the studies included in this review were from countries with a medium-to-high socioeconomic status; this leads to the limitation that such AVG interventions are difficult to implement or are not applicable in countries with a low socioeconomic status because of the lack of resources and lower purchasing power. It would be interesting to study the feasibility and possibilities of such interventions in societies with low socioeconomic status.

Furthermore, studies investigating the effects of AVG in children with obesity with some information from geohash dashboards related to the environment where the participants live and where the school is located can be useful to increase the strength of the hypothesis on how AVG affects the features addressed, as there are several features where weak correlation has been found with AVG use, as. As stated earlier, it is necessary to establish the guidelines for an effective intervention using AVGs. It seems that the most effective application of an AVG intervention is a sufficient duration and the structuring and planning of that

XSL•FO
RenderX

intervention. In addition, most of the scientific evidence that has studied AVG interventions combined with traditional exercise lacked a control group to compare the effects; these AVG interventions seem to be promising, but RCTs are needed to investigate the effects of these AVG interventions.

Finally, few studies have examined the effects of an AVG intervention on muscular fitness and motor competence, which, as mentioned earlier, are components of the pediatric inactivity triad observed by Faigenbaum et al [76] together with exercise deficit disorder. The importance of these 2 variables lies in their close relationship with PA and sports participation, which can improve the physical fitness and body composition of children and adolescents, which is especially important in those who are overweight or obese. By improving the muscle strength and motor skill of these children and adolescents, we may make them more active and therefore healthier.

RCTs are needed to investigate the effects of AVG interventions on children's muscle strength and motor skills to learn about the possibilities of AVG interventions in stopping the vicious circle of pediatric inactivity triad.

Limitations

The limitations of this review should be acknowledged. A wide variety of AVG interventions have been included, with different devices and training interventions (duration, frequency, training setting or training dynamic, and type of AVG), which makes it difficult to analyze all the articles together and to obtain generalized results. In addition, the potential risk of bias of some studies was not considered when interpreting the results. Finally, some subgroup analyses were not performed because of the small number of controlled trials. Gender, demographics, or race influence were not deeply addressed because the studies did not show the results divided by these covariates. However, it would be interesting to investigate whether such interventions are more effective depending on them.

Strengths

This study also has several strengths. To the best of our knowledge, this is the first meta-analysis to summarize the existing research on the effects of AVG on health-related physical fitness and motor competence in children and adolescents with overweight and obesity. This analysis included not only the effects of AVG on BMI, but also those on body composition, CRF, muscular fitness, and motor competence. This study allowed us to realize that more RCTs reporting motor competence and muscular fitness results are needed.

Conclusions

AVGs could be a good strategy to fight childhood obesity. AVG programs showed positive effects on BMI and body fat percentage. Improvements in CRF have been observed after an AVG intervention. Children and adolescents could benefit from AVGs to improve motor competence, but further research is needed to confirm these results. The effects of AVG programs on muscular fitness or fat-free mass are also unclear.

In conclusion, AVGs seem to be an effective tool to improve health-related physical fitness and is a promising tool for improving motor competence in children and adolescents. AVGs can even be considered as a prospective alternative to traditional exercise for enhancing health status during childhood.

Acknowledgments

CCC received a grant from Gobierno de Aragón (DGA IIU/2023/2017).

Authors' Contributions

All authors have been actively involved in the planning and execution of this study. AGA and JAC were the main researchers in this study, and CCC is the first author. AML, JMP, and GVR were coresearchers. CCC and AGA independently evaluated all the studies. CCC drafted the document, and JMP, AML, GVR, JAC, and AGA critically reviewed the document. All authors read and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 item Checklist.
[PDF File (Adobe PDF File), 231 KB - games_v9i4e29981_app1.pdf]

Multimedia Appendix 2
Descriptive characteristics of included noncontrolled trials.
[PDF File (Adobe PDF File), 195 KB - games_v9i4e29981_app2.pdf]

Multimedia Appendix 3
Quality assessment of randomized controlled trials.
[PDF File (Adobe PDF File), 13 KB - games_v9i4e29981_app3.pdf]

XSL•FO
RenderX

Multimedia Appendix 4
Descriptive characteristics of included studies with children with overweight and obesity.
[PDF File (Adobe PDF File), 29 KB - games_v9i4e29981_app4.pdf]

References

1. Ahluwalia N, Dalmasso P, Rasmussen M, Lipsky L, Currie C, Haug E, et al. Trends in overweight prevalence among 11-, 13- and 15-year-olds in 25 countries in Europe, Canada and USA from 2002 to 2010. Eur J Public Health 2015 Apr;25 Suppl 2:28-32. [doi: 10.1093/eurpub/ckv016] [Medline: 25805783]
2. Nittari G, Scuri S, Petrelli F, Pirillo I, di Luca NM, Grappasonni I. Fighting obesity in children from European World Health Organization member states. Epidemiological data, medical-social aspects, and prevention programs. Clin Ter 2019;170(3):223-230 [FREE Full text] [doi: 10.7417/CT.2019.2137] [Medline: 31173054]
3. McCrindle BW. Cardiovascular consequences of childhood obesity. Can J Cardiol 2015 Feb;31(2):124-130. [doi: 10.1016/j.cjca.2014.08.017] [Medline: 25661547]
4. Gurnani M, Birken C, Hamilton J. Childhood obesity: causes, consequences, and management. Pediatr Clin North Am 2015 Aug;62(4):821-840. [doi: 10.1016/j.pcl.2015.04.001] [Medline: 26210619]
5. Smith JJ, Eather N, Morgan PJ, Plotnikoff RC, Faigenbaum AD, Lubans DR. The health benefits of muscular fitness for children and adolescents: a systematic review and meta-analysis. Sports Med 2014 Sep;44(9):1209-1223. [doi: 10.1007/s40279-014-0196-4] [Medline: 24788950]
6. Cesa CC, Sbruzzi G, Ribeiro RA, Barbiero SM, de Oliveira Petkowicz R, Eibel B, et al. Physical activity and cardiovascular risk factors in children: meta-analysis of randomized clinical trials. Prev Med 2014 Dec;69:54-62. [doi: 10.1016/j.ypmed.2014.08.014] [Medline: 25175591]
7. American College of Sports Medicine. Acsm's Health-related Physical Fitness Assessment Manual. 5th Ed. Philadelphia, Pennsylvania, United States: Lippincott Williams & Wilkins; 2013.
8. Casonatto J, Fernandes RA, Batista MB, Cyrino ES, Coelho-E-Silva MJ, de Arruda M, et al. Association between health-related physical fitness and body mass index status in children. J Child Health Care 2016 Sep;20(3):294-303. [doi: 10.1177/1367493515598645] [Medline: 26396021]
9. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
10. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. Lancet Child Adolesc Health 2020 Jan;4(1):23-35 [FREE Full text] [doi: 10.1016/S2352-4642(19)30323-2] [Medline: 31761562]
11. Marques A, Santos R, Ekelund U, Sardinha LB. Association between physical activity, sedentary time, and healthy fitness in youth. Med Sci Sports Exerc 2015 Mar;47(3):575-580. [doi: 10.1249/MSS.0000000000000426] [Medline: 24977696]
12. Júdice PB, Silva AM, Berria J, Petroski EL, Ekelund U, Sardinha LB. Sedentary patterns, physical activity and health-related physical fitness in youth: a cross-sectional study. Int J Behav Nutr Phys Act 2017 Mar 04;14(1):25 [FREE Full text] [doi: 10.1186/s12966-017-0481-3] [Medline: 28259140]
13. Forde C, Hussey J. How children use active videogames and the association between screen time and physical activity. Games Health J 2015 Aug;4(4):312-317. [doi: 10.1089/g4h.2014.0135] [Medline: 26182219]
14. Childhood obesity surveillance initiative. World Health Organization. 2016. URL: https://www.euro.who.int/__data/assets/pdf_file/0006/372426/WH14_COSI_factsheets_v2.pdf [accessed 2021-01-04]
15. Hardy LL, Ding D, Peralta LR, Mihrshahi S, Merom D. Association between sitting, screen time, fitness domains, and fundamental motor skills in children aged 5-16 years: cross-sectional population study. J Phys Act Health 2018 Dec 01;15(12):933-940. [doi: 10.1123/jpah.2017-0620] [Medline: 30404565]
16. Han A, Fu A, Cobley S, Sanders RH. Effectiveness of exercise intervention on improving fundamental movement skills and motor coordination in overweight/obese children and adolescents: a systematic review. J Sci Med Sport 2018 Jan;21(1):89-102. [doi: 10.1016/j.jsams.2017.07.001] [Medline: 28728887]
17. De Meester A, Stodden D, Brian A, True L, Cardon G, Tallir I, et al. Associations among elementary school children's actual motor competence, perceived motor competence, physical activity and BMI: a cross-sectional study. PLoS One 2016;11(10):e0164600 [FREE Full text] [doi: 10.1371/journal.pone.0164600] [Medline: 27736964]
18. De Meester A, Maes J, Stodden D, Cardon G, Goodway J, Lenoir M, et al. Identifying profiles of actual and perceived motor competence among adolescents: associations with motivation, physical activity, and sports participation. J Sports Sci 2016 Nov;34(21):2027-2037. [doi: 10.1080/02640414.2016.1149608] [Medline: 26928601]
19. Bardid F, De Meester A, Tallir I, Cardon G, Lenoir M, Haerens L. Configurations of actual and perceived motor competence among children: associations with motivation for sports and global self-worth. Hum Mov Sci 2016 Dec;50:1-9. [doi: 10.1016/j.humov.2016.09.001] [Medline: 27620643]
20. Luz C, Rodrigues LP, Meester AD, Cordovil R. The relationship between motor competence and health-related fitness in children and adolescents. PLoS One 2017;12(6):e0179993 [FREE Full text] [doi: 10.1371/journal.pone.0179993] [Medline: 28658292]

XSL•FO
RenderX

21. Stoner L, Rowlands D, Morrison A, Credeur D, Hamlin M, Gaffney K, et al. Efficacy of exercise intervention for weight loss in overweight and obese adolescents: meta-analysis and implications. Sports Med 2016 May 2:1737-1751. [doi: 10.1007/s40279-016-0537-6] [Medline: 27139723]
22. The power of physical activity. Exercise is Medicine - American College of Sport Medicine. URL: https://www.exerciseismedicine.org/support_page.php/physical-activity-health-impact/ [accessed 2020-08-16]
23. Beedie C, Mann S, Jimenez A, Kennedy L, Lane AM, Domone S, et al. Death by effectiveness: exercise as medicine caught in the efficacy trap!. Br J Sports Med 2016 Mar;50(6):323-324 [FREE Full text] [doi: 10.1136/bjsports-2014-094389] [Medline: 25807166]
24. Oh Y, Yang S. Defining exergames and exergaming. In: Proceedings of the Meaningful Play 2010 Conference. 2010 Presented at: Meaningful Play 2010 Conference; Oct. 21-13, 2010; East Lansing, MI, USA URL: https://meaningfulplay.msu.edu/proceedings2010/
25. Liang Y, Lau PW. Effects of active videogames on physical activity and related outcomes among healthy children: a systematic review. Games Health J 2014 Jun;3(3):122-144. [doi: 10.1089/g4h.2013.0070] [Medline: 26196172]
26. Biddiss E, Irwin J. Active video games to promote physical activity in children and youth: a systematic review. Arch Pediatr Adolesc Med 2010 Jul;164(7):664-672. [doi: 10.1001/archpediatrics.2010.104] [Medline: 20603468]
27. Bailey BW, McInnis K. Energy cost of exergaming: a comparison of the energy cost of 6 forms of exergaming. Arch Pediatr Adolesc Med 2011 Jul;165(7):597-602. [doi: 10.1001/archpediatrics.2011.15] [Medline: 21383255]
28. Comeras-Chueca C, Villalba-Heredia L, Pérez-Llera M, Lozano-Berges G, Marín-Puyalto J, Vicente-Rodríguez G, et al. Assessment of active video games' energy expenditure in children with overweight and obesity and differences by gender. Int J Environ Res Public Health 2020 Sep 15;17(18):6714 [FREE Full text] [doi: 10.3390/ijerph17186714] [Medline: 32942663]
29. De BJ, Perrier-Melo R, Melo DO, De SP, Da CC. Physical effort, energy expenditure, and motivation in structured and unstructured active video games: a randomized controlled trial. Hum Move 2016;17(3):198. [doi: 10.1515/humo-2016-0021]
30. Lamboglia CM, da Silva VT, de Vasconcelos Filho JE, Pinheiro MH, Munguba MC, Silva Júnior FV, et al. Exergaming as a strategic tool in the fight against childhood obesity: a systematic review. J Obes 2013;2013:438364 [FREE Full text] [doi: 10.1155/2013/438364] [Medline: 24319594]
31. Gao Z, Chen S. Are field-based exergames useful in preventing childhood obesity? A systematic review. Obes Rev 2014 Aug;15(8):676-691. [doi: 10.1111/obr.12164] [Medline: 24602072]
32. Gao Z, Chen S, Pasco D, Pope Z. A meta-analysis of active video games on health outcomes among children and adolescents. Obes Rev 2015 Sep;16(9):783-794. [doi: 10.1111/obr.12287] [Medline: 25943852]
33. Kari T. Can exergaming promote physical fitness and physical activity?: a systematic review of systematic reviews. Int J Gaming Comput Med Simulations 2014;6(4):77. [doi: 10.4018/ijgcms.2014100105]
34. Joronen K, Aikasalo A, Suvitie A. Nonphysical effects of exergames on child and adolescent well-being: a comprehensive systematic review. Scand J Caring Sci 2017 Sep;31(3):449-461. [doi: 10.1111/scs.12393] [Medline: 27862145]
35. Fu Y, Burns RD, Gomes E, Savignac A, Constantino N. Trends in sedentary behavior, physical activity, and motivation during a classroom-based active video game program. Int J Environ Res Public Health 2019 Aug 07;16(16):2821 [FREE Full text] [doi: 10.3390/ijerph16162821] [Medline: 31394855]
36. Andrade A, Correia CK, Coimbra DR. The psychological effects of exergames for children and adolescents with obesity: a systematic review and meta-analysis. Cyberpsychol Behav Soc Netw 2019 Nov;22(11):724-735. [doi: 10.1089/cyber.2019.0341] [Medline: 31697604]
37. Santos IK, Medeiros RC, Medeiros JA, Almeida-Neto PF, 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):1-13 [FREE Full text] [doi: 10.3390/ijerph18041641] [Medline: 33572159]
38. Fu Y, Burns RD, Constantino N, Zhang P. Differences in step counts, motor competence, and enjoyment between an exergaming group and a non-exergaming group. Games Health J 2018 Oct;7(5):335-340. [doi: 10.1089/g4h.2017.0188] [Medline: 29989430]
39. McGann J, Issartel J, Hederman L, Conlan O. Hop.Skip.Jump.Games: The effect of “principled” exergameplay on children's locomotor skill acquisition. Br J Educ Technol 2019 Nov 13;51(3):798-816. [doi: 10.1111/bjet.12886]
40. Gao Z, Zeng N, Pope ZC, Wang R, Yu F. Effects of exergaming on motor skill competence, perceived competence, and physical activity in preschool children. J Sport Health Sci 2019 Mar;8(2):106-113 [FREE Full text] [doi: 10.1016/j.jshs.2018.12.001] [Medline: 30997256]
41. Chen H, Sun H. Effects of active videogame and sports, play, and active recreation for kids physical education on children's health-related fitness and enjoyment. Games Health J 2017 Oct;6(5):312-318. [doi: 10.1089/g4h.2017.0001] [Medline: 28704072]
42. Christison AL, Evans TA, Bleess BB, Wang H, Aldag JC, Binns HJ. Exergaming for health: a randomized study of community-based exergaming curriculum in pediatric weight management. Games Health J 2016 Dec;5(6):413-421. [doi: 10.1089/g4h.2015.0097] [Medline: 27911621]

XSL•FO
RenderX

43. Fu Y, Burns RD. Effect of an active video gaming classroom curriculum on health-related fitness, school day step counts, and motivation in sixth graders. J Phys Act Health 2018 Sep 01;15(9):644-650. [doi: 10.1123/jpah.2017-0481] [Medline: 29741441]
44. Foley L, Jiang Y, Ni Mhurchu C, Jull A, Prapavessis H, Rodgers A, et al. The effect of active video games by ethnicity, sex and fitness: subgroup analysis from a randomised controlled trial. Int J Behav Nutr Phys Act 2014 Apr 03;11(1):46 [FREE Full text] [doi: 10.1186/1479-5868-11-46] [Medline: 24694082]
45. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M. Cochrane handbook for systematic reviews of interventions version 6.2. Cochrane Training. 2021. URL: https://training.cochrane.org/handbook/current [accessed 2021-03-05]
46. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Br Med J 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
47. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. Br Med J 2015 Jan 02;350:g7647 [FREE Full text] [doi: 10.1136/bmj.g7647] [Medline: 25555855]
48. Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. Br Med J 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
49. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. Br Med J 2016 Oct 12;355:i4919 [FREE Full text] [doi: 10.1136/bmj.i4919] [Medline: 27733354]
50. Van Biljon A, Longhurst G. The influence of exergaming on the functional fitness in overweight and obese children. Afr J Phys Act Health Sci 2012;18(4):991 [FREE Full text]
51. Maddison R, Foley L, Ni MC, Jiang Y, Jull A, Prapavessis H, et al. Effects of active video games on body composition: a randomized controlled trial. Am J Clin Nutr 2011 Jul;94(1):156-163 [FREE Full text] [doi: 10.3945/ajcn.110.009142] [Medline: 21562081]
52. Staiano AE, Marker AM, Beyl RA, Hsia DS, Katzmarzyk PT, Newton RL. A randomized controlled trial of dance exergaming for exercise training in overweight and obese adolescent girls. Pediatr Obes 2017 Dec;12(2):120-128 [FREE Full text] [doi: 10.1111/ijpo.12117] [Medline: 26918815]
53. Wagener TL, Fedele DA, Mignogna MR, Hester CN, Gillaspy SR. Psychological effects of dance-based group exergaming in obese adolescents. Pediatr Obes 2012 Oct;7(5):68-74. [doi: 10.1111/j.2047-6310.2012.00065.x] [Medline: 22767495]
54. Staiano AE, Beyl RA, Guan W, Hendrick CA, Hsia DS, Newton RL. Home-based exergaming among children with overweight and obesity: a randomized clinical trial. Pediatr Obes 2018 Nov;13(11):724-733 [FREE Full text] [doi: 10.1111/ijpo.12438] [Medline: 30027607]
55. Trost SG, Sundal D, Foster GD, Lent MR, Vojta D. Effects of a pediatric weight management program with and without active video games a randomized trial. JAMA Pediatr 2014 May;168(5):407-413. [doi: 10.1001/jamapediatrics.2013.3436] [Medline: 24589566]
56. Staiano AE, Abraham AA, Calvert SL. Adolescent exergame play for weight loss and psychosocial improvement: a controlled physical activity intervention. Obesity (Silver Spring) 2013 Mar;21(3):598-601 [FREE Full text] [doi: 10.1002/oby.20282] [Medline: 23592669]
57. Maloney AE, Threlkeld KA, Cook WL. Comparative effectiveness of a 12-week physical activity intervention for overweight and obese youth: exergaming with "Dance Dance Revolution". Games Health J 2012 Apr;1(2):96-103. [doi: 10.1089/g4h.2011.0009] [Medline: 26193183]
58. Adamo KB, Rutherford JA, Goldfield GS. Effects of interactive video game cycling on overweight and obese adolescent health. Appl Physiol Nutr Metab 2010 Dec;35(6):805-815. [doi: 10.1139/H10-078] [Medline: 21164552]
59. Maddison R, Mhurchu CN, Jull A, Prapavessis H, Foley LS, Jiang Y. Active video games: the mediating effect of aerobic fitness on body composition. Int J Behav Nutr Phys Act 2012 May 03;9:54 [FREE Full text] [doi: 10.1186/1479-5868-9-54] [Medline: 22554052]
60. Goldfield GS, Adamo KB, Rutherford J, Murray M. The effects of aerobic exercise on psychosocial functioning of adolescents who are overweight or obese. J Pediatr Psychol 2012;37(10):1136-1147. [doi: 10.1093/jpepsy/jss084] [Medline: 23027721]
61. Bonney E, Ferguson G, Burgess T, Smits-Engelsman B. Benefits of activity-based interventions among female adolescents who are overweight and obese. Pediatr Phys Ther 2019 Oct;31(4):338-345. [doi: 10.1097/PEP.0000000000000636] [Medline: 31568378]
62. Irandoust K, Taheri M, H'mida C, Neto GR, Trabelsi K, Ammar A, et al. Exergaming and aquatic exercises affect lung function and weight loss in obese children. Int J Sports Med 2021 Jun;42(6):566-572. [doi: 10.1055/a-1289-9307] [Medline: 33176381]
63. Huang H, Cook L, Harrison J, O'Keefe N, Schultz K, Curtis D. Exercise by playing active video gaming in overweight and obese children. J Phys Ther Sport Med 2017;1(1):1000107. [doi: 10.35841/physical-therapy.1.1.25-31]
64. Duman F, Kokaçya MH, Doğru E, Katayıfcı N, Canbay O, Aman F. The role of active video-accompanied exercises in improvement of the obese state in children: a prospective study from Turkey. J Clin Res Pediatr Endocrinol 2016 Sep 01;8(3):334-340 [FREE Full text] [doi: 10.4274/jcrpe.2284] [Medline: 27087547]

XSL•FO
RenderX

65. Calcaterra V, Larizza D, Codrons E, De Silvestri A, Brambilla P, Abela S, et al. Improved metabolic and cardiorespiratory fitness during a recreational training program in obese children. J Pediatr Endocrinol Metab 2013;26(3-4):271-276. [doi: 10.1515/jpem-2012-0157] [Medline: 23327787]
66. Christison A, Khan HA. Exergaming for health: a community-based pediatric weight management program using active video gaming. Clin Pediatr (Phila) 2012 Apr;51(4):382-388. [doi: 10.1177/0009922811429480] [Medline: 22157430]
67. Argarini R, Herawati L, Irwadi I, Arum CP, Meinar SG. Fundamental movement skill approach to combat childhood obesity in Surabaya, Indonesia: potential effects of video games based exercises (Exergaming). Talent Dev Excell 2020;12(1):3026-3034 [FREE Full text]
68. Gao Z, Zeng N, McDonough DJ, Su X. A systematic review of active video games on youth's body composition and physical activity. Int J Sports Med 2020 Aug;41(9):561-573. [doi: 10.1055/a-1152-4959] [Medline: 32375184]
69. Hernández-Jiménez C, Sarabia R, Paz-Zulueta M, Paras-Bravo P, Pellico A, Ruiz Azcona L, et al. Impact of active video games on body mass index in children and adolescents: systematic review and meta-analysis evaluating the quality of primary studies. Int J Environ Res Public Health 2019 Jul 08;16(13):2424 [FREE Full text] [doi: 10.3390/ijerph16132424] [Medline: 31288460]
70. Zeng N, Gao Z. Exergaming and obesity in youth: current perspectives. Int J Gen Med 2016;9:275-284 [FREE Full text] [doi: 10.2147/IJGM.S99025] [Medline: 27536158]
71. LeBlanc AG, Chaput J, McFarlane A, Colley RC, Thivel D, Biddle SJ, et al. Active video games and health indicators in children and youth: a systematic review. PLoS One 2013;8(6):e65351 [FREE Full text] [doi: 10.1371/journal.pone.0065351] [Medline: 23799008]
72. Savva SC, Tornaritis M, Savva ME, Kourides Y, Panagi A, Silikiotou N, et al. Waist circumference and waist-to-height ratio are better predictors of cardiovascular disease risk factors in children than body mass index. Int J Obes Relat Metab Disord 2000 Nov;24(11):1453-1458. [doi: 10.1038/sj.ijo.0801401] [Medline: 11126342]
73. Van Brussel M, Bongers BC, Hulzebos EH, Burghard M, Takken T. A systematic approach to interpreting the cardiopulmonary exercise test in pediatrics. Pediatr Exerc Sci 2019 May 01;31(2):194-203. [doi: 10.1123/pes.2018-0235] [Medline: 30691356]
74. Armstrong N, Welsman J, Winsley R. Is peak VO2 a maximal index of children's aerobic fitness? Int J Sports Med 1996 Jul;17(5):356-359. [doi: 10.1055/s-2007-972860] [Medline: 8858407]
75. Tomlinson DJ, Erskine RM, Morse CI, Winwood K, Onambélé-Pearson G. The impact of obesity on skeletal muscle strength and structure through adolescence to old age. Biogerontology 2016 Jun;17(3):467-483 [FREE Full text] [doi: 10.1007/s10522-015-9626-4] [Medline: 26667010]
76. Faigenbaum AD, Rebullido TR, MacDonald JP. Pediatric inactivity triad: a risky PIT. Curr Sports Med Rep 2018 Feb;17(2):45-47. [doi: 10.1249/JSR.0000000000000450] [Medline: 29420346]
77. Cheng J, East P, Blanco E, Sim EK, Castillo M, Lozoff B, et al. Obesity leads to declines in motor skills across childhood. Child Care Health Dev 2016 May;42(3):343-350 [FREE Full text] [doi: 10.1111/cch.12336] [Medline: 27059409]

Abbreviations

AVG: active video game
CRF: cardiorespiratory fitness
DDR: Dance Dance Revolution
MD: mean difference
PA: physical activity
PICOS: Population, Intervention, Comparison, Outcomes and Study
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
ROBINS-I: Risk of Bias in Nonrandomized Studies-of Interventions

Edited by G Eysenbach; submitted 28.04.21; peer-reviewed by N Maglaveras; comments to author 30.05.21; revised version received 11.06.21; accepted 02.08.21; published 18.10.21.

Please cite as:
Comeras-Chueca C, Marin-Puyalto J, Matute-Llorente A, Vicente-Rodriguez G, Casajus JA, Gonzalez-Aguero A
Effects of Active Video Games on Health-Related Physical Fitness and Motor Competence in Children and Adolescents With Overweight or Obesity: Systematic Review and Meta-Analysis
JMIR Serious Games 2021;9(4):e29981
URL: https://games.jmir.org/2021/4/e29981
doi:10.2196/29981
PMID:34661549

XSL•FO
RenderX

©Cristina Comeras-Chueca, Jorge Marin-Puyalto, Angel Matute-Llorente, German Vicente-Rodriguez, Jose Antonio Casajus, Alex Gonzalez-Aguero. Originally published in JMIR Serious Games (https://games.jmir.org), 18.10.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An Interactive Physical-Cognitive Game-Based Training System Using Kinect for Older Adults: Development and Usability Study

Teerawat Kamnardsiri[1,2*], BSc, MSc, PhD; Kochaphan Phirom[3], BSc, MSc; Sirinun Boripuntakul[3*], BSc, MSc, PhD; Somporn Sungkarat[3*], BSc, MSc, PhD

[1]Research Group of Modern Management and Information Technology, College of Arts, Media and Technology, Chiang Mai University, Chiang Mai, Thailand

[2]Department of Digital Game, College of Arts, Media and Technology, Chiang Mai University, Chiang Mai, Thailand

[3]Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand

[*]these authors contributed equally

Corresponding Author:
Somporn Sungkarat, BSc, MSc, PhD
Department of Physical Therapy
Faculty of Associated Medical Sciences
Chiang Mai University
9th Floor, Faculty of Associated Medical Sciences
110 Intawaroros Rd, Sripoom
Chiang Mai, 50200
Thailand
Phone: 66 53936042
Fax: 66 53935072
Email: somporn.sungkarat@cmu.ac.th

Abstract

Background: Declines in physical and cognitive functions are recognized as important risk factors for falls in older adults. Promising evidence suggests that interactive game-based systems that allow simultaneous physical and cognitive exercise are a potential approach to enhance exercise adherence and reduce fall risk in older adults. However, a limited number of studies have reported the development of a combined physical-cognitive game-based training system for fall risk reduction in older adults.

Objective: The aim of this study is to develop and evaluate the usability of an interactive physical-cognitive game-based training system (game-based exercise) for older adults.

Methods: In the development phase (Part I), a game-based exercise prototype was created by integrating knowledge and a literature review as well as brainstorming with experts on effective fall prevention exercise for older adults. The output was a game-based exercise prototype that covers crucial physical and cognitive components related to falls. In the usability testing (Part II), 5 games (ie, Fruits Hunter, Where Am I?, Whack a Mole, Sky Falls, and Crossing Poison River) with three difficulty levels (ie, beginner, intermediate, and advanced levels) were tested in 5 older adults (mean age 70.40 years, SD 5.41 years). After completing the games, participants rated their enjoyment level while engaging with the games using the Physical Activity Enjoyment Scale (PACES) and commented on the games. Descriptive statistics were used to describe the participants' characteristics and PACES scores.

Results: The results showed that the average PACES score was 123 out of 126 points overall and between 6.66 and 7.00 for each item, indicating a high level of enjoyment. Positive feedback, such as praise for the well-designed interactions and user-friendly interfaces, was also provided.

Conclusions: These findings suggest that it is promising to implement an interactive, physical-cognitive game-based exercise in older adults. The effectiveness of a game-based exercise program for fall risk reduction has yet to be determined.

(JMIR Serious Games 2021;9(4):e27848) doi:10.2196/27848

KEYWORDS

digital game; interactive game-based training; physical-cognitive training; exergaming; Kinect sensors; older adults; falls; PACES; user-centered design; game-based exercise

XSL•FO
RenderX

Introduction

Declines in multiple physiological systems with ageing contribute to balance and gait deficits, leading to an increased risk of fall [1]. Fall is a serious public health problem, and its consequences have a marked adverse impact on physical and psychological aspects such as injuries, activity restriction, fear of falling, and loss of autonomy [2-4]. Given the substantial impact of falls on health, as well as their medical and economic burden, an effective strategy to prevent falls in older adults is warranted.

Several investigators have consistently reported a strong positive effect of physical exercise on fall prevention among older adults [5-9]. Researchers have also identified a critical role of cognition, especially executive function, attention, and memory, on balance and gait control [10-13]. Many examiners have demonstrated that cognitive training, which is an intervention program aimed at improving, maintaining, or restoring cognitive function via the repeated and structured practice of tasks, can improve balance and gait and reduce fall risk [14-17]. Taken together, incorporating a cognitive component into physical exercise may augment its benefits in fall prevention [11-13,18-20]. A growing number of investigators have documented the effects of combined physical-cognitive exercise training in a simultaneous form (dual-tasking) among older adults. Combined physical-cognitive exercise training programs have resulted in greater improvement in physical and cognitive performance than either type of single training alone [21-25].

As technology advances, a new alternative of rehabilitation approach targeting training of various physical and cognitive components in the form of interactive game-based exercises (exergames) is becoming available [26]. These interactive game-based exercises use technology-driven platforms that require users to move their body in order to complete assigned tasks via video game interface elements [27]. The interactive game-based exercises have advantages in terms of gamification features. Researchers have demonstrated that exergames are attractive because they provide real-time interaction and feedback to users, which enhances motivation and training adherence [28,29]. In addition, game-based exercises allow users to be active with repetitive practice and track progression, which is beneficial for training outcomes [30]. Another advantage of exergames is that they can offer an experience for daily-life task requirements based on concurrent training of physical and cognitive components [31-33]. Moreover, game-based exercises can be applied either in rehabilitation centers or community and home settings [33,34]. Several researchers have shown that game-based training using Nintendo Wii Fit, and Microsoft Kinect sensors were effective in improving physical abilities (eg, balance, gait performance), improving cognitive abilities (eg, executive function, speed of processing), and reducing the risk of falls among older adults [35-42]. However, in most existing training programs available for fall prevention, research concerning simultaneous training of physical and cognitive functions (dual tasking) with the use of exergames has remained scarce.

Among interactive game-based technology for training, the Kinect motion sensor (Microsoft Corporation) has been considered as a high-potential approach because it provides a markerless full-body 3D motion tracker and enables users to virtually interact hands-free with a computer system. Additionally, several examiners have demonstrated that among interactive game-based technology for exercising, Microsoft Kinect has an advantage in that it allows individuals to interact with games using their own body in a natural way [43,44]; this enhances the natural form of human-computer interaction [45]. In addition, the Microsoft Kinect motion sensor is an accurate input device; thus, it allows precise tracking and real-time feedback of user performance [46]. The Microsoft Kinect sensor is proposed to be a feasible and effective tool for training concurrent physical and cognitive components in older adults, with the aim of reducing the intrinsic causes (ie, physical and cognitive) of falls [47-51].

Based on the usability challenges faced by older adults, programmers should develop user interfaces that are user-friendly for older adults and specific for training purposes. A previous investigator has suggested that older adults accept innovative technology when they recognize its benefit and find it meaningful for their lives [52].

Therefore, this study focused on developing a prototype of a game-based exercise that accounts for the target user's expectation and requirement (eg, enjoyment, attractiveness, user skill-challenge balance, and benefit for training cognitive and physical components). In particular, following a user-centered design approach, acceptability and training adherence in future empirical studies were established.

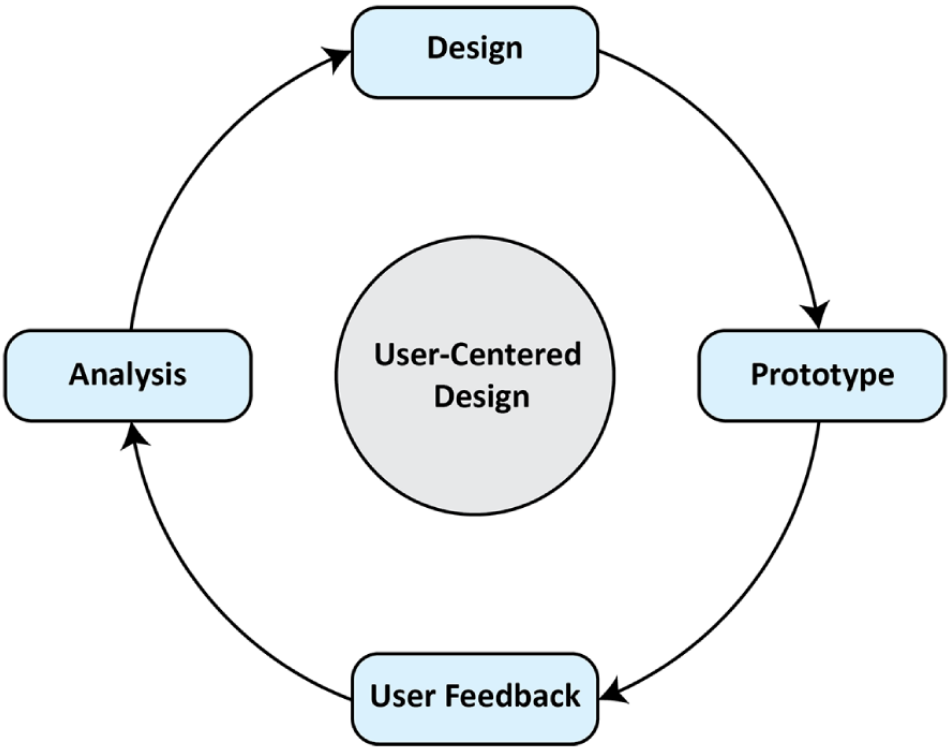
Methods

Study Design

The present study consists of two parts: (1) the development of a game-based exercise prototype, and (2) evaluation of target users' feedback, response, and satisfaction regarding the developed game-based exercise prototype. The concept of development and evaluation of the target user's experience for producing a final prototype followed a user-centered design (UCD) approach [27,46,53-62]. UCD is a user interface design process in which designers focus on gaining a target user's perspective to create a product with a high degree of usability. The UCD cycle is depicted in Figure 1.

XSL•FO
RenderX

Figure 1. Schematic of the user-centered design cycle [27].



Part I: Development of a Game-Based Exercise Prototype

Development Process

In this part, the 4-phase UCD process was applied [27,46,53-62]. The first phase, the design development process of a game-based exercise prototype, was conducted in the brainstorming phase. A total of 7 team members, including 3 physical therapists and 1 physician (3-20 years of experience in geriatric and cognitive rehabilitation), 2 game programmers (5 years of experience in the Unity 3D game engine), and 1 game designer (10 years of experience in game design and game theory), participated in the brainstorming session. This phase involved generating potential core game ideas by integrating the knowledge and literature review of previous physical and cognitive training programs and interactive exergame interventions for fall prevention in older adults [16,63-65]. In the present study, the core game was composed of two training elements: (1) a physical element, including stepping and balance training, and (2) a cognitive element related to balance and falls in older adults, including executive function, attention, and memory [18,66,67]. To ensure that the difficulty of the games was appropriate for each individual user, the progression of game difficulty was considered. The game-based training program underwent critical appraisal by the physical therapists and physician. In the second phase, after consensus, proven game ideas were used to create a game-based digital exercise game prototype using the Unity 3D game engine software with Kinect Sensor V2 for Windows. In the third phase, user feedback to improve the game-based exercise prototype was provided by end users using a think-aloud method [68]. Finally, the fourth phase concentrated on assessing physical activity enjoyment during game engagement. The Physical Activity Enjoyment Scale (PACES) questionnaire [69,70], an 18-item scale questionnaire, was used to analyze physical activity enjoyment. Moreover, the usability of the game-based exercise prototype was determined with a structured interview (feedback about the game, themes, user interface, sound effect, graphics, and interaction). Feedback from all participants was considered by the research team to improve the game-based exercise prototype.

Characteristics of the Game-Based Exercise Prototype

The characteristics of the game-based exercise prototype are presented in Table 1. The game can be described as individual interactive game-based training using Kinect. The prototype of the game-based exercise comprised 5 games, including (1) Fruits Hunter, (2) Where Am I ?, (3) Whack a Mole, (4) Sky Falls, and (5) Crossing Poison River. The games had three different levels: beginner, intermediate, and advanced. The level of game complexity progressed by increasing the difficulty of physical demand (ie, movement speed, distance, duration, base of support) and cognitive demand (ie, number of stimuli, complexity of the game's rules, and amount of cognitive load). The estimated play time was 45 to 60 minutes.

XSL•FO
RenderX

Table 1. Summary of the characteristics of the developed game-based exercise prototype.

Characteristic	Description
Basic characteristics	
Health topic	A game-based exercise prototype
Targeted age group	Older adults (age ≥65 years)
Short description of the game idea	The game-based exercise is a virtual, interactive game-based training system using Microsoft Kinect motion sensor technology. The game-based exercise comprises 5 games that include physical and cognitive components associated with balance and falls in older adults.
Target player	Individual
Behavior change procedure used	A game-based exercise is used to enhance motivation and engagement in older adults.
Clinical support needed	Physical therapist and geriatric physicians
Data shared with clinician	Data are saved and stored in the hard disk. However, reaction time, error, and score are given as feedback on the display screen (ie, the rubber mat) at the end of each game.
Type of game	Physical, action, real-time strategy
Game components	
Player's game goal/objective	
Physical components	• Improve static and dynamic balance • Improve stepping reaction and response time • Improve lower limb muscle strength
Cognitive components	• Fruits Hunter: improves response ability and speed of processing via a stepping task. • Where am I ?: improves semantic memory and visuospatial ability via visual sense • Whack a Mole: improves selective attention ability, visual attention performance, speed of processing, and inhibition ability • Sky Falls: improves sequencing and planning ability • Crossing Poison River: improve episodic memory via auditory sense

Rules	• Fruits Hunter: step on the presented fruits as fast as possible within a limited time.
	• Where am I ?: step to the presented objects and remember as many as of them possible. The recall questions are provided at the end of the game.
	• Whack a Mole: respond correctly to different rules of this game as follows:
	  • Mole or rabbit: steps on the target 1 time
	  • Mole or rabbit with helmet: steps on the target 2 times
	  • Bomb: do not step on the target
	• Sky Falls: step with alternating feet to collect as many dropping objects in the basket as possible
	• Crossing Poison River: listen to a short story and remember the content of the story while standing on one leg

Game mechanics	The game-based exercise system allows users to interact with the virtual games by stepping on the presented targets in different directions in pursuit of the game's goals. The game-based exercise also provides audio and visual feedback to the users while they are playing the games.
Virtual environment	A forest with fruits, animals, vegetables, and a river
Setting	The game-based exercise can be set in a room environment
Device requirements	Personal computer/notebook/laptop with LED projector
Sensors used	Microsoft Kinect Sensor V2
Estimated play time	45-60 minutes

Part II: Evaluation of Target User's Experience

Recruitment and Participants

A total of 5 community-dwelling older adults were enrolled as representative target users. The inclusion criteria were (1) age 65 years or older, (2) had normal cognitive function (determined by a Mental State Examination T10 [71] score ≥24 points or depending on the level of education), (3) ability to walk without an assistive device for at least 10 m, and (4) ability to step in all directions independently and safely. Exclusion criteria were (1) depressive symptoms (determined by a Thai Geriatric Depression Scale-15 [72] score >6 points), (2) orthopedic deficits, neurological deficits, and/or other significant health problems that precluded the participant from completing the testing protocol, and (3) uncorrected visual and hearing

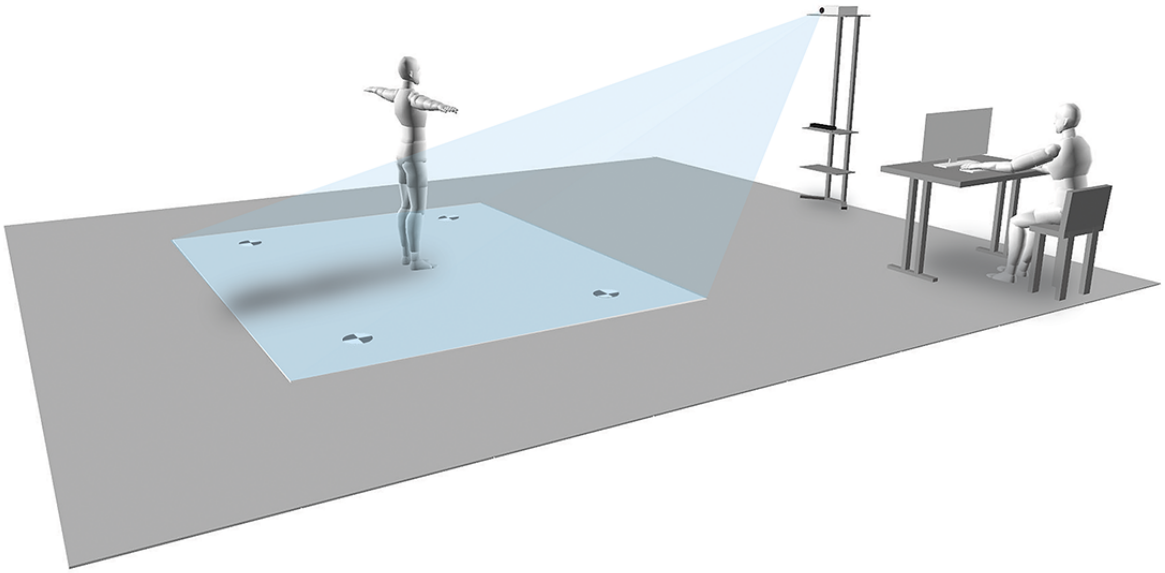
XSL•FO
RenderX

impairment. The study protocol was approved by the Human Ethical Review Board of the principal investigator's institute (AMSEC-61EX-078). All participants gave written informed consent prior to participating in the study. The demographic data of the participants, which consisted of age, height, weight, medication use, and history of falls in the previous 12 months, were recorded.

Hardware Configuration

To set up the system, the capture volume of the system was configured using the three main devices: Kinect Sensor V2 [73], LED projector, and laptop computer. In the present study, the Kinect Sensor V2 was used because it provides greater precision and more stable results compared to the Kinect Sensor V1 [74]. The Kinect Sensor V2 is a depth sensor camera manufactured by Microsoft that provides information about the depth, color, and skeleton of a user who is standing in front of the sensor. The Kinect sensor and LED projector were set on a portable metal storage rack at a height of 0.8 m and 2.0 m from the floor, respectively. The laptop computer that contained the developed game software was set near the Kinect sensor. The game was projected on a rubber mat (2.0 m width × 1.2 m height) that was placed on the floor; thus, the participants could virtually interact by stepping. The center of the rubber mat was set 2.5 m from the Microsoft Kinect sensor and LED projector. The configuration of hardware for playing the game-based exercise is illustrated in Figure 2.

Figure 2. Environment configuration of the game-based exercise system.



Protocol

The game-based exercise was connected with the Microsoft Kinect sensor, an LED projector, and the laptop computer. After that, the system was calibrated by moving four markers in the game-based training system over the four corner marks on the rubber mat. For individualized body position calibration, each participant was asked to perform a T-pose stand for approximately 5 seconds at the center of the rubber mat (Figure 2). After the calibration process, each participant received a comprehensive description of the rules of the games, including a demonstration, and was requested to play the game-based exercise. The American College Sport of Medicine exercise guidelines recommend that older adults should participate in aerobic activities for a minimum of 30 minutes per session to promote and maintain their health-related outcomes [75]. In this study, the exercise duration ranged between 30 and 50 minutes, with a rest interval of between 10 and 15 minutes, depending on the participant performance; this resulted in a total time of 45 to 60 minutes.

After completing the game-based exercise, participants were asked to rate their enjoyment using the PACES questionnaire [69,70]. The PACES is an 18-item scale questionnaire that assesses physical activity enjoyment during game engagement with a 7-point Likert scale (1, strongly disagree, to 7, strongly agree). A higher PACES score reflects a greater level of enjoyment. Moreover, using a structured interview, participants were interviewed about their impressions of the game-based exercise features in terms of rules, mechanics, interfaces, and scoring, as well as their physical and cognitive involvement while playing the games.

Data Analysis

Descriptive statistics were used to describe both the participants' characteristics and their scores on the 18-item PACES questionnaire. All data were analyzed using SPSS 21.0 (IBM Corporation).

XSL•FO
RenderX

Results

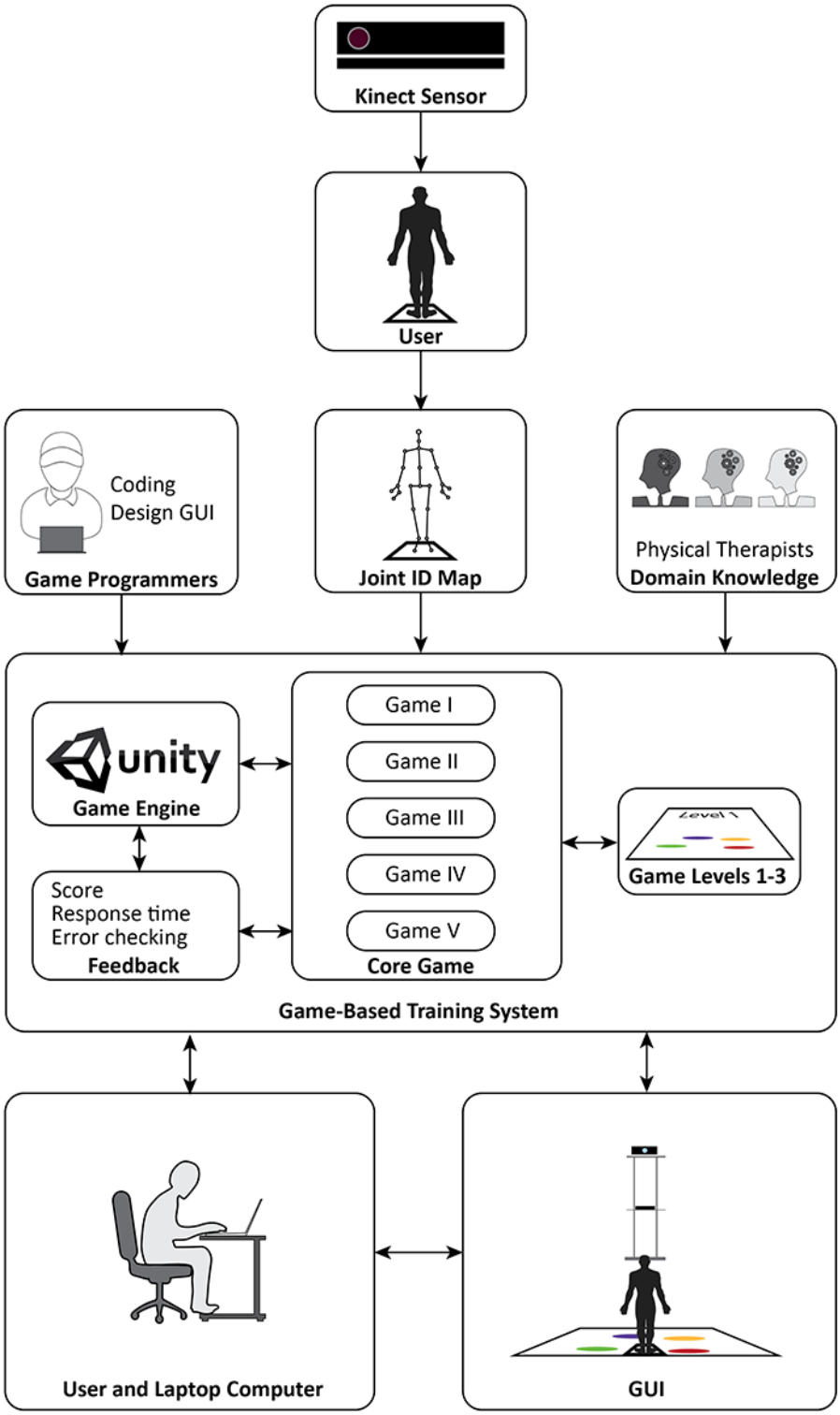
Part I: Development of a Game-Based Exercise Prototype

The framework of the game-based exercise comprised 6 components (Figure 3), including:

1. The Microsoft Kinect sensor: the depth sensor that was used to track and monitor full-body movements in 3D coordinates (ie, the x-, y-, and z-axes). The tracking data were then converted to the 24 points of the Joint ID Map (body skeleton model). In this study, 4 points of the Joint ID Map [43] (ANKLE_RIGHT, FOOT_RIGHT, ANKLE_LEFT and FOOT_LEFT) were used for the interaction between the user and the game.
2. Game programmers: the specialists who generated the digital game using the computer programming language and game engine software (the Unity 3D game engine software with Microsoft Kinect sensor V2 for Windows).
3. Domain knowledge: the experts having core knowledge and experience of physical and cognitive training programs for fall prevention in older adults.
4. Game-based training system: the digital game system, which comprised 5 games (Games I-V) with 3 levels (levels 1-3) and feedback (ie, score, response time, and error).
5. User and laptop computer: a system operator who was responsible for controlling the game-based training system while participants were playing the games.
6. Graphical user interface: a form of user interface that allows participants to interact with the game-based exercise (Figure 4). An example of a participant training with the game-based exercise prototype is displayed in Figure 5.

XSL•FO
RenderX

Figure 3. Framework of the game-based exercise system. GUI: graphical user interface.



XSL•FO
RenderX

Figure 4. Screenshots of the 5 games in the game-based exercise system: (A) Fruits Hunter, (B) Where Am I ?, (C) Whack a Mole, (D) Sky Falls, and (E) Crossing Poison River.

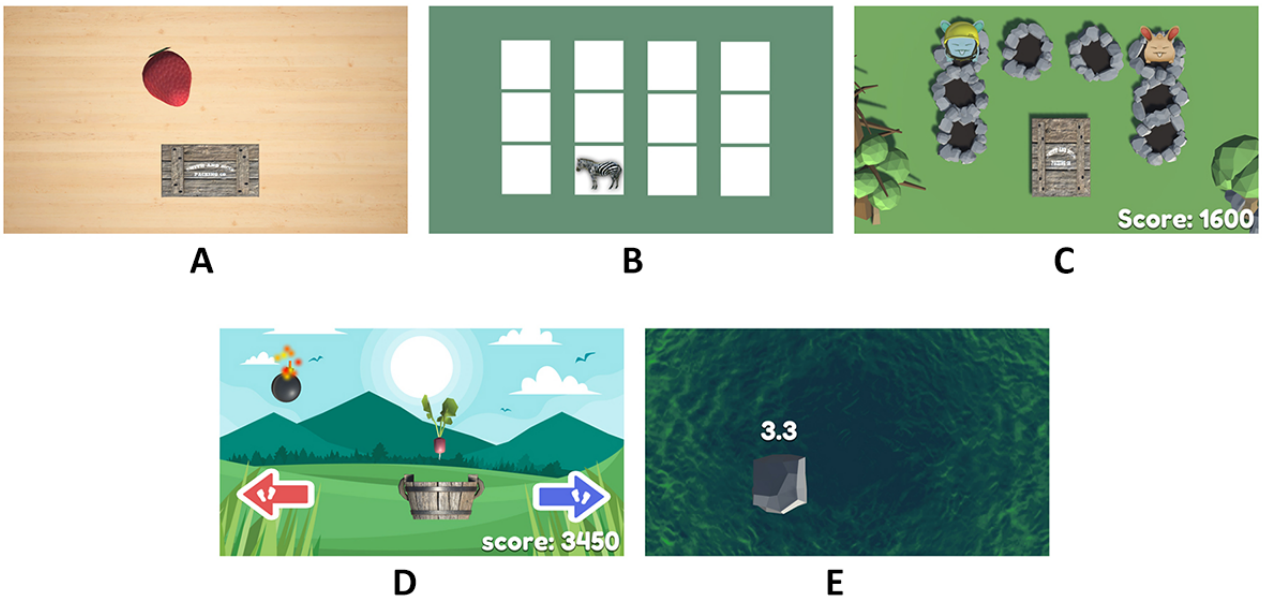
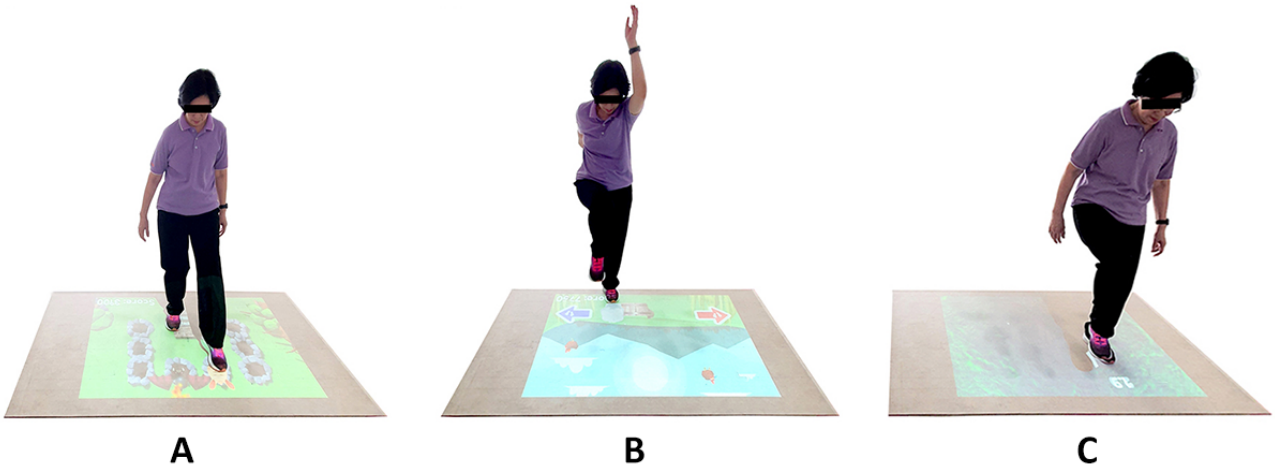


Figure 5. Examples of a participant playing the game-based exercise: (A) Whack a Mole, (B) Sky Falls, and (C) Crossing Poison River.



Part II: Evaluation of the Target Users' Experience

Participant Characteristics

A total of 5 community-dwelling older adults participated in the usability testing phase. Their mean age was 70.40 (SD 5.41) years (range 65-79 years). No participants had any experience with using exergames. They had low incidence rate of falls in the past 12 months, and they either did not take medication or took only one type. The participants' characteristics are summarized in Table 2.

Table 2. Characteristics of the study participants.

Characteristic	Value	Median	Range
Age (years), mean (SD)	70.40 (5.41)	68	65-79
Height (cm), mean (SD)	155.20 (6.06)	156	149-164
Weight (kg), mean (SD)	53.40 (8.62)	55	39-60
BMI (kg/m^2), mean (SD)	22.08 (2.63)	22.21	17.60-24.30
Education (years), mean (SD)	14.40 (5.90)	16	4-18
Types of medication, mean (SD)	0.20 (0.45)	0	0-1
Falls in the past year, n	1	0	0-1

XSL•FO
RenderX

User Experience in Using the Game-Based Exercise System

All 5 participants completed the game-based exercise and answered the PACES questionnaire. The average score on each item was between 6.66 and 7.00, which indicated greater levels of enjoyment. The 18-item PACES scores are illustrated in Table 3.

Table 3. Physical Activity Enjoyment Scale (PACES) rating scores (n=5). All items were rated on a 7-point scale from 1, strongly disagree, to 7, strongly agree.

Question	Participants					Response rating, mean (SD)
	S01	S02	S03	S04	S05	
1. I enjoy it; I hate it	7	6	7	7	6	6.80 (0.20)
2. I feel interested; I feel bored	7	6	7	7	7	7.00 (0.00)
3. I like it; I dislike it	7	7	7	7	7	7.00 (0.00)
4. I find it pleasurable; I find it unpleasurable	7	7	7	6	6	6.66 (0.24)
5. I am very absorbed in this activity; I am not at all absorbed in this activity	6	7	7	7	7	6.80 (0.20)
6. It's a lot of fun; it's no fun at all	7	7	7	7	7	7.00 (0.00)
7. I find it energizing; I find it tiring	7	6	7	7	7	6.80 (0.20)
8. It makes me happy; it makes me depressed	7	7	7	7	7	7.00 (0.00)
9. It's very pleasant; it's very unpleasant	7	7	7	7	6	6.80 (0.20)
10. I feel good physically while doing it; I feel bad physically while doing it	7	7	7	7	7	7.00 (0.00)
11. It's very invigorating; it's not at all invigorating	7	7	7	6	7	6.80 (0.20)
12. I am not at all frustrated by it; I am very frustrated by it	7	7	7	7	7	7.00 (0.00)
13. It's very gratifying; it's not at all gratifying	7	7	7	6	6	6.66 (0.24)
14. It's very exhilarating; it's not at all exhilarating	7	7	7	7	6	6.80 (0.20)
15. It's very stimulating; it's not at all stimulating	7	7	7	6	7	6.80 (0.20)
16. It give me a strong sense of accomplishment; it does not give me any sense of accomplishment	7	7	7	6	7	6.80 (0.20)
17. It's very refreshing; it's not at all refreshing	7	7	6	6	7	6.66 (0.24)
18. I felt as though there was nothing else I would rather be doing; I felt as though I would rather be doing something else	7	7	7	7	6	6.80 (0.20)
Rating scale of all items (total points: 126)	125	123	125	120	120	123.00 (1.26)

User Feedback and Suggestions

The feedback and suggestions provided by the users are presented in Table 4.

XSL•FO
RenderX

Table 4. Feedback and suggestions provided by the study participants for the game prototype.

Type of feedback or suggestion	Feedback
Feedback	
Positive	• "The games' feature and appearance were very attractive and enhanced my motivation to complete the games." • "The games provided my performance with visible outcomes and scores which motivated me to try harder to get a better score in the next trial." • "The difficultly of each game was optimal; it was not too easy and not too difficult." • "The games had variety of forms and rules that challenged my physical and cognitive abilities." • "The games had a meaningful sound effect which helped me to identify my right or wrong responses." • "The games' systems were quite simple to set up and easy to manage, thus it appeared to be feasible to use in the community or home settings."
Negative	• "In the Whack a Mole, sometimes I did not step on the bomb, but it eventually blew up." • "In the Sky Falls, sometimes it was quite hard to control the movement of the bamboo basket even though I tried to alternate my stepping rhythmically."
Suggestions	
Game rules	• Clearly state the game instructions and rules at the beginning
Level design	• Reduce the speed of dropping objects in the beginner level of Sky Falls • Use different types of animals and vegetables for each difficulty level of Where Am I?
Graphics/look and feel	• Adjust the distance of each presented object in Whack a Mole
Audio	• Increase the display volume • Use different background music for each game

Discussion

Principal Findings

In this study, we aimed to develop and test the usability of a virtual, interactive game-based training system that is focused on simultaneously training the physical and cognitive function (dual-tasking) of community-dwelling older adults. The core games were formulated by integrating the principal knowledge and existing evidence from the literature related to effective fall prevention exercise programs for older adults as well as by subjecting the games to critical appraisal from experts. We also assessed older adults' experiences in terms of enjoyment and game features using the PACES questionnaire and a structured interview.

Several intervention studies have reported high dropout rates and limited use of technology-supported platforms for delivering exercise training programs [76-78]. Interactive game-based training may be considered as a more efficient approach for empowering user engagement, which contributes to positive outcomes of an intervention. However, many older adults tend to be less engaged with modern digital technology, and not all are accepting of it. To overcome this limitation, the UCD approach, which incorporates game design principles (ie, goals, rules, feedback, points, time, reward structures, levels, and aesthetics), was used in the process of designing and developing a game-based exercise prototype with an aim to motivate and engage older adults in exercising [54-60]. Researchers have identified the potential benefits of using the UCD concept for developing exergames for older adults with and without health-related problems. For example, Hemingway et al [53] used the UCD process to develop a mobile game to influence the behavior of HIV service uptake among a key population. Lange et al [27] established an interactive game-based program constructed on a UCD design process for training the dynamic balance of individuals who have previously experienced a stroke. Howes et al [79] also used UCD to develop the bespoke Active Computer Gaming system to deliver strength and balance exercise programs for older adults. Together, the present and previous findings consistently suggest that the UCD approach is a core process that should be embedded in health games to ensure usability and acceptability. Therefore, the target users may benefit fully from exergames technology.

To our knowledge, our game-based exercise is the first game prototype that was mainly designed to support older adults in combined physical-cognitive exercising using the Microsoft Kinect motion sensor. In particular, the game-based exercise focused on the core impairment aspects that are related to falls in older adults, including balance and stepping performance as well as executive function, attention, and memory. Several investigators have consistently reported that the most important component of exercise programs for fall prevention is balance training [6,7]. In addition, stepping training, a form of highly specific balance training, has shown to be an effective fall prevention strategy [80]. Thus, balance and stepping training were included in the game-based exercise. Regarding cognition, declined executive function, attention, and memory have been identified as crucial contributors to falls [11-13,19]. Thus, adding these cognitive components to physical training may potentially enhance the efficacy of fall prevention programs for older adults. In addition, age-related perception and sensation decline in older adults were considered. Therefore, visual and

XSL•FO
RenderX

audio presentations, such as the size and distance of target objects as well as the level of sound volume, were included. Findings from the study demonstrated that the target users viewed the game-based exercise prototype as an enjoyable and practical intervention approach for their physical and cognitive training at home and in community settings. This may be because the development of the game-based exercise prototype incorporated the current knowledge regarding the key contributing factors for training continuation by using exergames among older adults. These factors gradually increase the level of game difficulty, provide clearer feedback, and offer a simple setup [81]. In this way, we expected that the newly developed exergames would overcome the barrier to exercise in older adults. This enjoyment (determined by PACES scores) and positive feedback from the users may be, at least in part, due to the fundamental elements of the game-based exercise, which feature a real-time interface display and feedback. Consistent with previous studies, our games system provides feedback, including scores and performance outcomes (ie, response time, error), which enhances the motivation of the users [82-84]. Moreover, the positive responses from older learners who are unfamiliar with new technologies can be attributed to the design features of the game-based exercise prototype, such as having a user-friendly interface and providing optimal levels of task difficulty [85]. Nevertheless, some comments indicated that further refinements are required prior to implementation of the game-based exercise among older adults in a realistic context.

Limitations

This study has certain limitations that need to be acknowledged. This study is a preliminary study that involved a small number of participants who were all female. Further, results on the enjoyment and experiences of using the game-based exercise prototype were obtained from a single training session. Thus, the findings should be considered preliminary and interpreted with caution. Future studies with larger sample sizes, a balanced gender ratio, and data obtained from multiple training sessions are warranted. Another limitation concerns the hardware specification of the Microsoft Kinect sensor. In this study, the capture volume of the Kinect sensor was restricted to 0.5 to 4.5 m, which partly limited the design of the configurations of the games. Further studies should consider using multiple Kinect sensors to cover a greater capture volume. Moreover, our game system was designed for the individual player. Exergames systems that allow group players should be considered for promoting social interaction. Finally, this study investigated the enjoyment during game engagement using the 18-item PACES questionnaire. In future studies, the 8-item version of PACES would be an appropriate questionnaire to reduce the completion time.

Conclusions

This preliminary study demonstrated a prototype of a game-based exercise for older adults using the Microsoft Kinect sensor. The game-based exercise prototype contained combined physical and cognitive training elements with different levels of difficulty. The developed game-based exercise was well accepted by the target users, with prominent enjoyment and positive feedback. Thus, the game-based exercise appears to be a promising tool for enhancing older adults' motivation to engage in physical-cognitive exercise with the aim to reduce the risk of falls.

Acknowledgments

This research was supported by the Research and Researchers for Industries (RRI) Project, Thailand Science Research and Innovation (TSRI) Grant MSD61I0015 and TSRI Grant RSA6180023 (SS). Moreover, this research was partially supported by Chiang Mai University and with the collaboration of a research group of Modern Management and Information Technology, College of Arts, Media and Technology, and the Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Thailand. In addition, the authors would like to thank K Intanon, P Sirinual, T Kunthadech, and S Korpraphan for their contributions to the development of the games.

Authors' Contributions

TK, SB, and SS share first authorships, conducted a major part of the methods and experimental design, developed the software, and contributed to the majority of the writing and reviewing of the manuscript. KP conducted a major part of the methods and experimental design, and commented on the manuscript and reviewed the final manuscript.

Conflicts of Interest

None declared.

References

1. Osoba MY, Rao AK, Agrawal SK, Lalwani AK. Balance and gait in the elderly: a contemporary review. Laryngoscope Investig Otolaryngol 2019 Feb 04;4(1):143-153. [doi: 10.1002/lio2.252] [Medline: 30828632]
2. Coutinho EDSF, Silva SDD. Uso de medicamentos como fator de risco para fratura grave decorrente de queda em idosos. Cad Saúde Pública 2002 Oct;18(5):1359-1366. [doi: 10.1590/s0102-311x2002000500029]
3. Schwartz AV, Nevitt MC, Brown Jr BW, Kelsey JL. Increased falling as a risk factor for fracture among older women: the study of osteoporotic fractures. Am J Epidemiol 2005 Jan 15;161(2):180-185. [doi: 10.1093/aje/kwi023] [Medline: 15632268]

XSL•FO
RenderX

4. Stevens JA, Corso PS, Finkelstein EA, Miller TR. The costs of fatal and non-fatal falls among older adults. Inj Prev 2006 Oct 01;12(5):290-295. [doi: 10.1136/ip.2005.011015] [Medline: 17018668]
5. Carlson JE, Ostir GV, Black SA, Markides KS, Rudkin L, Goodwin JS. Disability in older adults. 2: Physical activity as prevention. Behav Med 2017 Jun 21;24(4):157-168. [doi: 10.1080/08964289.1999.11879272] [Medline: 10023494]
6. Sherrington C, Whitney J, Lord S, Herbert R, Cumming R, Close J. Effective exercise for the prevention of falls: a systematic review and meta-analysis. J Am Geriatr Soc 2008 Dec 02;56(12):2234-2243. [doi: 10.1111/j.1532-5415.2008.02014.x] [Medline: 19093923]
7. Sherrington C, Michaleff ZA, Fairhall N, Paul SS, Tiedemann A, Whitney J, et al. Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. Br J Sports Med 2016 Oct 04;51(24):1750-1758. [doi: 10.1136/bjsports-2016-096547] [Medline: 27707740]
8. Ejupi A, Gschwind YJ, Valenzuela T, Lord SR, Delbaere K. A Kinect and inertial sensor-based system for the self-assessment of fall risk: a home-based study in older people. Hum Comput Interact 2015 Aug 25;31(3-4):261-293. [doi: 10.1080/07370024.2015.1085309]
9. Li J, Erdt M, Chen L, Cao Y, Lee S, Theng Y. The social effects of exergames on older adults: systematic review and metric analysis. J Med Internet Res 2018 Jun 28;20(6):e10486. [doi: 10.2196/10486] [Medline: 29954727]
10. Holtzer R, Friedman R, Lipton RB, Katz M, Xue X, Verghese J. The relationship between specific cognitive functions and falls in aging. Neuropsychology 2007;21(5):540-548. [doi: 10.1037/0894-4105.21.5.540] [Medline: 17784802]
11. Liu-Ambrose T, Ahamed Y, Graf P, Feldman F, Robinovitch SN. Older fallers with poor working memory overestimate their postural limits. Arch Phys Med Rehabil 2008 Jul;89(7):1335-1340. [doi: 10.1016/j.apmr.2007.11.052] [Medline: 18586136]
12. Segev-Jacubovski O, Herman T, Yogev-Seligmann G, Mirelman A, Giladi N, Hausdorff JM. The interplay between gait, falls and cognition: can cognitive therapy reduce fall risk? Expert Rev Neurother 2011 Jul 09;11(7):1057-1075 [FREE Full text] [doi: 10.1586/ern.11.69] [Medline: 21721921]
13. Martin K, Thomson R, Blizzard L, Wood A, Garry M, Srikanth V. Visuospatial ability and memory are associated with falls risk in older people. Dement Geriatr Cogn Disord 2009;27(5):451-457. [doi: 10.1159/000216840] [Medline: 19420939]
14. Valenzuela M, Sachdev P. Harnessing brain and cognitive reserve for the prevention of dementia. Indian J Psychiatry 2009 Jan;51 Suppl 1:S16-S21 [FREE Full text] [Medline: 21416010]
15. ten Brinke LF, Best JR, Crockett RA, Liu-Ambrose T. The effects of an 8-week computerized cognitive training program in older adults: a study protocol for a randomized controlled trial. BMC Geriatr 2018 Jan 30;18(1):1-11. [doi: 10.1186/s12877-018-0730-6] [Medline: 29378515]
16. Smith-Ray RL, Hughes SL, Prohaska TR, Little DM, Jurivich DA, Hedeker D. Impact of cognitive training on balance and gait in older adults. J Gerontol B Psychol Sci Soc Sci 2013 Nov 05;70(3):357-366. [doi: 10.1093/geronb/gbt097] [Medline: 24192586]
17. Smith-Ray RL, Irmiter C, Boulter K. Cognitive training among cognitively impaired older adults: a feasibility study assessing the potential improvement in balance. Front Public Health 2016 Oct 17;4:219:1-7. [doi: 10.3389/fpubh.2016.00219] [Medline: 27800473]
18. Amboni M, Barone P, Hausdorff JM. Cognitive contributions to gait and falls: evidence and implications. Mov Disord 2013 Sep 16;28(11):1520-1533. [doi: 10.1002/mds.25674] [Medline: 24132840]
19. Holtzer R, Friedman R, Lipton RB, Katz M, Xue X, Verghese J. The relationship between specific cognitive functions and falls in aging. Neuropsychology 2007 Oct;21(5):540-548. [doi: 10.1037/0894-4105.21.5.540] [Medline: 17784802]
20. Quigley A, MacKay-Lyons M, Eskes G. Effects of exercise on cognitive performance in older adults: a narrative review of the evidence, possible biological mechanisms, and recommendations for exercise prescription. J Aging Res 2020 May 14;2020:1-15. [doi: 10.1155/2020/1407896] [Medline: 32509348]
21. Fabre C, Chamari K, Mucci P, Massé-Biron J, Préfaut C. Improvement of cognitive function by mental and/or individualized aerobic training in healthy elderly subjects. Int J Sports Med 2002 Aug;23(6):415-421. [doi: 10.1055/s-2002-33735] [Medline: 12215960]
22. Silsupadol P, Shumway-Cook A, Lugade V, van Donkelaar P, Chou L, Mayr U, et al. Effects of single-task versus dual-task training on balance performance in older adults: a double-blind, randomized controlled trial. Arch Phys Med Rehabil 2009 Mar;90(3):381-387. [doi: 10.1016/j.apmr.2008.09.559] [Medline: 19254600]
23. Theill N, Schumacher V, Adelsberger R, Martin M, Jäncke L. Effects of simultaneously performed cognitive and physical training in older adults. BMC Neurosci 2013 Sep 23;14:103-114 [FREE Full text] [doi: 10.1186/1471-2202-14-103] [Medline: 24053148]
24. Quigley A, MacKay-Lyons M, Eskes G. Effects of exercise on cognitive performance in older adults: a narrative review of the evidence, possible biological mechanisms, and recommendations for exercise prescription. J Aging Res 2020;2020:1407896 [FREE Full text] [doi: 10.1155/2020/1407896] [Medline: 32509348]
25. de Bruin ED, Schoene D, Pichierri G, Smith ST. Use of virtual reality technique for the training of motor control in the elderly. Some theoretical considerations. Z Gerontol Geriatr 2010 Aug;43(4):229-234. [doi: 10.1007/s00391-010-0124-7] [Medline: 20814798]

XSL•FO
RenderX

26. Oh Y, Yang S. 2010 Oct 21 Presented at: Meaningful Play 2010; October 21-23, 2010; East Lansing, MI p. 1-17 URL: https://meaningfulplay.msu.edu/proceedings2010/mp2010_paper_63.pdf
27. Lange B, Flynn S, Proffitt R, Chang C, Rizzo AS. Development of an interactive game-based rehabilitation tool for dynamic balance training. Top Stroke Rehabil 2010;17(5):345-352. [doi: 10.1310/tsr1705-345] [Medline: 21131259]
28. Hoffmann K, Sportwiss D, Hardy S, Wiemeyer J, Göbel S. Personalized adaptive control of training load in cardio-exergames—a feasibility study. Games Health J 2015 Dec;4(6):470-479. [doi: 10.1089/g4h.2014.0073] [Medline: 26484388]
29. 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. Entertain Comput 2016 May;14:55-65. [doi: 10.1016/j.entcom.2015.10.002]
30. Deutsch JE, Brettler A, Smith C, Welsh J, John R, Guarrera-Bowlby P, et al. Nintendo wii sports and wii fit game analysis, validation, and application to stroke rehabilitation. Top Stroke Rehabil 2011;18(6):701-719. [doi: 10.1310/tsr1806-701] [Medline: 22436308]
31. van Diest M, Lamoth CJ, Stegenga J, Verkerke GJ, Postema K. Exergaming for balance training of elderly: state of the art and future developments. J Neuroeng Rehabil 2013 Sep 25;10:101 [FREE Full text] [doi: 10.1186/1743-0003-10-101] [Medline: 24063521]
32. Skjæret N, Nawaz A, Morat T, Schoene D, Helbostad JL, Vereijken B. Exercise and rehabilitation delivered through exergames in older adults: an integrative review of technologies, safety and efficacy. Int J Med Inform 2016 Jan;85(1):1-16. [doi: 10.1016/j.ijmedinf.2015.10.008] [Medline: 26559887]
33. Adcock M, Sonder F, Schättin A, Gennaro F, de Bruin ED. A usability study of a multicomponent video game-based training for older adults. Eur Rev Aging Phys Act 2020 Jan 11;17:3:1-15. [doi: 10.1186/s11556-019-0233-2] [Medline: 31938075]
34. Laver K, George S, Thomas S, Deutsch JE, Crotty M. Virtual reality for stroke rehabilitation. Stroke 2012 Feb 15;43(2):20-21. [doi: 10.1161/strokeaha.111.642439]
35. Chan CLF, Ngai EKY, Leung PKH, Wong S. Effect of the adapted Virtual Reality cognitive training program among Chinese older adults with chronic schizophrenia: a pilot study. Int J Geriatr Psychiatry 2010 Jun;25(6):643-649. [doi: 10.1002/gps.2403] [Medline: 19806599]
36. Maillot P, Perrot A, Hartley A. Effects of interactive physical-activity video-game training on physical and cognitive function in older adults. Psychol Aging 2012 Sep;27(3):589-600. [doi: 10.1037/a0026268] [Medline: 22122605]
37. Bieryla KA, Dold NM. Feasibility of Wii Fit training to improve clinical measures of balance in older adults. Clin Interv Aging 2013 Jun;8:775-781. [doi: 10.2147/cia.s46164] [Medline: 23836967]
38. Fu AS, Gao KL, Tung AK, Tsang WW, Kwan MM. Effectiveness of exergaming training in reducing risk and incidence of falls in frail older adults with a history of falls. Arch Phys Med Rehabil 2015 Dec;96(12):2096-2102. [doi: 10.1016/j.apmr.2015.08.427] [Medline: 26360975]
39. Park E, Kim S, Lee C. The effects of virtual reality game exercise on balance and gait of the elderly. J Phys Ther Sci 2015 Apr;27(4):1157-1159. [doi: 10.1589/jpts.27.1157] [Medline: 25995578]
40. Fang Q, Ghanouni P, Anderson SE, Touchett H, Shirley R, Fang F, et al. Effects of exergaming on balance of healthy older adults: a systematic review and meta-analysis of randomized controlled trials. Games Health J 2020 Feb;9(1):11-23. [doi: 10.1089/g4h.2019.0016] [Medline: 31800322]
41. Mansor NS, Chow CM, Halaki M. Cognitive effects of video games in older adults and their moderators: a systematic review with meta-analysis and meta-regression. Aging Ment Health 2019 Feb 20;24(6):841-856. [doi: 10.1080/13607863.2019.1574710] [Medline: 30784296]
42. Klompstra LV, Jaarsma T, Strömberg A. Exergaming in older adults: a scoping review and implementation potential for patients with heart failure. Eur J Cardiovasc Nurs 2014 Oct 01;13(5):388-398. [doi: 10.1177/1474515113512203] [Medline: 24198306]
43. Yeung LF, Yang Z, Cheng KC, Du D, Tong RK. Effects of camera viewing angles on tracking kinematic gait patterns using Azure Kinect, Kinect v2 and Orbbec Astra Pro v2. Gait Posture 2021 Jun;87:19-26. [doi: 10.1016/j.gaitpost.2021.04.005] [Medline: 33878509]
44. Sápi M, Fehér-Kiss A, Csernák K, Domján A, Pintér S. The effects of exergaming on sensory reweighting and mediolateral stability of women aged over 60: usability study. JMIR Serious Games 2021 Jul 21;9(3):e27884. [doi: 10.2196/preprints.27884] [Medline: 34287215]
45. Zhang Z. Microsoft Kinect Sensor and Its Effect. IEEE Multimedia 2012 Feb 27;19(2):4-10. [doi: 10.1109/mmul.2012.24]
46. Lange B, Chang C, Suma E, Newman B, Rizzo AS, Bolas M. Development and evaluation of low cost game-based balance rehabilitation tool using the Microsoft Kinect sensor. Annu Int Conf IEEE Eng Med Biol Soc 2011;2011:1831-1834. [doi: 10.1109/IEMBS.2011.6090521] [Medline: 22254685]
47. Yang CM, Hsieh JS, Chen YC, Yang SY, Lin HC. Effects of Kinect exergames on balance training among community older adults: a randomized controlled trial. Medicine 2020 Jul 10;99(28):e21228. [doi: 10.1097/md.0000000000021228] [Medline: 32664177]

XSL•FO
RenderX

48. Yu TC, Chiang CH, Wu PT, Wu WL, Chu IH. Effects of exergames on physical fitness in middle-aged and older adults in Taiwan. Int J Environ Res Public Health 2020 Apr 08;17(7):2565. [doi: 10.3390/ijerph17072565] [Medline: 32276515]
49. Adcock M, Fankhauser M, Post J, Lutz K, Zizlsperger L, Luft AR, et al. Effects of an in-home multicomponent exergame training on physical functions, cognition, and brain volume of older adults: a randomized controlled trial. Front Med 2020 Jan 28;6:321. [doi: 10.3389/fmed.2019.00321] [Medline: 32047751]
50. Liao YY, Chen IH, Wang RY. Effects of Kinect-based exergaming on frailty status and physical performance in prefrail and frail elderly: a randomized controlled trial. Sci Rep 2019 Jun 27;9(1):9353. [doi: 10.1038/s41598-019-45767-y] [Medline: 31249332]
51. Zhao Y, Feng H, Wu X, Du Y, Yang X, Hu M, et al. Effectiveness of exergaming in improving cognitive and physical function in people with mild cognitive impairment or dementia: systematic review. JMIR Serious Games 2020 Jun 30;8(2):e16841. [doi: 10.2196/16841] [Medline: 32602841]
52. Hanson VL. Influencing technology adoption by older adults. Interact Comput 2010 Nov;22(6):502-509. [doi: 10.1016/j.intcom.2010.09.001]
53. Hemingway C, Baja ES, Dalmacion GV, Medina PMB, Guevara EG, Sy TR, et al. Development of a mobile game to influence behavior determinants of HIV service uptake among key populations in the Philippines: user-centered design process. JMIR Serious Games 2019 Dec 20;7(4):e13695. [doi: 10.2196/13695] [Medline: 31859673]
54. Pagulayan R, Keeker K, Wixon D, Romero R, Fuller T. User-centered design in games. In: Jacko J, Sears A, editors. Handbook for Human-Computer Interaction in Interactive Systems. Mahwah, NJ: Lawrence Erlbaum Associates Inc; 2003:883-906.
55. Fullerton T, Swain C, Hoffman S. Game design workshop. In: Designing, Prototyping and Playtesting Games. San Francisco, CA: CMP Books; Jan 2004.
56. Salen K, Tekinba K, Zimmerman E. Rules of play. In: Game Design Fundamentals. Cambridge, MA: MIT Press; 2004.
57. Preece J, Rogers Y, Sharp H, Benyon D, Holland S, Carey T. Human-Computer Interaction. Edinburgh Gate Harlow, UK: Addison-Wesley Longman Ltd; Oct 01, 1994:82-85.
58. Carroll JM. Encountering others: reciprocal openings in participatory design and user-centered design. Hum Comput Interact 2009 Nov 11;11(3):285-290. [doi: 10.1207/s15327051hci1103_5]
59. Schättin A, Häfliger S, Meyer A, Früh B, Böckler S, Hungerbühler Y, et al. Design and evaluation of user-centered exergames for patients with multiple sclerosis: multilevel usability and feasibility studies. JMIR Serious Games 2021 May 7;9(2):e22826. [doi: 10.2196/22826] [Medline: 33960956]
60. Chu CH, Biss RK, Cooper L, Quan AML, Matulis H. Exergaming platform for older adults residing in long-term care homes: user-centered design, development, and usability study. JMIR Serious Games 2021 Mar 9;9(1):e22370. [doi: 10.2196/22370] [Medline: 33687337]
61. Harte R, Glynn L, Rodríguez-Molinero A, Baker PM, Scharf T, Quinlan LR, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. JMIR Hum Factors 2017 Mar 16;4(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.5443] [Medline: 28302594]
62. Ergonomics of human-system interaction. Human-centred design for interactive systems. British Standards Institution. 2010. URL: https://doi.org/10.3403/30186637 [accessed 2021-08-20]
63. Lord S, Tiedemann A, Chapman K, Munro B, Murray S, Gerontology M, et al. The effect of an individualized fall prevention program on fall risk and falls in older people: a randomized, controlled trial. J Am Geriatr Soc 2005 Aug;53(8):1296-1304. [doi: 10.1111/j.1532-5415.2005.53425.x] [Medline: 16078954]
64. Choi SD, Guo L, Kang D, Xiong S. Exergame technology and interactive interventions for elderly fall prevention: a systematic literature review. Appl Ergon 2017 Nov;65:570-581. [doi: 10.1016/j.apergo.2016.10.013] [Medline: 27825723]
65. Lipardo DS, Aseron AMC, Kwan MM, Tsang WW. Effect of exercise and cognitive training on falls and fall-related factors in older adults with mild cognitive impairment: a systematic review. Arch Phys Med Rehabil 2017 Oct;98(10):2079-2096. [doi: 10.1016/j.apmr.2017.04.021] [Medline: 28554873]
66. Montero-Odasso M, Oteng-Amoako A, Speechley M, Gopaul K, Beauchet O, Annweiler C, et al. The motor signature of mild cognitive impairment: results from the gait and brain study. J Gerontol A Biol Sci Med Sci 2014 Sep 02;69(11):1415-1421. [doi: 10.1093/gerona/glu155] [Medline: 25182601]
67. Zhang W, Low LF, Schwenk M, Mills N, Gwynn JD, Clemson L. Review of gait, cognition, and fall risks with implications for fall prevention in older adults with dementia. Dement Geriatr Cogn Disord 2019 Nov 19;48(1-2):17-29. [doi: 10.1159/000504340] [Medline: 31743907]
68. van Someren MW, Barnard YF, Sandberg JAC. The think aloud method: a practical approach to modelling cognitive processes. London, United Kingdom: Academic Press; 1994. URL: http://echo.iat.sfu.ca/library/vanSomeren_94_think_aloud_method.pdf [accessed 2021-08-18]
69. Kendzierski D, DeCarlo K. Physical activity enjoyment scale: two validation studies. J Sport Exerc Psychol 1991 Mar;13(1):50-64. [doi: 10.1123/jsep.13.1.50]
70. Mullen SP, Olson EA, Phillips SM, Szabo AN, Wójcicki TR, Mailey EL, et al. Measuring enjoyment of physical activity in older adults: invariance of the physical activity enjoyment scale (paces) across groups and time. Int J Behav Nutr Phys Act 2011 Sep 27;27(8:103):1-9. [doi: 10.1186/1479-5868-8-103] [Medline: 21951520]

XSL•FO
RenderX

71. Boongird P. Mental state examination T10. Webpage in Thai. Dementia Association of Thailand. 2018. URL: http://www.wongpakaran.com/images/sub_1550692605/MSET10_GPsy_8Mar2021.pdf [accessed 2021-08-11]
72. Wongpakaran N, Wongpakaran T, Van Reekum R. The use of GDS-15 in detecting MDD: a comparison between residents in a Thai long-term care home and geriatric outpatients. J Clin Med Res 2013 Apr;5(2):101-111. [doi: 10.4021/jocmr1239w] [Medline: 23518497]
73. Kinect for Windows Human Interface Guidelines v2.0. Microsoft Corporation. 2014. URL: http://download.microsoft.com/download/6/7/6/676611b4-1982-47a4-a42e-4cf84e1095a8/kinecthig.2.0.pdf [accessed 2021-08-18]
74. Gonzalez-Jorge H, Rodríguez-Gonzálvez P, Martínez-Sánchez J, González-Aguilera D, Arias P, Gesto M, et al. Metrological comparison between Kinect I and Kinect II sensors. Measurement 2015 Jun;70:21-26. [doi: 10.1016/j.measurement.2015.03.042]
75. American College of Sports Medicine. ACSM's guidelines for exercise testing and prescription. In: Riebe D, Ehrman JK, Liguori G, Magal M, editors. American College of Sports Medicine Series. Philadelphia, PA: Wolters Kluwer Health; Dec 13, 2016.
76. Brouwer W, Kroeze W, Crutzen R, de Nooijer J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011 Jan 6;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
77. Nyman SR, Victor CR. Older people's recruitment, sustained participation, and adherence to falls prevention interventions in institutional settings: a supplement to the Cochrane systematic review. Age Ageing 2011 Jul;40(4):430-436. [doi: 10.1093/ageing/afr016] [Medline: 21502163]
78. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012 Apr 30;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
79. Howes SC, Charles D, Pedlow K, Wilson I, Holmes D, McDonough S. User-centred design of an active computer gaming system for strength and balance exercises for older adults. J Enabling Technol 2019 Apr 16;13(2):101-111. [doi: 10.1108/jet-12-2018-0057]
80. Okubo Y, Schoene D, Lord SR. Step training improves reaction time, gait and balance and reduces falls in older people: a systematic review and meta-analysis. Br J Sports Med 2017 Apr;51(7):586-593. [doi: 10.1136/bjsports-2015-095452] [Medline: 26746905]
81. Nawaz A, Skjæret N, Ystmark K, Helbostad J, Vereijken B, Svanæs D. Assessing seniors' user experience (UX) of exergames for balance training. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction. 2014 Oct 26 Presented at: 8th Nordic Conference on Human-Computer Interaction; October 26-30, 2014; Helsinki, Finland p. 578-587. [doi: 10.1145/2639189.2639235]
82. Tabak M, Weering MDV, van Dijk H, Vollenbroek-Hutten M. Promoting daily physical activity by means of mobile gaming: a review of the state of the art. Games Health J 2015 Dec;4(6):460-469. [doi: 10.1089/g4h.2015.0010] [Medline: 26397179]
83. Lewis ZH, Swartz MC, Lyons EJ. What's the point?: a review of reward systems implemented in gamification interventions. Games for Health Journal 2016 Apr;5(2):93-99. [doi: 10.1089/g4h.2015.0078] [Medline: 26812253]
84. Da Silva Júnior JLA, Biduski D, Bellei EA, Becker OHC, Daroit L, Pasqualotti A, et al. A bowling exergame to improve functional capacity in older adults: co-design, development, and testing to compare the progress of playing alone versus playing with peers. JMIR Serious Games 2021 Jan 29;9(1):e23423. [doi: 10.2196/23423] [Medline: 33512319]
85. Lu MH, Lin W, Yueh HP. Development and evaluation of a cognitive training game for older people: a design-based approach. Front Psychol 2017 Oct 17;8:1837. [doi: 10.3389/fpsyg.2017.01837] [Medline: 29089914]

Abbreviations

PACES: Physical Activity Enjoyment Scale
UCD: user-centered design

Edited by N Zary; submitted 10.02.21; peer-reviewed by M Jordan-Marsh, Y Zhang, DR Colombo Dias, R Dörner, C Sik-Lanyi; comments to author 17.04.21; revised version received 08.09.21; accepted 24.09.21; published 27.10.21.

Please cite as:
Kamnardsiri T, Phirom K, Boripuntakul S, Sungkarat S
An Interactive Physical-Cognitive Game-Based Training System Using Kinect for Older Adults: Development and Usability Study
JMIR Serious Games 2021;9(4):e27848
URL: https://games.jmir.org/2021/4/e27848
doi:10.2196/27848
PMID:34704953

XSL•FO
RenderX

©Teerawat Kamnardsiri, Kochaphan Phirom, Sirinun Boripuntakul, Somporn Sungkarat. Originally published in JMIR Serious Games (https://games.jmir.org), 27.10.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Immersion Experiences in a Tablet-Based Markerless Augmented Reality Working Memory Game: Randomized Controlled Trial and User Experience Study

Bo Zhang[1], BSc, MSc, PhD; Nigel Robb[2], PhD

[1]Department of Education Information Technology, Faculty of Education, East China Normal University, Shanghai, China
[2]Research Faculty of Media and Communication, Hokkaido University, Sapporo, Japan

Corresponding Author:
Bo Zhang, BSc, MSc, PhD
Department of Education Information Technology
Faculty of Education
East China Normal University
No 3663, North Zhongshan Road
Putuo District
Shanghai, 200062
China
Phone: 86 13505194318
Email: bzhang@ed.ecnu.edu.cn

Abstract

Background: In recent years, augmented reality (AR), especially markerless augmented reality (MAR), has been used more prevalently to create training games in an attempt to improve humans' cognitive functions. This has been driven by studies claiming that MAR provides users with more immersive experiences that are situated in the real world. Currently, no studies have scientifically investigated the immersion experience of users in a MAR cognitive training game. Moreover, there is an observed lack of instruments on measuring immersion in MAR cognitive training games.

Objective: This study, using two existing immersion questionnaires, investigates students' immersion experiences in a novel MAR n-back game.

Methods: The n-back task is a continuous performance task that taps working memory (WM) capacity. We compared two versions of n-back training. One was presented in a traditional 2D format, while the second version used MAR. There were 2 experiments conducted in this study that coordinated with 2 types of immersion questionnaires: the modified Immersive Experiences Questionnaire (IEQ) and the Augmented Reality Immersion (ARI) questionnaire. Two groups of students from two universities in China joined the study, with 60 participants for the first experiment (a randomized controlled experiment) and 51 participants for the second.

Results: Both groups of students experienced immersion in the MAR n-back game. However, the MAR n-back training group did not experience stronger immersion than the traditional (2D) n-back control group in the first experiment. The results of the second experiment showed that males felt deeply involved with the AR environment, which resulted in obtaining higher levels of immersion than females in the MAR n-back game.

Conclusions: Both groups of students experienced immersion in the MAR n-back game. Moreover, both the modified IEQ and ARI have the potential to be used as instruments to measure immersion in MAR game settings.

Trial Registration: UMIN Clinical Trials Registry UMIN000045314; https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000051725

(JMIR Serious Games 2021;9(4):e27036) doi:10.2196/27036

KEYWORDS

augmented reality; markerless augmented reality; immersion experience; cognitive training games; working memory; markerless augmented reality n-back game

XSL•FO
RenderX

Introduction

In the past decade, studies have reported significant improvements in cognitive functions through game-based cognitive training in different groups [1-3]. In recent years, research shows that the application of augmented reality (AR) technology in games is related to the user's cognitive functionality, as it has a positive impact on mental processes and psychological reactions [4,5]. Many studies also started to explore the affordances of using AR technology to create cognitive training tasks. There is some evidence that AR-based cognitive training is effective in improving users' cognitive functions, such as attention and memory [6-8]. Other research shows that AR-based tasks play a significant role in increasing students' motivation and attention during the learning process [9].

By overlapping computer-generated 3D graphics onto a real-world environment, AR enables borderless interactions between the digital and physical worlds [10]. AR has the following key features: combination of virtual content and real context, real-time interaction between virtual objects and real contexts, and the registered 3D virtual object [11]. Through mobile devices (eg, smartphones and tablets), AR apps capture the real scene with the camera and then present the digital contents to the real context surrounding the user. This means of interaction creates a new mode of learning and training, which can be easily conducted even by students who have no experience using digital devices [12]. AR apps enable users to have ubiquitous access to surrounding real contextual environments while playing AR games on a mobile interface [13].

According to how AR apps represent virtual graphics in a real context through device camera screens, AR games are divided into 2 types: markerless AR (MAR) games and marker-based AR games [10]. Generally, marker-based AR requires a fiducial or artificial marker placed in a real context to implement AR experiences. These markers, such as printed barcodes and QR codes on cards, generally have specific geometric or color properties, which makes them easy to extract and identify in a video frame. Marker-based AR uses an external marker to calibrate the camera's pose and then achieves a successful projection of virtual content into the captured real environment [14]. However, this approach has been shown to have many drawbacks, such as virtual content being easily lost in the tracking process if a live camera moves quickly and visual markers have to remain in sight due to a limited range [13,15]. Moreover, AR with fiducial markers is not scalable for outdoor scenes and mobile learning [16].

By contrast, the tracking system (based on a simultaneous localization and mapping technique) of MAR can rely on natural features instead of using external markers to trigger augmentations in the real environment [17]. Through an eligible mobile device, markerless tracking can manipulate any part of a real location as a marker to place virtual objects [18,19]. MAR games are similar to location-aware AR games [20], which are geo-based augmented realities that do not need any special markers for identifying where to place a virtual object in the real environment. In this way, these games can provide more realistic and interactive AR experiences for users, accurately track the real context, and detect locations through smart devices [21]. A MAR app brings a markedly different experience for its users, which enhances interactive involvement with and perception of the real surroundings without physical constrains [5,18,22]. Thus, MAR apps have been argued to provide users with more immersive and interactive experiences, which can be beneficial for promoting enjoyment and engagement in learning and training [21]. As AR technology can enable users to experience a sense of immersion [23,24], MAR game playing may provide users with a stronger sense of immersion experience.

Immersion has been widely discussed in the context of AR games [10,20], especially ones related to cognitive training [6,25]. Immersion, a form of cognitive and emotional absorption, has been used to promote enjoyment and engagement in tasks and learning [26,27]. Immersion is widely considered as a desired outcome of the gaming experience [27]. Performing and training related to immersive digital experiences are assumed to be reliant on the degree of achieved immersion, specifically the degree to which users become cognitively and emotionally engaged with a given digital app [28,29]. These statements indicate that measuring the immersion in the context of an AR game developed for cognitive training is a primary tool to evaluate the effects of AR games. Hence, to discover the potential of MAR games in the field of cognitive training, we need to investigate users' immersion experiences in MAR games playing. However, there is an observed lack of studies investigating immersion in MAR cognitive training games.

Immersion in the field of digital games has been widely discussed in recent years [20,27,29-31]. The definition of immersion was first proposed in the field of virtual environments. The research by Georgiou and Kyza [20] argues that immersion can be seen as a reliable dimension to objectively access the properties of a virtual environment. On the other hand, Witmer and Singer [32] described the concept of immersion as a "psychological state characterized by perceiving oneself to be enveloped by, included in and interacting with an environment that provides a continuous stream of stimuli and experiences." In addition, Brooks [28] argued that an immersive experience can not only occur in virtual reality (VR) technologies but also appear when playing simple computer games, such as video games. The definition of immersion also considers the di erent degrees of cognitive and a ective absorption when playing a digital game, as they are closely related to task enjoyment and engagement [20,33]. Moreover, many studies suggest that gender affects the immersion level in digital games [34]. Specifically, men and women report different levels of immersion when interacting with virtual environments [35] and when playing video games [36].

However, the virtual environments described in the definitions were generally created by VR technology. By wearing head-mounted displays, users were immersed in virtual environments, isolating completely from real contexts. While AR games are played using mobile devices, the games can also provide users with immersive experiences by interacting with a continuous stream of digital-based stimuli situated in the real

XSL•FO
RenderX

contexts. Hence, users may perceive a different immersive experience when playing AR apps, compared with virtual or other digital environments.

Previously, in the absence of valid immersion measurements, some researchers discussed AR immersive experiments through the evaluation of flow and presence [37,38]. Csikszentmihalyi [39] defined flow as "the state in which individuals are so involved in an activity that nothing else seems to matter." As for the immersion state, flow can be described as people so absorbed in their activities that irrelevant thoughts and perceptions are filtered out from their minds [27]. The definition of presence is a psychological sense of being in a virtual environment [40]. The degree to which the virtual environment mimics real-world experiences is impacted by the degree of participants' presence in the virtual environment [32].

Some studies argued that the sense of flow and presence could barely be created and maintained in AR games [38], as potential external distractions exist in the context of an AR system. For instance, many factors such as temperature, light and noise can be easily controlled in the virtual environment, while these parameters are hardly controlled by designers of AR apps in a real context [41]. These potential and uncontrolled external distractions may prevent users from giving their attention and thus disrupt their immersive experiences in AR games [38,41].

In addition, Benyon [42] stated that immersion in AR experiences is different from feelings of presence, in terms of being in a virtual space. Yet, immersion in AR experiences is the feeling of involvement in a blended space of real and digital elements. Moreover, a good AR game should be able to immerse its users into playing the game and decrease their focus on external distractions situated in real contexts. Hence, the evaluation of immersion provides a valuable option to describe the user experiences in the context of AR (or MAR) games.

There are several previously developed validated instruments for evaluating immersion in digital games [20,27,29,30]. Although these existing instruments are validated in some games, they may be invalidated in other types of games (eg, AR games) when measuring the immersion experience. As described above, we can see MAR games are technically and functionally different from other digital software (eg, 2D games), as they can present 3D elements and software assets in real environments without external markers.

It is uncertain which existing valid immersion questionnaire is valid for measuring user immersion in emerging MAR apps. As a potential benefit of MAR technology in developing cognitive training games, identifying valid instruments for evaluating immersion in MAR games is significant for the researchers and developers in this field. Therefore, this study aims to explore students' immersion experiences in a novel MAR cognitive training task by using 2 existing immersion instruments for digital games.

Methods

Development of a MAR n-Back Task

An n-back task is used extensively in the literature as a working memory (WM) training task [43,44]. Participants in an n-back task are presented with a series of stimuli. They are required to retain some aspects of each stimulus in WM (eg, the location of the stimulus in a grid). During the n-back task, the trainee is instructed to respond whenever a current stimulus is presented on-screen that matches the one presented n positions back in the sequence [43]. In this study, a MAR n-back game and a 2D n-back game were developed (Figure 1) based on the same principles of the n-back task described in previous studies [45-47]. The MAR n-back game was developed using a combination of the Unity game engine and Apple's ARKit database, based on the MAR game framework developed by Chen et al [14] that could achieve real-time 3D context reconstruction by using mobile devices. The main differences between the 2 versions of the game were how the player perceived the location of the stimuli: whether these stimuli were more stimulating graphics and whether these stimuli were perceived as being located on the screen of the device (non-MAR) or as being in real-world surroundings (MAR).

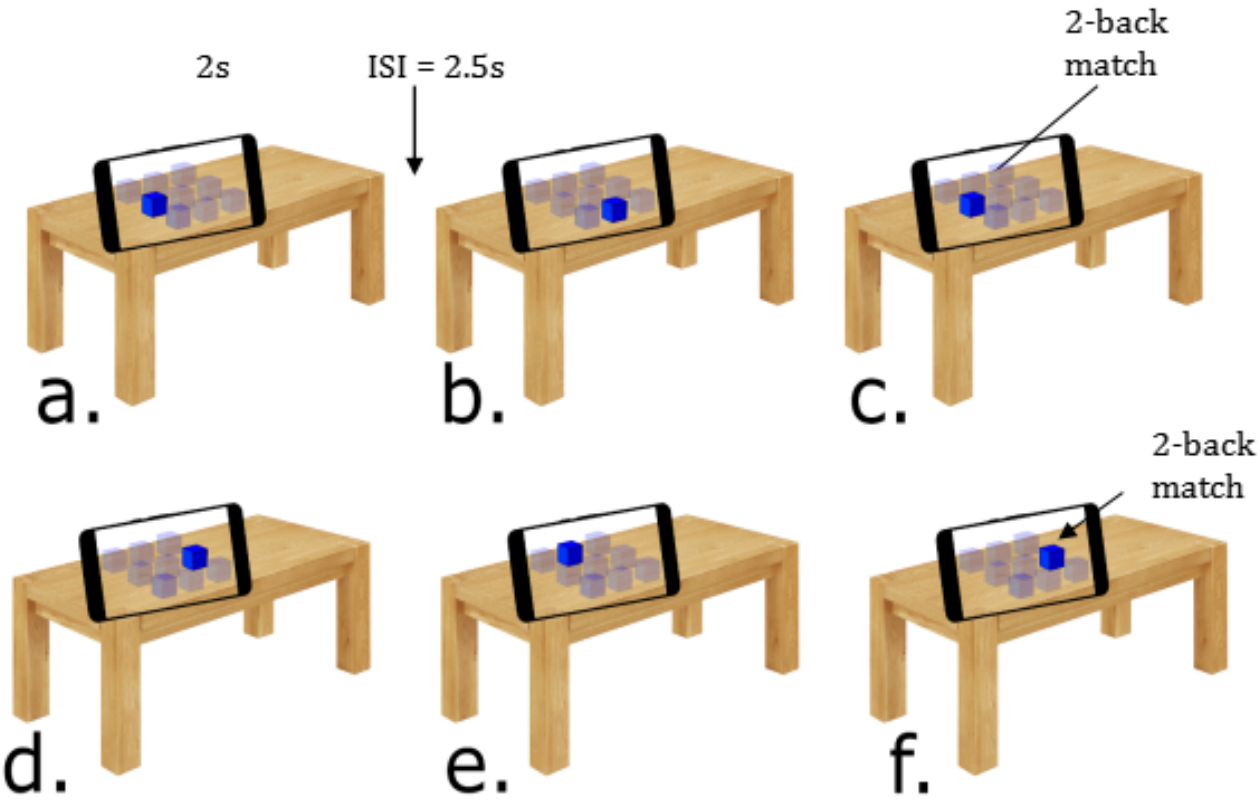
Figure 1. Examples of (a) markerless augmented reality n-back game and (b) traditional (2D) n-back game.

Figure 2 shows 6 successive trials from the MAR n-back game. Each trial consisted of a stimulus that was presented for 2 seconds (s), followed by an interstimulus interval of 2.5 s and the next stimulus [43]. In this example, with n (n=1, 2, or 3) set

XSL•FO
RenderX

to 2 (2-back), the correct responses would be to indicate a match on trial (c) because the location matches the trial's previous 2 steps (a), and to indicate a match on trial (f) because the location matches the trial's previous 2 steps (d).

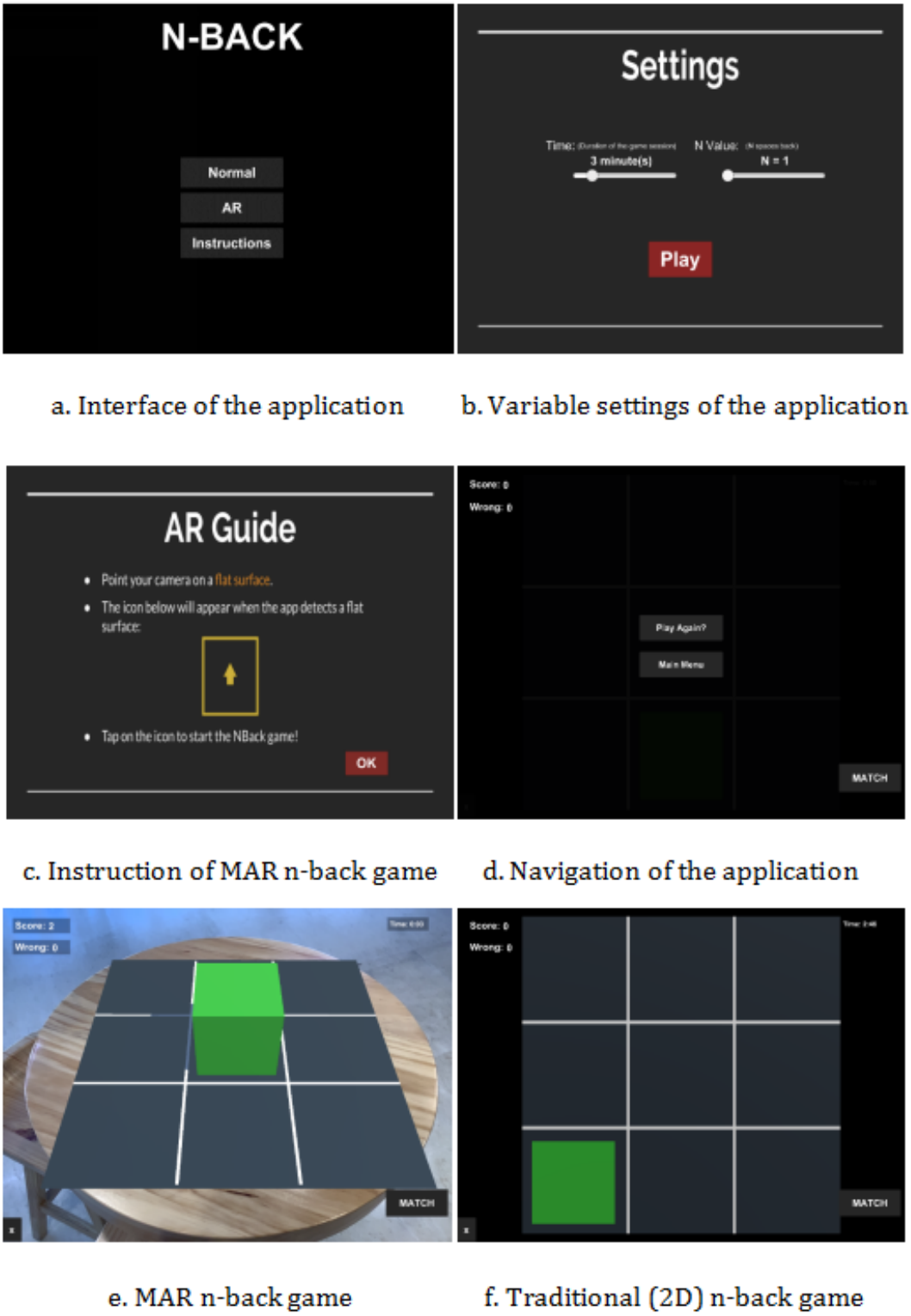
Figure 2. Six successive trials from the augmented reality version of the n-back game.



Hence, by playing the MAR n-back game, participants can interact with stimulus within the real context via smart devices. We assume that MAR n-back games could provide users with more immersive and interactive experiences, thereby being potentially beneficial for promoting enjoyment and engagement in cognitive training. As AR technology has been increasingly applied to various domains, it is urgent to investigate the essential theories and practices related to the key aspects of AR experiences, such as immersive features. In this study, both versions of the n-back games were created with 3 load levels (1-back, 2-back, and 3-back). Figure 3 illustrates the app settings of both versions of the n-back game, while Figure 4 displays the flowchart of a full gaming session (spanning 3 minutes). Hence, the motivation of n-back game training is based on the participants responding to the correct targets as quickly and accurately as possible, while the difficulty increases by raising the value of n. Through this way, participants were motivated to take the n-back game training seriously and perform effectively, as their scores were cumulatively displayed against other players.

XSL•FO
RenderX

Figure 3. App settings of both versions of the n-back game. MAR: markerless augmented reality.



a. Interface of the application

b. Variable settings of the application

c. Instruction of MAR n-back game

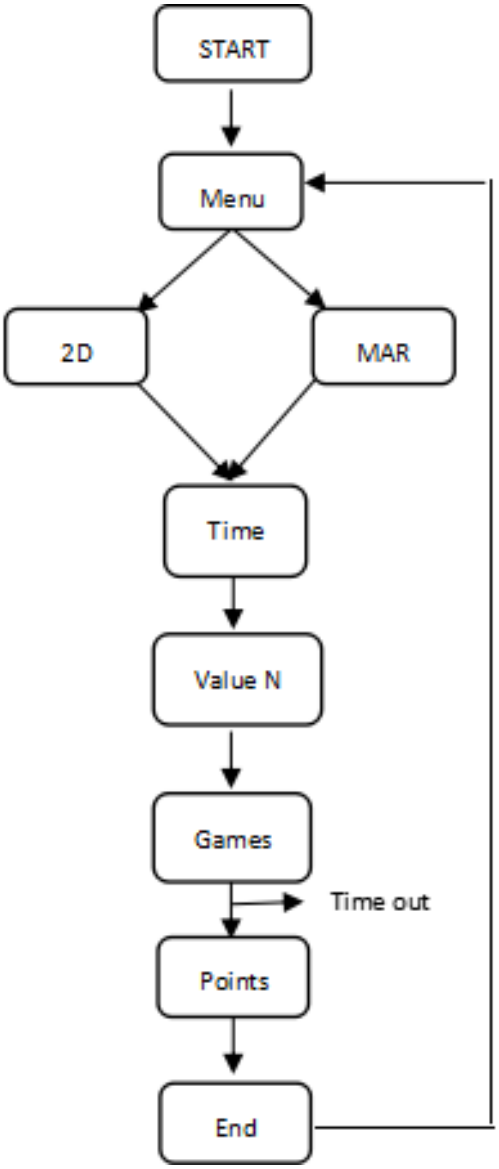
d. Navigation of the application

e. MAR n-back game

f. Traditional (2D) n-back game

XSL•FO
RenderX

Figure 4. Flowchart of a full gaming session with both versions of the n-back game. MAR: markerless augmented reality.



Immersion Measurements

Background

In this study, two existing validated immersion instruments: the Immersive Experiences Questionnaire (IEQ) and the Augmented Reality Immersion (ARI) questionnaire have been selected as the measurements with some reasons.

In 2004, Brown and Cairn [30] created a famous model of immersion with the 3 sequential levels of engagement, engrossment, and total immersion to describe the degrees of involvement with digital games. The level of engagement includes access and investment, which means that the player needs to invest time and effort in learning how to play the game and get familiar with the game's controls. From the engagement level, participants may be able to become more interested in and further involved with the game and then move into the engrossment level. During this level, gamers' attention spans and emotions were directly affected by the game and game controls became invisible because gamers became less aware of their surroundings [27]. Total immersion occurred when gamers reached a sense of presence, became lost in the game's world, and achieved a sense of feeling that the game was all that mattered.

On the basis of immersion model [30], flow [39], cognitive absorption [33] and presence [32], the IEQ was developed for measuring immersion in digital games [27]. In particular, the validity of the IEQ was proven through experiments with 2 different types of games: a video (2D) and a VR game. As virtual environments, although AR differs from VR in displaying digital content, both technologies are claimed to represent effective immersive technology [48]. Moreover, the focus of the IEQ covers 5 factors: cognitive involvement, emotional involvement, real-world dissociation, challenge, and control [27]. These factors are more related to the purpose of measuring immersion in a cognitive training game (eg, n-back game). Hence, using the IEQ, this study sets a similar experiment to explore differences in students' immersion experiences with 2 different types of games: a video (2D) game and a MAR game.

The ARI questionnaire, based on the immersion theory of Brown and Cairns [30], was developed particularly for the purpose of

XSL•FO
RenderX

measuring immersion in location-aware AR settings, which, as explained earlier, is similar to MAR games. It has also been used for measuring immersion in other types of AR apps used for training [31]. Georgiou and Kyza [20] emphasize that immersion in AR apps is not about getting disengaged from the real world but shifting attention toward AR games, which in turn results in a decreased focus on any potential distractors. Hence, to investigate students' immersion experiences in a MAR game, the ARI questionnaire is necessary for this study.

Immersive Experiences Questionnaire

The IEQ includes 33 items; 32 questions (Q1-Q32) were all designed on 5-point scales (1 = strongly disagree and 5 = strongly agree). One question used a 10-point scale, asking participants to indicate how immersed the felt overall (1 = not at all and 10 = very much).

However, as discussed in the previous section, users perceive a feeling of being surrounded by a blended world of real and digital elements in AR games, unlike virtual environments in which users can totally immerse themselves [36].

Therefore, to make the IEQ suitable for assessing immersion in AR game settings, some questions were omitted:

- Q19: It was as if I could interact with the world of the game as if I was in the real world.
- Q20: Interacting with the world of the game did not feel as real to me as it would be in the real world.
- Q23: I felt detached from the outside world.
- Q30: I still felt as if I was in the real world while playing.

Moreover, certain potential external distractions, such as weather and noises, exist in the context of an AR system. Therefore, the Q21 "I was unaware of what was happening around me" should also be removed from the IEQ [20]. Hence, a modified IEQ, with 27 5-point scale questions (omitting questions Q19, Q20, Q21, Q23, Q30) and one 10-point scale question, was used to measure participant immersion experiences in this study.

ARI Questionnaire

The ARI questionnaire is a 21-item, 7-point Likert-type instrument ranging from totally disagree (1) to totally agree (7) [20]. The ARI questionnaire has a 3-level construct (with each construct broken down further into 2 subcategories) according to the following: engagement (interest and usability), engrossment (emotional investment and focus of attention), and total immersion (presence and flow). These factors measure the immersion level of students while using AR apps.

First Experiment

Participants

Recruitment of participants began in October 2019. A total of 60 participants took part in the first experiment (average age 18.97 [SD 1.09] years; 42 females). They were a group of undergraduate students from a university in China. No eligibility criteria were specified. Written informed consent forms were collected from these students before they began the WM training. Participants were informed that they could drop out of the study at any time. Prior to beginning the study, ethics approval was obtained from the ethics committee at East China Normal University (HR 055-2019). The trial was registered at the University Hospital Medical Information Network Clinical Trials Registry [UMIN000045314]. All participants were paid US $15.

Procedure

At the beginning of the first experiment, participants were randomly divided into 2 groups, with 30 participants in the control group (who trained on the video n-back game) and 30 in the experimental group (who trained on the MAR n-back game). Participants were not blinded. Random allocation was based on random numbers generated in Microsoft Excel by BZ. During the training, six 10.5-inch iPad Pro tablets (Apple Inc) were used to present the n-back game (see Figure 5). The instructions for the n-back task were demonstrated to the participants prior to the WM training. Two groups of participants then received separate WM training in 2 versions of n-back tasks for 8 sessions (2 sessions a day, with each session lasting 3 minutes) within 4 days. All training sessions were conducted in a laboratory. At this stage, the aim of the training was not to investigate its effects on WM but to focus on comparing the level of immersion between the 2 groups. After all the training sessions were completed, each participant completed a modified IEQ through a link to an online questionnaire using their mobile phones.

Figure 5. Visual depictions of the first experiment.

XSL•FO
RenderX

Second Experiment

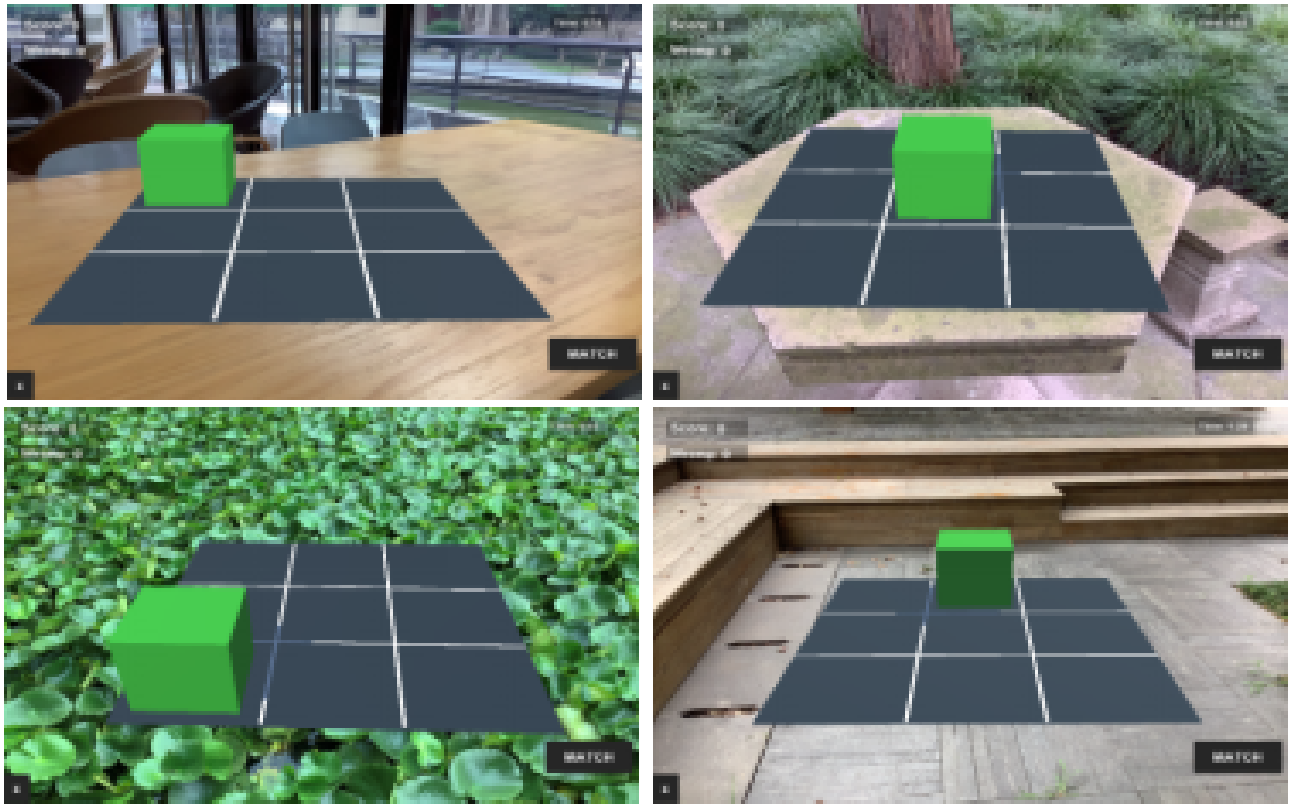
Participants

A total of 51 participants took part in the second experiment (average age 21.22 [SD 1.69] years; 30 females). Participants, studying in the major of Educational Technology, were recruited from another university in China. To complete the assignments on the evaluation of an innovative software, these students need to test an innovative app. All participants were fully debriefed at the end of the experiment. Written informed consent forms were also collected from these students before they completed the survey.

Procedures and Materials

In this experiment, 51 participants practiced in a MAR n-back game for once a day, 3 minutes each time, using 6 iPad tablets for 1 week. Participants were allowed to take the tablets with them outside of the classes and could play the MAR n-back game at any location within the university's campus. The instructions for the WM training with MAR n-back game were explained to the participants and several practice trials were conducted before the initiation of the study. Some visuals illustrating the practice process are presented in Figure 6.

Figure 6. Visuals demonstrating the second experiment.



After 1 week of playing, the ARI questionnaire was issued to students to complete through an online link using mobile phones, and data were collected. The practice process was given separately to each participant and the data collection took place over a 1-week period. As the ARI questionnaire is particularly developed to measure immersion in AR settings, the second experiment aims to further investigate students' immersion experiences in a MAR game based on the study of Georgiou and Kyza (2016).

Results

First Experiment

Following the example of Jennett et al [27], we first examined correlations between the 2 immersion measures of 27 5-point scale questions and one 10-point scale question in the modified IEQ. Taking both groups (control and experimental groups) together, the 2 immersion measures were positively, significantly correlated (Pearson r=0.466, P<.001). This finding appears reliable, as variables such as participant age, gender, and education level were all found to have nonsignificant correlations with both immersion measures. Taking the groups individually, the correlation was stronger in the MAR n-back group (Pearson r=0.518, P=.003) than the 2D n-back group (Pearson r=0.362, P=.049). Based on a total 28 items of the modified IEQ, while the MAR n-back group had a higher mean value on all immersion measures (see Table 1) than the 2D n-back group, the differences between the 2 groups were not significant on either measure.

XSL•FO
RenderX

Table 1. Mean and standard deviation for the measures of immersion in the modified Immersive Experiences Questionnaire.

Characteristic	5-point scale (27 items)		10-point scale (1 item)	
	MAR[a] (n=30)	2D[b] (n=30)	MAR (n=30)	2D (n=30)
Mean	3.584	3.421	7.9	7.433
Standard deviation	0.619	0.516	1.668	1.501

[a]MAR: markerless augmented reality n-back game.
[b]2D: n-back game.

Second Experiment

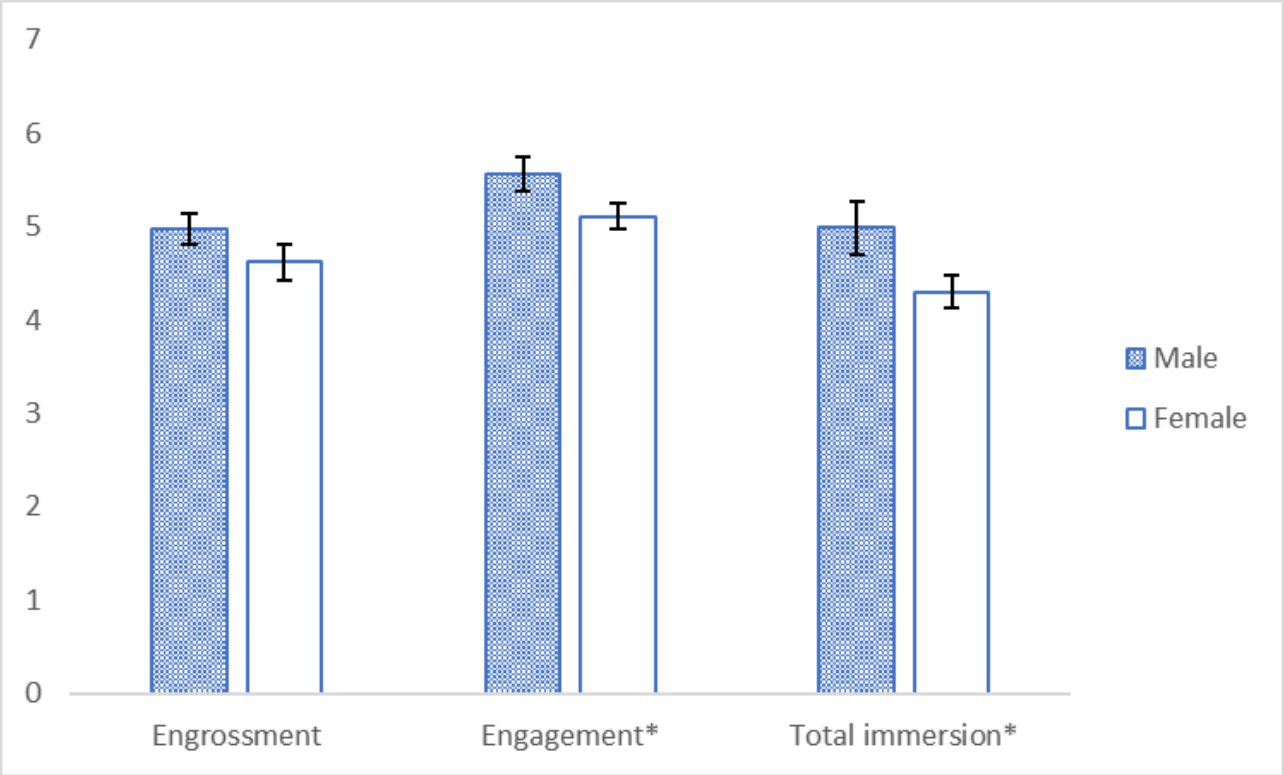
Cronbach alpha for the ARI scale was 0.783 overall and ranged from 0.711 to 0.836 for the 6 subdimensions (ie, interest, usability, emotional investment, focus of attention, presence and flow), indicating acceptable reliability. The mean values for engagement, engrossment, and total immersion with the subdimensions of interest, usability, emotional investment, focus of attention, presence, and flow (see Table 2) were similar to the research by Georgiou and Kyza [20] using the ARI to investigate immersion in AR experiences but noticeably lower than that found in the research of Salar et al [31].

Table 2. Mean values for interest, usability, emotional investment, focus of attention, presence and flow.

Factor	Value, mean (SD)
Interest	5.853 (0.156)
Usability	3.995 (0.961)
Emotional investment	4.580 (1.697)
Focus of attention	4.954 (0.300)
Presence	4.299 (0.388)
Flow	4.848 (0.150)

However, males reported higher mean values than females on all 3 constructs of the immersion model (see Figure 7). These differences were significant in the case of engagement (Student *t* test P=.048, P<.05) and total immersion (Student *t* test P=.033, P<.05; see Table 3 and Figure 7).

Figure 7. Mean values for engrossment, engagement, and total immersion for males and females. Significant differences (P<.05) are marked with an asterisk (*).



XSL•FO
RenderX

Table 3. Differences between males and females on mean values for engrossment, engagement and total immersion.

Construct	Male	Female	Student *t* test		
	Mean (SD)	Mean (SD)	*t*	*df* [a]	*P* value
Engagement	5.565 (0.849)	5.108 (0.750)	–2.029	49	.048
Engrossment	4.976 (0.775)	4.622 (1.063)	–1.301	49	.20
Total immersion	4.994 (1.307)	4.303 (0.950)	–2.189	49	.03

[a]*df*: degree of freedom.

Discussion

Principal Findings

Recently, many researchers have paid increasing attention to the application of AR games for cognitive development. However, with the lack of validated instruments for measuring users' immersion in AR apps, the scientific evaluation of emerging AR games is still missing in the literature. Therefore, this study aimed to investigate students' immersion experiences in a novel MAR n-back game for WM training by using 2 existing immersion questionnaires. As explained earlier, IEQ has the potential to compare users' immersion with 2 different types of games [27], such as a traditional (2D) n-back and a MAR game; therefore, we used the modified IEQ as the immersion instrument in a randomized controlled experiment. This aims to explore students' immersion experience with both versions of n-back games and investigate whether students can obtain stronger immersion in the MAR n-back game than the 2D n-back game. However, as the ARI questionnaire has been widely accepted as a valid tool for assessing immersion in the context of AR apps [20], it was used to measure only a single group of students, with the MAR n-back game in the second experiment. The purpose of this experiment was to further explore students' immersion experiences with the MAR n-back game.

The results show that the correlation between the 2 immersion measures of 27 5-point scale questions and one 10-point scale question of the modified IEQ was positive and significant. This indicates that these 2 different sets of questions were still strongly consistent and connected with each other, which proves the reliability of the modified IEQ. The findings show that while the MAR n-back group had a higher mean value on all immersion measures (see Table 1) than the 2D n-back group, there were no significant differences in the immersion measures between groups on the modified IEQ (27 5-point scale items and one 10-point scale item). The lack of significant differences in immersion may imply that the 2 versions of the n-back games used in this research, although being visually different in many aspects, did not produce obvious different levels of immersion in players. As Figure 1 shows, the traditional (2D) n-back game was played on the screen of a device, and the MAR n-back game was performed in the real environment. Although the MAR n-back game has more simulating graphics, both games have the same motivation rules. Moreover, the participants may be influenced by real-world surroundings, such as sound and light, during the MAR n-back game training in Figure 5 [38,41,42]. It may be that for MAR apps to be more immersive, more is required than simply translating an existing 2D game into an AR version, as we did in this study. On the other hand, participants' scores on the modified IEQ, in both the 2D and MAR conditions, were similar to the scores in previous research using immersive experiences [27]. This suggests that participants were experiencing immersion in both versions of the n-back training games.

Furthermore, Table 2 shows that the mean values of the second experiment for 6 subdimensions with 51 students all showed relatively higher results compared with the average of the 7-point (average 3.500). We can see participants paid attention to the MAR n-back game and felt it was interesting. These results illustrate that a good level of immersion within participants has been established. However, there is still potential for improvement, as the mean value of the 6 dimensions (see Table 2) are all slightly less than those measured in the study by Salar et al [31]. At this stage, the results of using the ARI questionnaire with participants in the second experiment proved that a MAR n-back game can establish participants' immersion experiences.

Previous studies argue that immersion is a desired outcome of the gaming experience and can promote participants' enjoyment and engagement in tasks if their level of involvement increases in the games [27]. Performing and training related to experiences based on immersive technology are assumed to be reliant on the degree of achieved immersion in the games [28,29]. MAR apps bring a different experience for its users, which enhances interactive involvement with the real surroundings without physical constrains [5,18,22]. The findings show that the MAR n-back game (WM training task) can provide users with an immersive and interactive experience, which is beneficial for promoting enjoyment and engagement in training [21]. As the potential affordances of AR-based tasks can improve cognitive functions, such as attention and memory [6-8], this study indicates that the MAR games also have the potential to promote cognition functions (eg, WM in this study), as all the participants experienced immersion during the MAR n-back training.

Unexpectedly, the findings of the second experiment revealed that males gained higher mean values than females on all 3 constructs of the immersion model (see Figure 7), with significantly stronger values for engagement and total immersion (see Table 3 and Figure 7). The immersion theory of Brown and Cairn [30] defines that engagement is the first level of immersion, which indicate that player is willing to invest time and make effort in learning and controlling the game. With the same assignment of manipulating the MAR n-back game, we can see males were willing to make more efforts in controlling the MAR games than females at the first learning stage.

XSL•FO
RenderX

However, males and females experienced no significant differences in the engrossment level, which involves more emotional investment on the games. This result may be influenced by the fact that all the students were required instead of self-driven to play the MAR n-back game, hence the emotional input process is short but directly enters the total immersion level. The results also show that male students researched significantly stronger sense of becoming lost in the MAR games (total immersion level) than female students. These results coincide with the finding that male players experienced deeper involvement with the AR environment, which in turn resulted in feeling higher levels of immersion than female players [34].

In conclusion, by conducting 2 experiments by using the modified IEQ and ARI questionnaire, we found that both groups of students were experiencing immersion in the MAR n-back game. As discussed in the Introduction section, MAR games are very different from other digital games as they can present 3D elements and software assets in real environments without external markers. However, there is an observed lack of studies on investigating immersion in MAR games. This study shows that the modified IEQ and ARI questionnaire have the potential to be used as an instrument to measure immersion in MAR settings. Also, as results show that MAR cognitive games can provide users with an immersive and interactive experience, this suggests the affordance of MAR games in improving users' cognitive functions. However, the MAR n-back game did not produce obvious different levels of immersion in players than the traditional (2D) n-back game in this study, this may imply that the design of the MAR games should have obvious differences from 2D versions or other type of games in order to increase immersion levels. Finally, we found that male players experienced stronger immersion than female players in the MAR n-back game, which further supports that males and females perceive different levels of immersion when interacting with virtual environments [36]. In the end, we believe these findings can contribute to the development of MAR games in the field of cognitive training for AR designers and also for the researchers who are interested in exploring users' immersion in MAR settings.

Limitations and Future Study

Although this study has a number of practical implications, there are a number of limitations that must be considered. First, in both experiments, the immersion questionnaire was completed by participants after the practice sessions were completed. Although students finished immersion questionnaires immediately after the game practices, their immersion experiences may not be accurate, as immersion is an instantaneous state that would disappear after the activity is finished. Second, in the second experiment, only a single group of students completed the ARI questionnaire to measure their immersion with the MAR n-back game. Further, a randomized controlled experiment needs to be considered to explore whether the ARI questionnaire can be used to measure immersion within the non-AR games. Third, an experiment with the modified IEQ should be repeated with an app that really differs in an AR version from its 2D version game to see whether still no obvious different levels of immersion in players. It is important to further investigate the affordances of AR technology in building games. Finally, as only one MAR game has been tested using the modified IEQ and ARI questionnaires in this study, more MAR games need to be tested in the future to strengthen the findings of this study.

Conclusion

In this study, we used two questionnaires, the IEQ and ARI, to investigate immersion in an AR working memory game. In addition, we compared immersion levels between the AR game and a 2D version of the same game. We found that the AR working memory game produced reasonable levels of immersion in players (on both questionnaires), which may contribute to its effectiveness as a cognitive training program. However, there were no significant differences between immersion levels in the AR game and the 2D version (on the IEQ). We also found different levels of immersion experience in men and women on two constructs of the ARI, with men showing significantly higher levels of engagement and total immersion than women in the AR game.

Acknowledgments

This study was sponsored by grant CCA210256 from the National Youth Project of National Education Sciences Planning Project: the Intervention Study of Markerless Augmented Reality Cognitive Training Game on Working Memory of Children with Autism Spectrum Disorder in China. The authors would like to thank the reviewers for providing valuable comments that helped us to greatly improve this manuscript.

Authors' Contributions

BZ designed and conducted the study, contributed to the design of the n-back games, collected the data, performed statistical analyses, and wrote the manuscript. NR advised on the design of the n-back games, assisted with analyses, and read and approved the final manuscript.

Conflicts of Interest

None declared.

Editorial Notice

XSL•FO
RenderX

This randomized study was only retrospectively registered, explained by authors as follows: "The funders of the research ... did not require the registration." The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and the study was considered formative, guiding the development of the application. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1074 KB - games_v9i4e27036_app1.pdf]

References

1. Tulbure BT, Siberescu I. Cognitive training enhances working memory capacity in healthy adults: a pilot study. Procedia Soc Behav Sci 2013 May;78:175-179. [doi: 10.1016/j.sbspro.2013.04.274]
2. van Dongen-Boomsma M, Vollebregt MA, Buitelaar JK, Slaats-Willemse D. Working memory training in young children with ADHD: a randomized placebo-controlled trial. J Child Psychol Psychiatry 2014 Aug 15;55(8):886-896. [doi: 10.1111/jcpp.12218] [Medline: 24628438]
3. Bahar-Fuchs A, Clare L, Woods B. Cognitive training and cognitive rehabilitation for mild to moderate Alzheimer's disease and vascular dementia. Cochrane Database of Systematic Reviews 2013;6(6):10. [doi: 10.1002/14651858.cd003260.pub2]
4. Scanlon L, O'Shea E, O'Caoimh R, Timmons S. Usability and validity of a battery of computerised cognitive screening tests for detecting cognitive impairment. Gerontology 2016 Jun 24;62(2):247-252. [doi: 10.1159/000433432] [Medline: 26113397]
5. Brito PQ, Stoyanova J. Marker versus markerless augmented reality. Which has more impact on users? Int J of Human Comput Interact 2017 Nov 07;34(9):819-833. [doi: 10.1080/10447318.2017.1393974]
6. Boletsis C, McCallum S. Augmented reality cubes for cognitive gaming: preliminary usability and game experience testing. IJSG 2016 Mar 01;3(1):3-8. [doi: 10.17083/ijsg.v3i1.106]
7. Hoe Z, Lee I, Chen C, Chang K. Using an augmented reality-based training system to promote spatial visualization ability for the elderly. Univ Access Inf Soc 2017 Nov 14;18(2):327-342. [doi: 10.1007/s10209-017-0597-x]
8. Shema-Shiratzky S, Brozgol M, Cornejo-Thumm P, Geva-Dayan K, Rotstein M, Leitner Y, et al. Virtual reality training to enhance behavior and cognitive function among children with attention-deficit/hyperactivity disorder: brief report. Dev Neurorehabil 2019 Aug 17;22(6):431-436. [doi: 10.1080/17518423.2018.1476602] [Medline: 29771624]
9. Bacca AJ, Baldiris NS, Fabregat GR, Graf S. Augmented reality trends in education: a systematic review of research and applications. J Educ Technol Soc 2014;17(4):133-134 [FREE Full text]
10. Kim MJ. A framework for context immersion in mobile augmented reality. Automation Construction 2013 Aug;33:79-85. [doi: 10.1016/j.autcon.2012.10.020]
11. Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B. Recent advances in augmented reality. IEEE Comput Grap Appl 2001;21(6):34-47. [doi: 10.1109/38.963459]
12. Lu S, Liu Y. Integrating augmented reality technology to enhance children’s learning in marine education. Env Educ Res 2014 May 06;21(4):525-541. [doi: 10.1080/13504622.2014.911247]
13. Sungkur RK, Panchoo A, Bhoyroo NK. Augmented reality, the future of contextual mobile learning. ITSE 2016 Jun 20;13(2):123-146. [doi: 10.1108/itse-07-2015-0017]
14. Chen C, Chen W, Peng J, Cheng B, Pan T, Kuo H. A real-time markerless augmented reality framework based on SLAM technique. 2017 Presented at: 2017 14th International Symposium on Pervasive Systems, Algorithms and Networks & 2017 11th International Conference on Frontier of Computer Science and Technology & 2017 Third International Symposium of Creative Computing (ISPAN-FCST-ISCC); 2017; Exeter p. 27-132. [doi: 10.1109/ispan-fcst-iscc.2017.87]
15. Zhou F, Duh H, Billinghurst M. Trends in augmented reality tracking, interaction and display: a review of ten years of ISMAR. 2008 Presented at: 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality; 2008; Cambridge p. 193-202. [doi: 10.1109/ismar.2008.4637362]
16. Hammady R, Ma M, Temple N. Augmented reality and gamification in heritage museums. In: Marsh T, Ma M, Oliviera M, Baalsrud Hauge J, Göbel S, editors. Serious Games. Cham: Springer; 2016:181-187.
17. Chowdhary Y, Garland K, Blumenfeld S. System and method of markerless injection of ads in AR. Civic Resource Group International Inc 2017 Dec;US patent No. 9,852,550:1.
18. Comport A, Marchand E, Pressigout M, Chaumette F. Real-time markerless tracking for augmented reality: the virtual visual servoing framework. IEEE Trans Visual Comput Graphics 2006 Jul;12(4):615-628. [doi: 10.1109/tvcg.2006.78]
19. Hammady R, Ma M, Powell A. User experience of markerless augmented reality applications in cultural heritage museums:"Museumeye" as a case study. In: De Paolis L, Bourdot P, editors. Augmented Reality, Virtual Reality, and Computer Graphics. Cham: Springer; 2018:349-369.

XSL•FO
RenderX

20. Georgiou Y, Kyza EA. The development and validation of the ARI questionnaire: an instrument for measuring immersion in location-based augmented reality settings. Int J Human Comput Studies 2017 Feb;98:24-37. [doi: 10.1016/j.ijhcs.2016.09.014]
21. Dilek U, Erol M. Detecting position using ARKit. Phys Educ 2018 Jan 10;53(2):025011. [doi: 10.1088/1361-6552/aaa0e6]
22. Gao Q, Wan T, Tang W, Chen L. A stable and accurate marker-less augmented reality registration method. 2017 Presented at: International Conference on Cyberworlds (CW); 2017; Chester p. 41-47. [doi: 10.1109/cw.2017.44]
23. Lee H, Chung S, Lee W. Presence in virtual golf simulators: the effects of presence on perceived enjoyment, perceived value, and behavioral intention. New Media Soc 2012 Nov 26;15(6):930-946. [doi: 10.1177/1461444812464033]
24. Lee Y, Shan L, Chen C. System development of immersive technology theatre in museums. In: Virtual, Augmented and Mixed Reality. Systems and Applications. Cham: Springer; 2013:400-408.
25. Boletsis C, Mccallum S. Augmented reality cube game for cognitive training: an interaction study. Stud Health Technol Inform 2014;200:81-87. [Medline: 24851967]
26. Georgiou Y, Kyza EA. Relations between student motivation, immersion and learning outcomes in location-based augmented reality settings. Comput Human Behav 2018 Dec;89:173-181. [doi: 10.1016/j.chb.2018.08.011]
27. Jennett C, Cox AL, Cairns P, Dhoparee S, Epps A, Tijs T, et al. Measuring and defining the experience of immersion in games. Int J Human Comput Studies 2008 Sep;66(9):641-661. [doi: 10.1016/j.ijhcs.2008.04.004]
28. Brooks K. There is nothing virtual about immersion: narrative immersion for VR and other interfaces. 2003. URL: http://alumni.media.mit.edu/~brooks/storybiz/immersiveNotVirtual.pdf [accessed 2007-05-01]
29. Cheng M, She H, Annetta L. Game immersion experience: its hierarchical structure and impact on game-based science learning. J Comput Assisted Learning 2014 Apr 14;31(3):232-253. [doi: 10.1111/jcal.12066]
30. Brown E, Cairns P. A grounded investigation of game immersion. Human Factors Comput Syst 2004:1279-1300. [doi: 10.1145/985921.986048]
31. Salar R, Arici F, Caliklar S, Yilmaz RM. A model for augmented reality immersion experiences of university students studying in science education. J Sci Educ Technol 2020 Jan 02;29(2):257-271. [doi: 10.1007/s10956-019-09810-x]
32. Witmer BG, Singer MJ. Measuring presence in virtual environments: a presence questionnaire. Presence 1998 Jun;7(3):225-240. [doi: 10.1162/105474698565686]
33. Agarwal R, Karahanna E. Time flies when you're having fun: cognitive absorption and beliefs about information technology usage. MIS Quarterly 2000 Dec;24(4):665. [doi: 10.2307/3250951]
34. Kallioniemi P, Heimonen T, Turunen M, Hakulinen J, Keskinen T, Pihkala-Posti L, et al. Collaborative navigation in virtual worlds: how gendergame experience influence user behavior. Proc 21st ACM Symp Virtual Reality Softw Technol 2015:173-182. [doi: 10.1145/2821592.2821610]
35. Slater M, Steed A, McCarthy J, Maringelli F. The influence of body movement on subjective presence in virtual environments. Hum Factors 1998 Sep 23;40(3):469-477. [doi: 10.1518/001872098779591368] [Medline: 9849105]
36. Lachlan K, Krcmar M. Experiencing presence in video games: the role of presence tendencies, game experience, gender, and time spent in play. Commun Res Rep 2011 Feb 02;28(1):27-31. [doi: 10.1080/08824096.2010.518924]
37. Bressler D, Bodzin A. A mixed methods assessment of students' flow experiences during a mobile augmented reality science game. J Comput Assisted Learning 2013 Apr 01;29(6):505-517. [doi: 10.1111/jcal.12008]
38. McCall R, Wetzel R, Löschner J, Braun A. Using presence to evaluate an augmented reality location aware game. Pers Ubiquit Comput 2010 Jun 30;15(1):25-35. [doi: 10.1007/s00779-010-0306-8]
39. Csikzentmihaly M. Flow: The Psychology of Optimal Experience. New York: Harper & Row; 1990.
40. Slater M, Usoh M, Steed A. Depth of presence in virtual environments. Presence Teleoperators Virtual Envir 1994 Jan;3(2):130-144. [doi: 10.1162/pres.1994.3.2.130]
41. Reid J, Hull R, Clayton B, Melamed T, Stenton P. A research methodology for evaluating location aware experiences. Pers Ubiquit Comput 2010 Jun 30;15(1):53-60. [doi: 10.1007/s00779-010-0308-6]
42. Benyon D. Presence in blended spaces. Interact Comput 2012 Jul;24(4):219-226. [doi: 10.1016/j.intcom.2012.04.005]
43. Jaeggi SM, Buschkuehl M, Perrig WJ, Meier B. The concurrent validity of the N-back task as a working memory measure. Memory 2010 May;18(4):394-412. [doi: 10.1080/09658211003702171] [Medline: 20408039]
44. Jones MR, Katz B, Buschkuehl M, Jaeggi SM, Shah P. Exploring n-back cognitive training for children with ADHD. J Atten Disord 2020 Mar 07;24(5):704-719 [FREE Full text] [doi: 10.1177/1087054718779230] [Medline: 29877128]
45. Jaeggi S, Buschkuehl M, Etienne A, Ozdoba C, Perrig W, Nirkko A. On how high performers keep cool brains in situations of cognitive overload. Cogn Affect Behav Neurosci 2007 Jun 01;7(2):75-89. [doi: 10.3758/cabn.7.2.75] [Medline: 17672380]
46. Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A 2008 May 13;105(19):6829-6833 [FREE Full text] [doi: 10.1073/pnas.0801268105] [Medline: 18443283]
47. Jaeggi SM, Schmid C, Buschkuehl M, Perrig WJ. Differential age effects in load-dependent memory processing. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2009 Dec 19;16(1):80-102. [doi: 10.1080/13825580802233426] [Medline: 18686052]
48. Wang M, Callaghan V, Bernhardt J, White K, Peña-Rios A. Augmented reality in education and training: pedagogical approaches and illustrative case studies. J Ambient Intell Human Comput 2017 Jul 8;9(5):1391-1402. [doi: 10.1007/s12652-017-0547-8]

XSL•FO
RenderX

Abbreviations

AR: augmented reality
ARI: Augmented Reality Immersion
IEQ: Immersive Experiences Questionnaire
MAR: markerless augmented reality
VR: virtual reality
WM: working memory

Edited by N Zary; submitted 10.01.21; peer-reviewed by H Söbke, M Eckert, R Lundin, S Kruger; comments to author 19.03.21; revised version received 21.04.21; accepted 29.05.21; published 12.10.21.

Please cite as:
Zhang B, Robb N
Immersion Experiences in a Tablet-Based Markerless Augmented Reality Working Memory Game: Randomized Controlled Trial and User Experience Study
JMIR Serious Games 2021;9(4):e27036
URL: https://games.jmir.org/2021/4/e27036
doi:10.2196/27036
PMID:34636738

©Bo Zhang, Nigel Robb. Originally published in JMIR Serious Games (https://games.jmir.org), 12.10.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Game Design in Mental Health Care: Case Study–Based Framework for Integrating Game Design Into Therapeutic Content

Panote Siriaraya[1], PhD; Valentijn Visch[1], PhD; Marilisa Boffo[2], PhD; Renske Spijkerman[3], PhD; Reinout Wiers[4,5], PhD; Kees Korrelboom[6], PhD; Vincent Hendriks[3], PhD; Elske Salemink[7], PhD; Marierose van Dooren[1], PhD; Michael Bas[8], PhD; Richard Goossens[1], PhD

[1]Faculty of Industrial Design Engineering, Delft University of Technology, Delft, Netherlands
[2]Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, Netherlands
[3]Parnassia Addiction Research Centre, Parnassia Psychiatric Institute, The Hague, Netherlands
[4]Addiction, Development and Psychopathology Lab, Department of Psychology, University of Amsterdam, Amsterdam, Netherlands
[5]Centre for Urban Mental Health, University of Amsterdam, Amsterdam, Netherlands
[6]Tilburg School of Social and Behavioural Sciences, Tilburg University, Tilburg, Netherlands
[7]Social and Behavioral Sciences, Utrecht University, Utrecht, Netherlands
[8]&RANJ Serious Games, Rotterdam, Netherlands

Corresponding Author:
Panote Siriaraya, PhD
Faculty of Industrial Design Engineering
Delft University of Technology
Mekelweg 5
Delft, 2628 CD
Netherlands
Phone: 31 152789111
Email: spanote@gmail.com

Abstract

While there has been increasing interest in the use of gamification in mental health care, there is a lack of design knowledge on how elements from games could be integrated into existing therapeutic treatment activities in a manner that is balanced and effective. To help address this issue, we propose a design process framework to support the development of mental health gamification. Based on the concept of experienced game versus therapy worlds, we highlight 4 different therapeutic components that could be gamified to increase user engagement. By means of a Dual-Loop model, designers can balance the therapeutic and game design components and design the core elements of a mental health care gamification. To support the proposed framework, 4 cases of game design in mental health care (eg, therapeutic protocols for addiction, anxiety, and low self-esteem) are presented.

(JMIR Serious Games 2021;9(4):e27953) doi:10.2196/27953

KEYWORDS

design models; gamification; case studies; mental health; eHealth

Introduction

There has been growing interest in the use of gamification in the field of health care. The concept of gamification proposes that motivational elements drawn from the field of game design could be applied to a positive effect in a nongame context such as health care [1,2]. For example, this approach has been used to help users manage symptoms related to diabetes [3] and allow patients with cancer to better understand their condition [4]. The field of mental health care in particular has seen increasing interest in the use of gamification, especially to enhance the effectiveness of psychological therapies, often by improving adherence and engagement in the therapeutic activities (see [5,6]). It is often essential for clients in treatment to adhere to their given therapeutic assignments [7], but in practice this is quite difficult [8]. Therefore, various gamified therapeutic activities have been developed and used for the treatment of conditions such as attention deficit hyperactivity disorder [9].

Incorporating game elements into existing mental health interventions can be difficult as designers must constantly take into account the opportunities and restrictions in both the domain of game design and mental health treatment [10]. For example,

XSL•FO
RenderX

when modifying an underlying treatment activity based on a specific game design strategy, designers must determine in what way this can be done to provide enough *game value* (ie, to motivate or persuade clients to adopt specific healthy behaviors or change maladaptive behaviors), while justifying the potential impact on the *therapeutic value* (ie, the capacity to reduce psychological symptoms). Adding more game elements to improve engagement results in a more enjoyable game experience; however, the underlying therapeutic aspect might be negatively influenced [10]. Arbitrarily adding game elements without careful consideration of the target therapeutic audience risks creating a gamification which alienates players [11,12], whereas introducing game mechanics without considering their influence on the relationship with the underlying therapeutic activity risks corroding the effectiveness of the therapeutic mechanisms themselves [13]. As such, when creating a mental health care gamification, it is extremely important to consider the balance and relationship between the game and therapy elements in gamification. Often, the very nature of such issues means that there is no clear-cut guideline for the integration of game design and therapy elements in gamified health interventions [14].

This paper aims to address this issue by providing a design framework to enhance both the conceptual and practical knowledge of gamification in mental health care. It is based on our reflections and practical experiences of designing and evaluating 4 different gamifications. These gamifications cover 2 intervention protocols commonly used in mental health treatment (ie, cognitive training and cognitive behavioral therapy) and 2 categories of disorders (ie, externalizing and internalizing disorders). More specifically, in this paper:

- We propose the concept of a "game world" versus "therapy world" experience and identified 4 components of a "therapy world" that could be gamified by design into a "game world" experience with increased user engagement. We then highlight 3 different strategies for how integrated game therapy worlds could be created using the 4 components.
- We created a design process framework to help designers analyze the procedures used in a therapeutic intervention and design the core elements of a mental health care gamification based on the concept of a core-game loop.
- To support our framework, we provide 4 case studies of gamifications of mental health care therapies and analyzed their design process in detail through our framework.

Previous Research on the Design of Digital Games and Gamification in Mental Health Care

Overview

Previously, digital games that have been developed for mental health care often embodied similar characteristics as those found in fully functional entertainment games, although with a health-related purpose at its core [14]. Such games usually contain a carefully designed fully functioning "game space" with their own gameplay mechanics and interactive aesthetics (eg, [15-17]). In some cases the therapeutic tasks and activities themselves are embedded in a game-level structure and players would need to achieve various health-related objectives in order to progress through the game [15,16]. For example, in the SPARX game, users go through different levels that challenge them to acquire core skills that would help them better cope with depression [16]. In other cases, the game space itself is designed to provide awareness (often by encouraging players to reflect through gameplay) and persuades players to adopt and maintain a behavior that improves well-being [4,17]. An example of this could be seen in the Playsafety game where players are shown various scenarios that depict dangerous situations related to drug use and are shown the consequences of making the wrong decision in those scenarios [17]. Overall, the design process of these games generally resembles those that are used to develop serious games in other application domains, such as in education. To design such games, the designer first defines "serious objective or outcome" and then designs an engaging game space (including story flow, rules and interaction–feedback system, etc.) around the desired outcome (ie, the therapeutic goal) [18]. In these games, the original "serious" nongame content generally remains independent from the game space. As such, users might experience a separation between the game content and the nongame content when playing the game and could result in players becoming immersed in the game space while not being engaged with the serious content.

The more recent interests in gamification have led some designers to further examine ways in which game methods and design approaches can be integrated as part of a therapeutic process. Instead of creating a full game space to help motivate clients in mental health treatment to realize a therapeutic objective (which tends to be costly and time consuming), recent approaches in gamification suggest that perhaps discrete game elements could instead be applied to enhance existing tasks within the treatment process (see [19]). It is believed that this would help transform the experience of doing an otherwise mundane task into one which is more engaging for users [20]. A particular appeal of this approach is that it allows for various commonly used game mechanics (such as point systems and leaderboards) and game experience designs (challenge, competition, etc.) to be viewed as "templates," which could be carefully tested and then used as a blueprint to provide gamelike experiences to enhance different therapeutic tasks (see [19]). In the health care field in particular, which values an evidence-based approach when designing interventions, this approach can be particularly appealing as the effect of different game design decisions can be empirically tested and more clearly understood (serious games employing a full-blown game space have traditionally been difficult to validate and examine [21]). Examples of such game mechanics and elements used that have been transposed to a therapeutic task include in-game reward systems (badges, tangible rewards, and social feedback), a narrative driven quiz system [22], and a gamelike audio and visual feedback system combined with a point-based reward system [13,23].

However, many researchers caution against the difficulties and pitfalls of this kind of "integrative" gamification approach. Some researchers argue against the practice of employing discrete game elements without carefully designing for the

XSL•FO
RenderX

emerging player experience [18]. These researchers propose that gamification should be viewed more as a way to create emerging experience of playfulness instead of relying on template game mechanics [24]. More specifically, these researchers argue that focusing only on the mechanics themselves (eg, only arbitrarily adding in elements such as points, badges, and levels) risks striping the games from their most engaging aspects, as the "engagingness" of a game comes not only from a single mechanic, but also rather from a well-designed and thought-out combination of the game mechanics with other elements of the game (such as the rules and narration) and with the serious content itself to create a ludic system that allows users to experience meaningful play and have fun [24,25]. Such researchers posit that to design such a system, it is important to carefully consider the underlying context and understand the goals of users within the serious activity when integrating the game mechanics (possibly by employing a user-centered design approach) [11,26]. Merely adding game elements to a therapeutic activity without careful consideration of the interests of the targeted audience and the nature of the target activity risks the gamification becoming less relevant to the user interests and thus becoming ineffective in motivating users [11]. Overall, in the domain of mental health care in particular, there have been few proposed design models that show the process of how elements from game design could be integrated effectively with the content of a therapeutic activity.

Integrating Game Elements Into Therapeutic Activity Through a Game Therapy Worlds Concept

A key essence in the process of mental health care gamification is the question of how various elements from game design could be incorporated into an existing therapeutic activity to create a more engaging and (at least) equally effective treatment process. Key issues such as which game elements can be added without conflicting with the underlying therapeutic processes, which parts of the treatment can be mixed with game elements, and which parts of the therapeutic content should remain unchanged for the gamified therapy to be effective would need to be addressed. The difficulty of such a process is compounded further by the fact that the therapeutic activities themselves are often black boxes [21]. It is often unknown whether changing the procedure, rules, or interactions of a certain therapeutic activity (to make it more engaging) would also jeopardize the therapeutic effect.

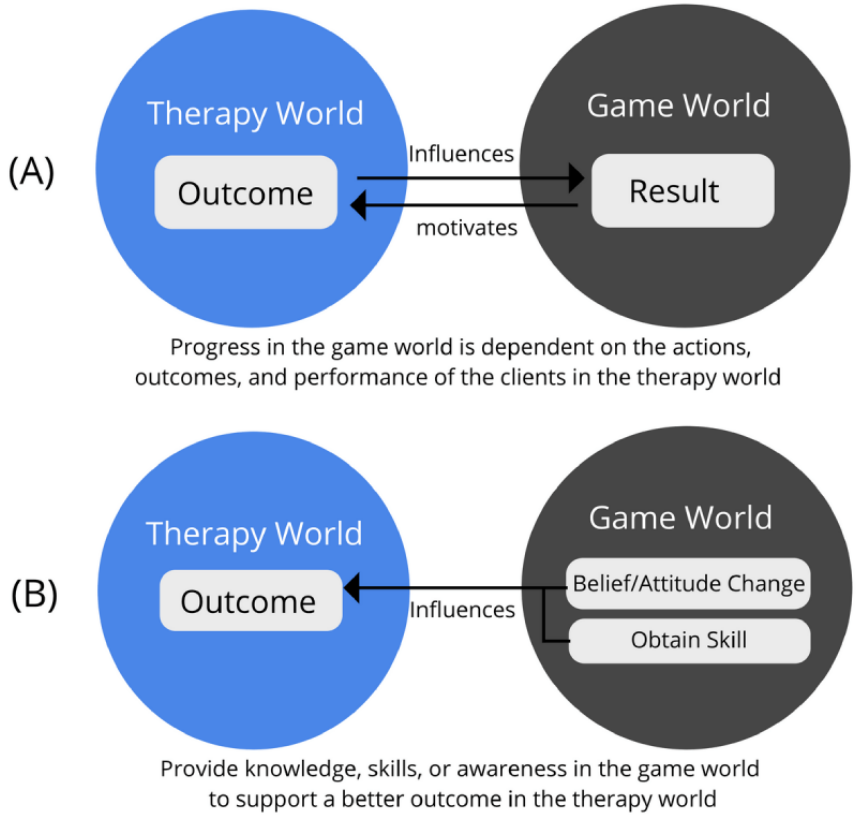
To provide a better understanding of how game elements could be combined with existing therapeutic activities, we draw upon the notion of "worlds," which represent the changing experiences of a person as we examine gamification design in past studies. This notion suggests that when playing an immersive game, players gradually leave their world of daily life and move toward a fun and engaging game world [27]. Their experience changes and they feel "transported" [28] into a world of play. Huizinga [27] described this process as "creating a temporary "game" world within the ordinary world" for which he proposed the metaphor of a "magic circle" having its own time and space boundaries and absorbing the player "intensely and utterly." It is exactly this absorbing and engaging quality of game experiences that can be used to enrich mental health care therapies. As such, in our "game therapy worlds" concept, we view games developed for mental health care as similarly consisting of 2 worlds: the game world, which denotes how users experience, interact, and are affected by the designed game experience; and the therapy world, which denotes the experiences, interactions, and effect of the designed therapeutic process within the mental health care game.

Therapy Games Formed From Separated Game and Therapy Worlds

Two strategies have been commonly used in previous studies when designing games for mental health care. The first involves creating a discrete game world (with its own rules, narration, and interaction mechanics), where the performance and actions of users in a therapeutic activity are used to drive their progress in the game world (Figure 1A). An example of such strategy is the Changamoto game, which was developed to encourage youth to fill in a diary for their triggers in cannabis addiction. In the Changamoto game, a separate turn-based strategy game world was created in which players would need to command their own robot team to defeat an opposing team. Progression in the game is dependent on players successfully filling in entries in their diary [29].

XSL•FO
RenderX

Figure 1. Strategies for incorporating a separate game world experience into the therapy world.



The second strategy is to create a discrete game world to provide users with the knowledge and awareness to improve the outcome of an existing treatment (Figure 1B). An example of this is the Re-mission game, which was designed to provide children with knowledge related to their treatment and in doing so improve their adherence to the treatment [4]. The game itself was designed as a third-person action game with the interactions inside the game (ie, destroying cancer cells with weapons such as chemotherapy) designed to make players more aware and understand the importance of complying with their treatment regime.

In both strategies, the game world itself is usually not built upon an existing therapeutic activity but as a separate game space with its own rules and mechanics. The advantage of such a strategy is that there is only a limited relationship between the therapy world and the game world, and therefore there is much less potential for the game world to impact the targeted therapeutic activity. In addition, the engagement potential of the newly created game world is not hindered by the therapy rules, thus allowing for more flexibility when designing the game. These discrete game worlds often function as fully fledged games, containing their own mechanics, rules, and interactions, which are separate from the therapy. Players experience first the game world and then the therapy world. However, the disadvantage of this strategy is that the effectiveness of such games often depends on the entertainment quality of the game world itself and it could be quite costly to develop fully fledged games that are entertaining enough to be effective [14].

Therapeutic Games Formed From Integrated Game and Therapy Worlds

The Elemental Tetrad

Instead of creating a discrete game world to encourage users to achieve a specific therapeutic outcome, different aspects within the therapeutic activity itself could be transformed to become more engaging by using specific elements and approaches drawn from the field of game design [26]. Although there have been different suggestions of what components exist in a game (such as [30,31]), the concept of the elemental tetrad proposed by Schell [32] provides a useful distinction of the broad elements found in a game system. This tetrad proposes 4 interconnected game elements: the mechanics (the procedures, rules, and possible interaction space in the game), the aesthetics (the visual and audio qualities of the game), the technology (the medium in which a game has been implemented), and storyline (the prescripted events that emerge from the user interaction) [32]. While there are serious games in mental health care that have used technology (eg, virtual reality to create a more engaging experience [33]), aesthetics (eg, visualizing the consequence of how different bodily organs are affected by drugs [34]), or the storyline (eg, a game where users play the role of an investigator to learn about the consequences of drug overdoses [35]) as the core element, most gamifications focus on the use of the game mechanics to enhance the underlining experience (such as adding a point base system to create a sense of challenge) with the technology, aesthetics, and storyline being designed around the game mechanic elements [36]. Therefore, we focus on this aspect in our design framework as we analyze the components of an integrated therapy and game world.

XSL•FO
RenderX

The Components of an Integrated Therapy and Game Worlds

To allow us to better understand how a specific game element can be combined with the different aspects of the therapeutic procedure, we look at the components of a therapeutic activity from a game systems perspective from the viewpoint of the game mechanics element proposed by Schell [32]. When viewed from this perspective, there are many aspects of a therapeutic task that are similar in nature to the mechanics of a game system. For example, in a therapeutic activity, clients are generally presented with a list of tasks with different content (referred to as *content samples*) that they would need to resolve to achieve success in the therapy (eg, in the form of homework and out-of-session activities, which are common tasks in psychotherapy and cognitive behavioral treatment procedures [37,38]). These tasks are usually presented in a *structured* format (eg, Clark's model [39] proposes that the situations that clients would confront as part of their social phobia treatment be presented in a hierarchical manner based on difficulty). In each task, users are presented with a meaningful choice where they can decide to take action to complete the task, after which feedback is provided by the therapist to help clients evaluate their performance (referred to as the *performance space*). An example could be seen from an Interpretation Bias Modification training procedure for social anxiety, where clients would choose whether a word representing a threat or benign interpretation is related to an ambiguous situation and would receive positive feedback when they relate benign interpretations toward ambiguous sentences [40]. Whether or not the user is successful in his/her task is determined by the *rules* of the therapeutic activity. For example, the retraining exercises which are used to change maladaptive cognitive biases based on the Dual-Process model of addiction often have specific rules that would need to be followed to be effective (such as to make avoidance actions in response to alcoholic stimulus within a set period) [41]. The cumulative completion of each therapeutic task is expected to result in a better health-related outcome. Overall, Table 1 provides an overview of the similarities between the components in the game world and the therapy procedures and provides examples of several treatment approaches and therapeutic activities where these components could be identified.

Table 1. Components of a therapeutic activity when viewed from a game systems perspective.

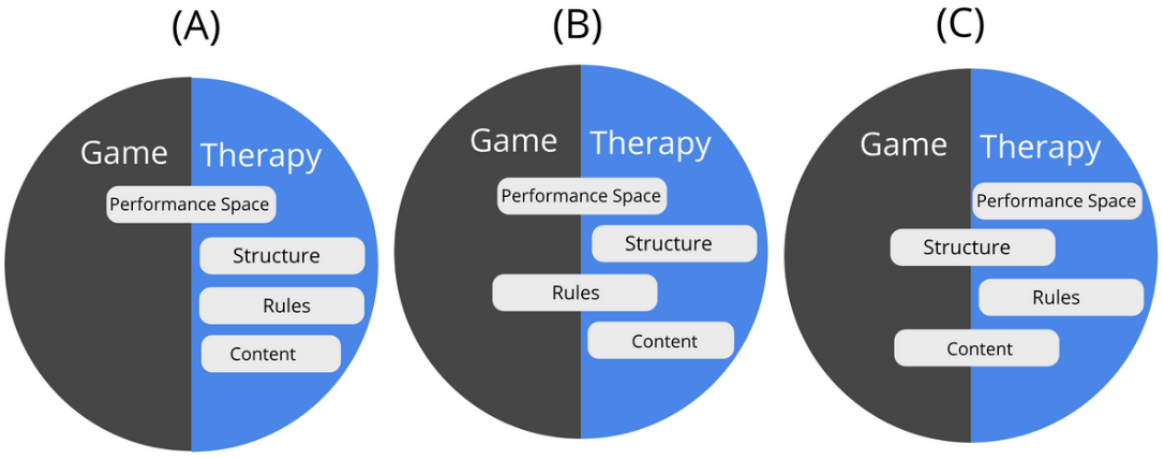
Component to be balanced	In a game world	In a therapeutic world
Performance space	The actions or meaningful choices available to players to act upon. Feedback is given to users as indicators to convey the results of their actions in the game.	The actions and choices are presented to the clients in a therapeutic activity, which can be acted upon to achieve success in the therapy. Feedback is conveyed to clients to report the results of their actions and is generally aimed at providing guidance in future choices. Example: *In contingency management programs, clients are presented with a task to abstain from drug use within a set period. They are presented with 2 choices of potential actions that could be taken: either choose to use drugs or not to use drugs within the set period. If clients choose not to use drugs, they would be presented with monetary rewards as a form of feedback to reinforce their positive behavior* [42].
Structure	Referred to as the "steps of play," the structure component highlights the progression of players in the game [31].	A structured organization of tasks presented to guide clients to progress through the therapy. Example: *As part of weekly homework assignments in the cue exposure therapy to treat social phobia, clients are exposed to different social situations that they normally avoid. These situations are presented in a hierarchal structure based on their difficulty to help clients build their confidence, with easier tasks being presented before more difficult ones* [39].
Rules	The principles that govern the consequences of user actions and results within the game.	The rules or principles set by the therapy procedure that relate to the success or failure in achieving the health-related outcome. Example: *In the selective response inhibition training module used as part of the cognitive bias modification training to treat alcohol addiction, clients are presented with images representing alcoholic and nonalcoholic beverages. Clients would need to respond when shown images of nonalcoholic beverages in a timely manner by pressing a key and withhold their response when shown images of alcoholic beverages. Responding to alcoholic beverages or not responding in time results in a failure* [43].
Content Sample	The different events given to players through the game that they must overcome to achieve success. Referred to as a pulse [31].	The different content samples drawn from a therapeutic activity framework that are presented to the clients, who need to resolve them to achieve success in the therapeutic activity. Example: *In the interpretation bias modification training used to treat anxiety, clients are presented with ambiguous scenarios representing anxiety-related situations. Each scenario ends with a word fragment which clients need to solve toward a positive solution. Clients need to go through a series of multiple different scenarios (ie, content sample) representing different anxiety-related situations and complete the word fragment in each one, reinforcing a beneficial interpretation of ambiguous information* [44].

XSL•FO
RenderX

Strategies for Creating an Integrated Therapy and Game World

Based on game and therapy world components, we have identified 3 different strategies that have been generally used to create an integrated game world that overlaps with an existing therapeutic activity (Figure 2). One strategy, the *reward systems strategy* (Figure 2A), is built around enhancing motivation by creating a reward system based on *the performance* of users in the therapeutic activity. For instance, if social competition is thought to be particularly appealing toward a target audience, simple mechanics such as a "leader board" could be used to create a feeling of competition based on how well users perform in the therapeutic activity in relation to others (see [45]).

Figure 2. Three examples of strategies to integrate game worlds from existing therapeutic activities: (A) Creating a reward system based on the results. (B) Restructuring the rules and performance space. (C) Restructuring the content and structure.



Alternatively, another strategy is to *restructure the rules and performance space* (Figure 2B) of the target therapeutic activity to provide a more engaging experience. For example, in our gamification of the goal setting activity used in cognitive behavioral therapy, we added a time constraint *rule* and wager-based mechanism to the *performance space* to provide users with an experience of challenge when completing their goals. Users would set a self-imposed time limit for each goal (with less time receiving more points) and place a wager to reflect how confident they are in their ability to accomplish the task [10]. In the third strategy (Figure 2C), *the content and structure within the therapeutic activity* could be enhanced through *a narrative guidance framework*. This example could be seen in a gamified therapeutic activity used in cognitive behavioral therapy for burnout [22]. In this example, clients are asked to choose positive thoughts in response to stress-inducing work scenarios. In the gamified version, the therapeutic process is presented through the narrative guidance of a fictional character, with metaphors such as monsters representing negative statements, which players need to "shoot down," and the cognitive state of players represented as an overloaded file cabinet that players need to clean up and fill instead with compliments. The whole therapy procedure is segmented into different levels, with players being guided through the levels as they complete each aspect of the therapy (ie, first recognizing negative thoughts, then identifying alternative positive ones and finally practicing with real-life situations) [22].

Overall, integrating a game world directly into an existing therapeutic activity provides several distinct advantages. First, interactions within the game world can be linked more directly to the interactions within the therapeutic activity, thus making them more meaningful and relevant to the therapeutic outcome. This is opposed to a nonintegrated approach with a separate game world (eg, Figure 1), where core player actions in the game space (such as going through various scenarios to learn about the dangers of drugs) play a more supportive role toward the outcome of the therapeutic activities [17]. In the integrated game therapy world used in the burnout gamification for instance, the actions of players in the game world (shooting down negative statements to earn points) are shared with those of the therapy world (rejecting negative thoughts as part of the therapy process) [22]. In addition, as the integrated game world also contains many elements that are shared with an original therapeutic activity, the intrinsic motivation of clients within the therapy world would more likely remain in the integrated game world as well. As such, users are likely to ascribe more value toward their actions in the integrated game world, as they feel it would lead them to a serious therapeutic outcome that helps improve their well-being. This is opposed to the effectiveness of discrete game worlds, which often depend on the entertainment quality of the game world itself. However, there is also the risk that altering different aspects of the therapy when integrating it with the game world could (negatively) influence the potential value of the therapy. For example, in gamifying a cognitive bias modification (CBM) task used to treat alcohol addiction, the authors cautioned that adjustments made to the original CBM paradigm as part of the gamification process (eg, changing the way in which the stimuli is presented, ie, the speed or frequency in which the alcohol image is shown, to make the training more challenging or changing the control paradigm, ie, requiring users to use different keyboard buttons or control devices to respond to the stimulus) could risk making the original training ineffective [46]. Cognitive training tasks in this domain (such as the Go/No-go task) require the stimuli and feedback to be presented at a specific time interval (1500 and 500 ms, respectively) [43], and it is unclear whether changing this original mechanism (eg, shortening the response

XSL•FO
RenderX

time to create different challenge levels) would impact the therapeutic effectiveness.

Our Proposed Method: Designing Gamified Therapeutic Interventions Through a Dual-Loop Design Model

During the design of the gamifications, we noticed how the core interaction process within a therapeutic activity shares many similarities with the concept of the core-game loop [47]. In game design, the core-game loop represents the fundamental action–feedback loop of a game, where players complete a specific task fulfilling an in-game objective and receive feedback based on their action. Players then use this feedback to improve their understanding of the game rules, update their strategies, and move on to complete the next task. The process then repeats itself throughout the game with the tasks becoming progressively difficult. Similarly, a therapeutic intervention or activity could also be viewed as taking place within a recurrent loop (ie, the therapy activity loop), as it generally consists of repeated tasks (although with different content) that clients must undertake to achieve the cognitive or behavioral change that leads to the desired health goal. After completing a task they would receive feedback on their performance and progress onto the next task with a different content but governed by the same rules and principles.

This observation led us to conceptualize the Dual-Loop Design model in which we propose that a core-game loop could be designed around the therapeutic activity loop, thereby enhancing the user experience of each of the therapy components. An earlier example of a gamified therapeutic activity being designed around this concept can be found in our previous studies, where we designed a goal setting gamification around the core-game loop of selecting different tasks leading to desirable therapeutic outcomes [10]. Based on our Dual-Loop Design model, the 3 components of a therapeutic activity which form the main part of a therapeutic loop are presented in Textbox 1.

Textbox 1. Components of a therapeutic activity.

Sample

The sample refers to a specific task presented to the client, which is drawn from a list of possible tasks. For example, in the cue-specific response inhibition training task [43], clients are presented with a specific alcohol stimulus image (an image of a full or empty bottle of beer, etc.) which they would need to respond to.

User action

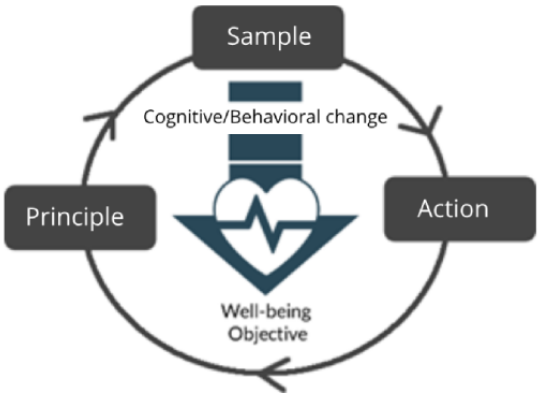
The user actions refer to the actions required of clients after being presented with the sample. This could occur both in the digital domain and in the physical world. For example, in the cue-specific response inhibition training task used in the treatment of alcohol addiction, clients can either choose to respond to the image of the stimulus in the task by pressing a button on the keyboard or choose not to respond by withholding their action.

Therapy principle

The principles refer to the logic within the therapeutic task that determines whether the actions of the user constitute a correct or appropriate response, which when carried out should progressively lead to the desired positive cognitive or behavioral change. For example, in the cue-specific response inhibition training task, clients must respond by making a correct action based on the stimulus presented within a specific time limit to be successful (ie, the desired response would be to press a key on the keyboard when shown a nonalcoholic image and withhold their response when an alcoholic image appears and letting it disappear).

These components as well as the intended cognitive/behavioral change and the overall well-being/health objective are what constitutes the therapeutic activity loop (Figure 3). More specifically, in each iteration of the therapeutic activity loop, the client is given a *sample* task. They would attempt to carry out the *actions* required to complete the task based on the governing *therapy principles*. After completing the task, the client is then presented with the next sample task to complete and the process would repeat itself. The tasks presented to clients could cascade in difficulty or be used to represent a wide range of situations in the therapy (eg, in the CBM training for alcohol addiction, clients are presented with stimuli depicting different types of alcoholic beverages to better cover the large variety of alcohol drinks they may encounter in daily life [44]). By repeating this process, clients are able to change their cognitive state or behavior in a way which leads to a better health outcome.

Figure 3. The therapeutic activity loop.

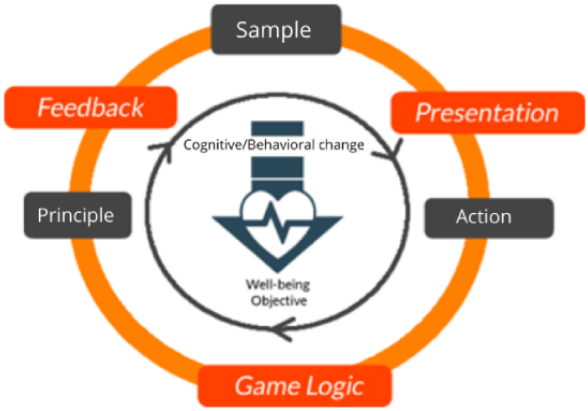


XSL•FO
RenderX

The main gamification process starts by formulating a core-game loop upon the therapeutic activity loop. More specifically, the gamification designer decides on the following 3 aspects (Figure 4).

- Can "presentation design" make the therapy more "gamey" and motivating and what are the implications toward the existing "sample" and "actions" of the therapy? For example, when a narrative metaphor is used to represent negative thoughts or maladaptive characteristic traits as enemy creatures [29], would this make the "sample" and "action" (eg, the task of recording the various triggers of addiction) more interesting?
- Can "game logic design" make the therapy more fun and what are the implications toward the existing "action" and "principles" of the therapy? For example, in our gamification of the approach bias training for alcohol addiction, we added a game logic in which players would need to push alcoholic and pull nonalcoholic pictures based on the beat of the music, instead of at a set time interval as in the original version of the training. This was done to enhance the underlying "action" (ie, the task of pushing away alcohol-related stimuli and pulling closer nonalcohol-related stimuli) and "principles" (ie, players must complete the tasks correctly within a set time limit) of the original therapy.
- Can "feedback design" make the therapy more fun and what are the implications toward the existing "principles" and "sample" of the therapy? For example, when a chain bonus system is added to the cue-specific response inhibition task (ie, users get an increasing number of points for consecutive correct actions, which resets when they make a mistake), would this influence the "principles" and "sample" of the therapy in a positive manner? More specifically, when players are presented with the next "sample" (ie, the task to inhibit their response toward an alcohol-related stimulus), would they put more effort in avoiding mistakes, as they would lose their chain bonus?

Figure 4. A core-game loop encompassing the therapeutic activity loop.

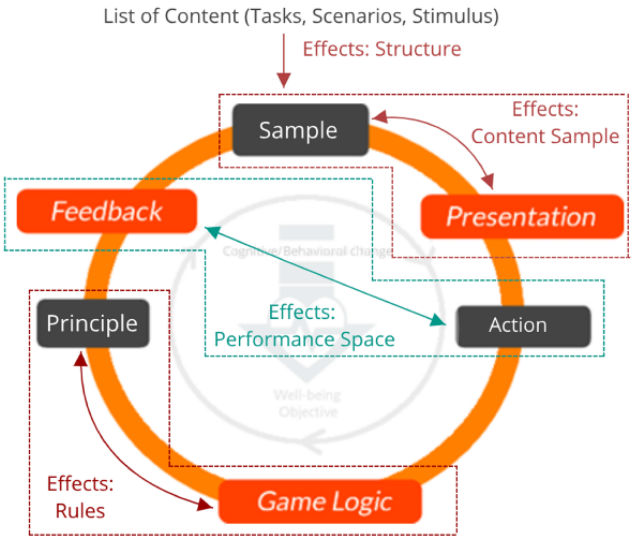


After deciding on the 3 aspects, the designer could then choose to add structural game elements to further enhance the core-game loop experience and encourage sustained engagement across the loops [48]. For instance, a "level progression system" could be added where players would be provided with experience points upon successful accomplishment of each task in the gameplay loop, which they then could use to "level up" their in-game characters, providing them with a sense of progression. Players could also be ranked based on their performance on a "leaderboard" to create a sense of competition. Another example is a "game-level system," where completion of the tasks could be used to unlock different portions of the narrative experience, encouraging players to complete more tasks in the loop to complete the story.

When viewed from the Dual-Loop Design model, the components in the therapy world concept discussed in the "Strategies for Creating an Integrated Therapy and Game World" section serve to highlight how the different aspects of the original therapy could be affected through the gamification design process. Figure 5 provides a summary of the relationship between the therapy game world model and the dual-core–game loop. More specifically, the "performance space" in the therapy is affected mainly by the "feedback" design choice and the "user action" within the original therapeutic activity loop. The "rules" of the therapy are affected mainly by the "game logic" design choice and the "therapy principles" within the original therapeutic activity loop. The "content sample" is affected mainly by the "presentation" design choice and the "sample" within the original therapeutic activity loop. Finally, the structure of the game and therapy world is influenced mainly by the structural game elements, which the designers choose to use in their game design as this affects the nature of how each sample in the loop is drawn and presented to players (eg, a game-level system could be implemented which would mean that in each game loop, the sample that users interact with would be progressively more challenging). To balance the game and therapy world experience, players would need to take into account how their choices in designing the core-game loop influence the elements of the therapy game world.

XSL•FO
RenderX

Figure 5. Effect of the core-game loop on the therapy game world.



The Case Studies

Designing of The Case Study Gamifications

To provide more knowledge regarding the principles that can be used in the gamification of mental health care treatments, we designed and developed 4 different gamifications of existing therapeutic interventions. These gamifications are presented as case studies in the manuscript to show the result of how our proposed model could be used in the practical design and analysis of gamifications for mental health care. The gamifications were designed in collaboration with experts from serious games design as well as researchers and clinicians in the domain of psychological interventions. To ensure the generic applicability of the principles, the gamifications covered 2 main classes of intervention types (ie, cognitive training and cognitive behavioral therapy) and 2 main categories of mental disorders (ie, externalizing and internalizing disorders). In addition, 2 of the gamifications were of computerized cognitive training modules that are mostly implemented without the interference of a therapist, and the other 2 were designed to be used as part of a blended therapy program through face-to-face sessions with a therapist. The 4 gamifications include the ReadySetGoals mobile app (*cognitive behavioral therapy, blended therapy, externalizing disorder*), the Addiction Beater computer application (*cognitive training, computer based, externalizing disorder*), the Zen Garden mobile app (*cognitive behavioral therapy, blended therapy, internalization disorder*), and the Één klein probleempje (Small problems) computer application (*cognitive training, computer based, internalizing disorder*).

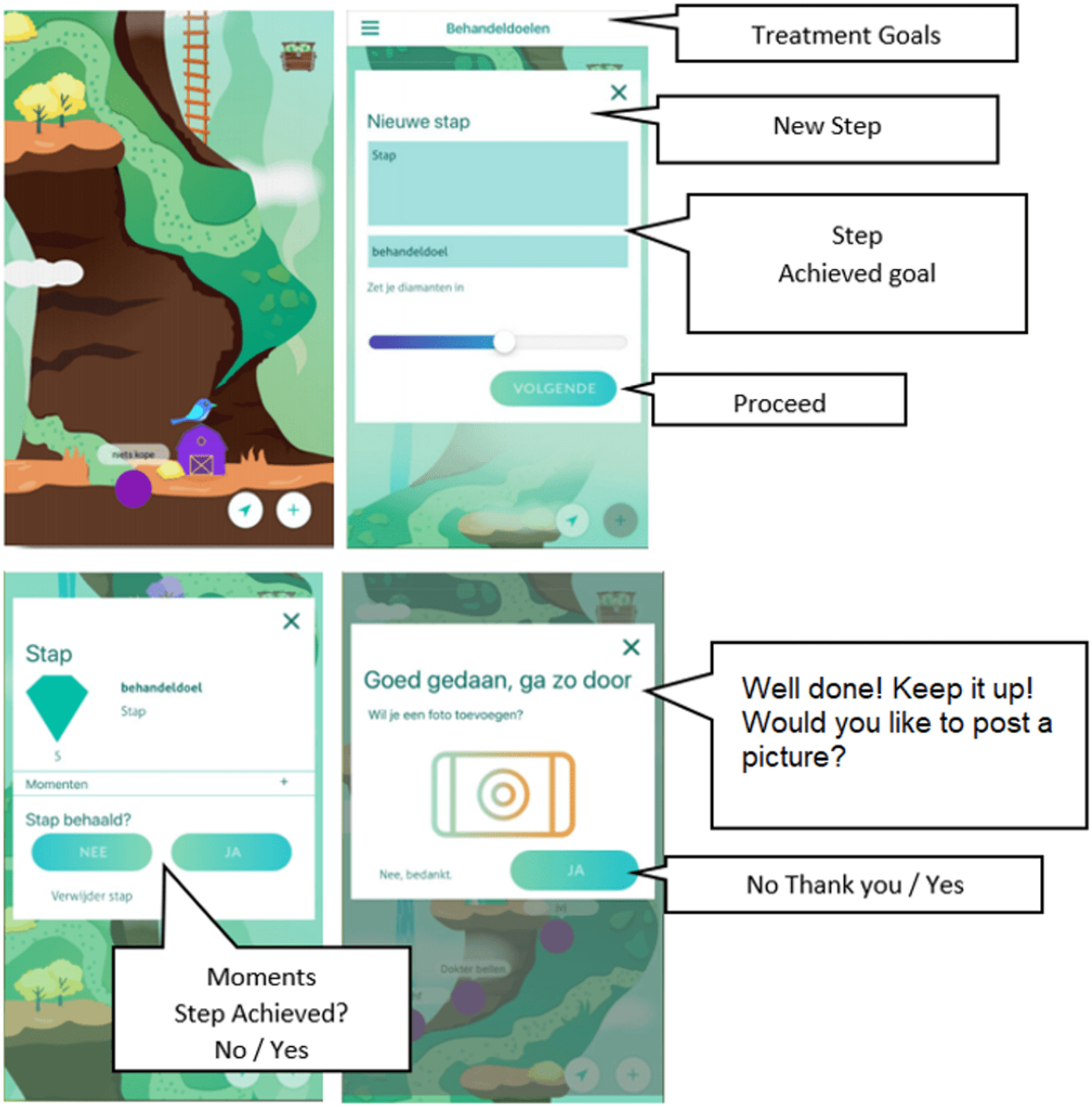
The ReadySetGoals Design

Purpose

The ReadySetGoals is a gamification of the goal setting activity commonly used as part of a treatment plan at the beginning of therapy in mental health care. In addiction treatment, this activity plays a key role in addressing not only substance use problems, but also anxiety and depression symptoms. The goals tend to focus on encouraging clients to change their typical behaviors or aspects of daily life, in order to improve mood and decrease substance consumption. However, it is generally difficult for them to adhere to such goals and clients tend to experience difficulty in putting therapeutic insights into practice. Therefore, the ReadySetGoals was created as a gamified mobile app based on a risk-taking mechanism to encourage users to set and complete goals that are beneficial to their treatment. Players set goals to achieve various therapy-related tasks and then place a wager on how likely they feel they are able to achieve those goals. Structural game elements such as a progression-based reward system (eg, users would progress along a path for each goal set) were implemented. As part of a blended therapy program, the app was used during the face-to-face session between the client and the therapists. During each session, the client would decide together with the therapist on what long-term goals and tasks would be appropriate. In addition, the level of difficulty (and thus the reward which would be received) of each task was decided in collaboration with the therapist. In the following session, the therapist would review the tasks and credit points to the client for each task that was completed satisfactorily, evaluate the overall progress of the long-term goals, and set new tasks for the following week through the ReadySetGoals app. Figure 6 shows screenshots of the gamification (see [10] for in-depth details of the gamification design process).

XSL•FO
RenderX

Figure 6. Screenshots of the ReadySetGoals gamification.

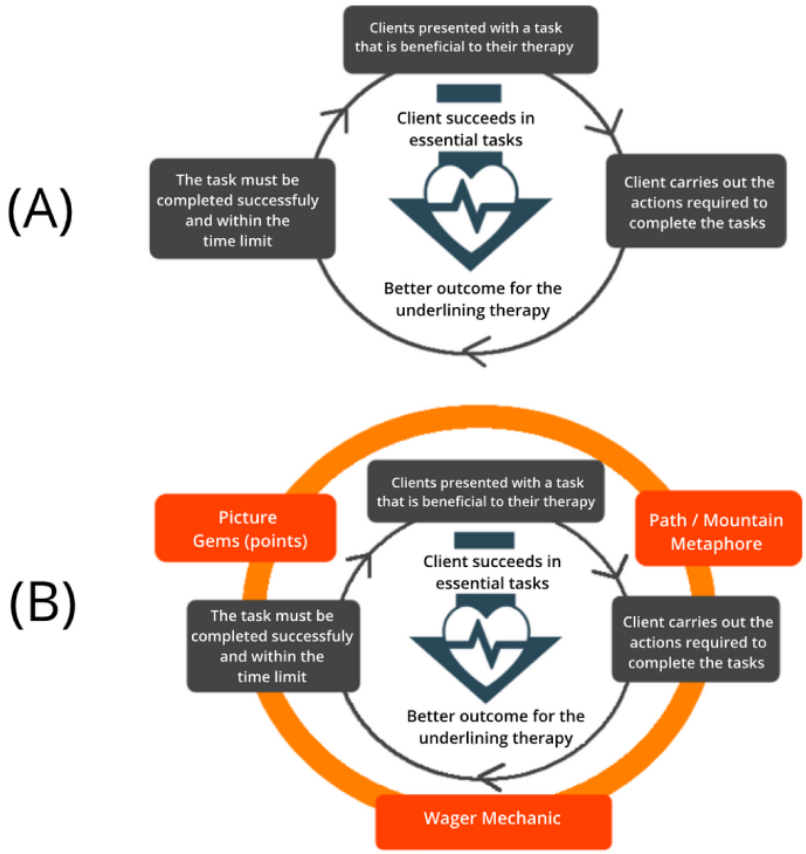


The Gamification Strategy and Dual-Loop Design of the ReadySetGoals App

In the gamification of the ReadySetGoals app, the desired behavioral change was defined as *encouraging users to persist and succeed in tasks that are essential in therapy*. Users are encouraged to set goals to complete the various therapeutic activities that are given to them in different domains (eg, find and carry out alternative rewarding activities such as pleasant hobbies, as a way to become less dependent on drugs). In the original therapeutic goal setting activity, clients discuss with the therapist and decide on a goal which they would need to complete to achieve a better outcome in their treatment. They are then required to break down the goal into smaller tasks (*sample*), decide on a time limit, and then carry out the actions required to complete the tasks (*action*). Clients are successful if they complete the tasks within the agreed time limit (*principles*; see Figure 7A).

XSL•FO
RenderX

Figure 7. The core-game loop and gamification design of ReadysetGoals.



To determine what game elements could be used to increase user engagement in the therapy loop, we carried out multiple sessions including a discussion session; brainstorming session; and survey study with therapists, care managers, domain experts, and clients of an addiction center (see [10]). The results showed that personality traits such as sensation seeking and impulsivity are very prevalent among clients with addiction problems [49] and thus we determined that a risk-taking game mechanic would be appropriate in engaging these clients. Therefore, we decided to add a wager rule as the main game mechanic in the gamification (*game logic*). Based on this mechanic, players receive points (in the form of diamonds) that they could use to "bet" on themselves based on how likely they feel they can complete a specific task. Players are then rewarded in proportion to the risk taken (the amount of wager placed and time limit they set themselves). In addition, we represented the small steps that users are taking to progress within the therapy through the metaphor of climbing a mountain (*presentation*). As users complete tasks, they move further up the mountain until they reach the destination on the top. Users are rewarded with points for successfully completing tasks and they are also able to view pictures taken as proof when they complete their tasks on the mountain (*feedback*; Figure 7B).

The Therapy Game World of the ReadySetGoals App

When designing the ReadySetGoals gamification, we determined that the *structure* and *rules* of the therapy were flexible enough to be adapted without severely impacting the therapeutic effect. By using the mountain metaphor and presenting the smaller steps to obtain the goals on a progression map, we restructured the way tasks are presented to users. For example, in the initial version of ReadySetGoals, users need to complete easy tasks at the beginning before moving on to more difficult ones and they were limited in the number of tasks they could challenge themselves at one time. Adding the wager mechanic also created a rule where players are rewarded more for achieving tasks in a shorter period, thus changing the original rule where players are provided with a fixed deadline. We expected the design of the gamification to have a minimal impact on the original *content* and *performance space* of the therapy (Figure 8). The goals that users need to complete were the same as those given in the original therapeutic setting and the actions that users would need to take also stayed the same. Similar to the original activity, the therapist and client would collaborate together to set the long-term goals and task difficulty during each therapy session (eg, the therapist and client would first select the relevant area in life which they would like to work on and set the goals and tasks together). The client would also set a specific date and time deadline they feel they could accomplish the task and in the following week, the therapist would review the progress of the goals and tasks and provide feedback to the client.

XSL•FO
RenderX

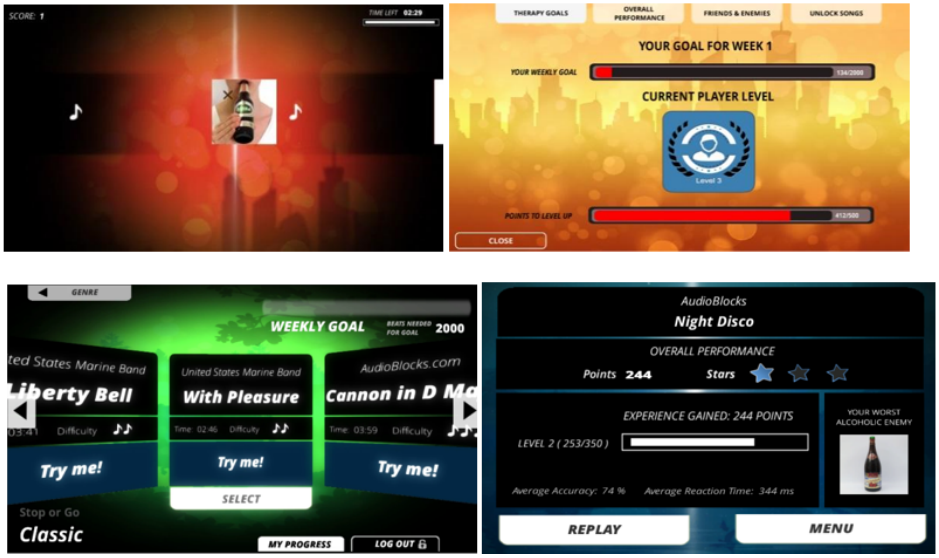
Figure 8. The therapy game world of the ReadySetGoals gamification.

The Addiction Beater

Purpose

The Addiction Beater is a gamification of the therapeutic activities included in different CBM training modules based on a music rhythm game concept and was designed to support the treatment of alcohol addiction. Such CBM modules have been used in digital cognitive training programs as an add-on therapy for addiction disorders, by helping clients to retrain maladaptive cognitive processes in substance abuse (see [50,51]). In the original training, people's motivation to engage and persist in the training was quite low due to the repetitive nature of the task. To address this problem, the Addiction Beater gamification was developed based on a music rhythm game concept to provide a more engaging training experience. Two modules used in CBM training were gamified: the cue-specific response inhibition training and the approach bias training. In the gamification, users are challenged to respond to the alcoholic stimulus based on the beat of the music. For instance, in the approach bias training module, users are challenged to "push away" images of alcoholic drinks using the "up" arrow key and "pull closer" images of nonalcoholic drinks using the "down" arrow key on their keyboard based on the musical beat of different songs. When carrying out the training tasks, users can select from a list of different songs to use as part of their training. The current version has over 20 songs in more than 4 different genres (Rock, Electronic, Classic, International, etc.). As users continue to train, they would be able to gain experience points, which will unlock more difficult songs (Figure 9).

Figure 9. Screenshots of the Addiction Beater game.

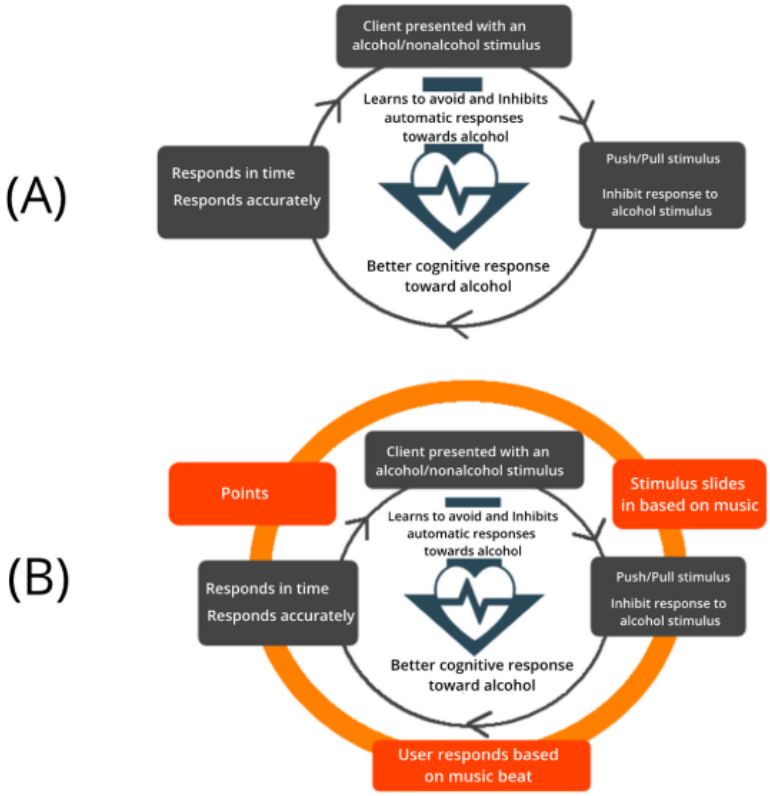


The Gamification Strategy and Core-Game Loop Design of the Addiction Beater

When viewed from the Dual-Loop Design model, in the original cognitive training clients are presented with an image of either an alcoholic or nonalcoholic beverage (*sample*). In the cue-specific response inhibition training task for example, clients need to withhold their response when an image of an alcoholic beverage is presented and respond when the image displays nonalcoholic beverages by pressing the spacebar key (*action*). In both training tasks, clients would need to respond with the correct action based on the image shown within a specific time limit to be successful (*principle*). The repeated actions of clients in both training tasks enable them to learn to inhibit their automatic responses toward alcohol and reduce their automatic tendency to approach alcohol, which in turn allow them to better respond toward alcohol-related cues in real life (Figure 10A).

XSL•FO
RenderX

Figure 10. The core-game loop and gamification design of Addiction Beater.



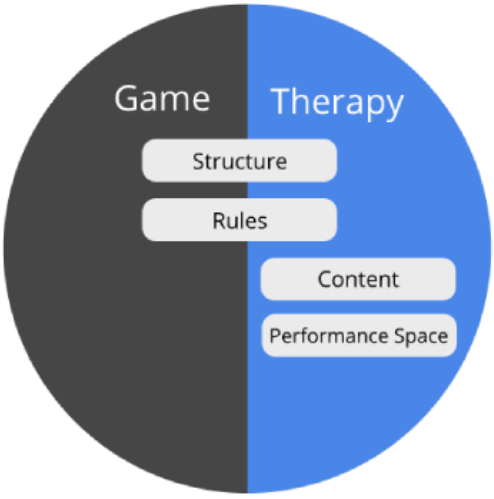
The gamification process for the Addiction Beater started with an initial analysis of the digital training modules. Initial brainstorming sessions were carried out to generate ideas on how the original *sample*, *action*, and *principle* of the cognitive training could be enhanced through gamification. Through a discussion with addiction researchers and clinical experts, different concepts were generated and refined into playable prototypes. In the concept that was later selected by the domain experts, the alcohol-related and nonalcohol-related images were presented to players based on the beat of the music. A beat detection algorithm was developed to ensure that the images would slide into the center of the screen based on the beat of the songs (*presentation*). Players would then need to respond in a timely and correct manner based on the rhythm of the music (similar to rhythm games where players tap keys or hit drums during different musical beats in the song), according to the original therapeutic principles (*game logic*). Players would receive points for each correct response, which was afterward upgraded into a "combo system" in which players would receive an increasing amount of bonus points for providing multiple correct response consecutively (+1 for the second correct response, +2 for the third correct response, etc.; *feedback*). Afterward, the feasibility of our music rhythm game concept was further examined in a preliminary evaluation session with heavy drinkers of different age groups. The results of the evaluation prompted us to add structural game elements, such as a level progression system in which players would earn experience points and level up for each training round completed. In addition, as the target audience of this gamification were people who were suffering from alcohol addiction, we avoided taking design choices (such as adding structural game play elements) that would result in our gamification sharing key characteristics found in addictive games (eg, 24-hour online games with an extensive in-game social network footprint, containing role-playing elements that strongly appeal to the sense of escapism) [52]. Furthermore, when implementing the gamification, we ensured that Addiction Beater was played under the supervision of the therapy staff, thus making it difficult for excessive gameplay.

The Therapy Game World of Addiction Beater

Early discussions carried out during the design of the CBM gamification revealed concerns about whether modifying the *performance space* and *content* of the original training modules might reduce the therapeutic effectiveness (Figure 11). Therefore, these 2 elements were kept constant in the gamification. The type and number of alcohol-related and nonalcohol-related images shown to the players (*content*), the actions players need to perform (pushing away alcohol-related images and pulling in nonalcohol-related images), and the feedback received (text feedback indicating they responded correctly or not; *performance space*) were kept the same as in the original training. While we kept most of the *principles* the same (eg, the time given to users to respond to each stimulus), we added an additional rule that players would need to time their response based on the beat of the music as a way to provide players with an experience of challenge and immersion. Finally, the *structure* of the therapy was modified in that the number of stimuli presented to the users in each training round was based on the length of the song played (and not based on a fixed amount as in the original therapy).

XSL•FO
RenderX

Figure 11. Therapy game world analysis of the Addiction Beater Gamification.

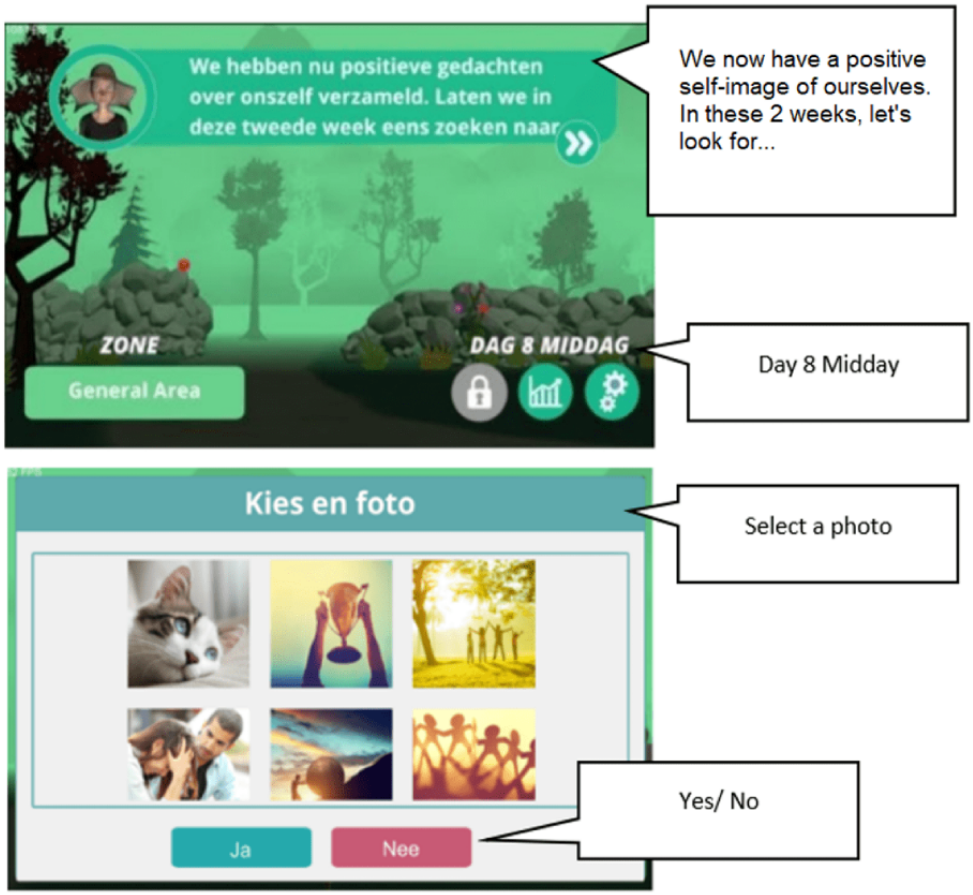


The Zen Garden App

Purpose

The Zen Garden is a gamification of the Competitive Memory Training (COMET) therapy used in the treatment of low self-esteem problems [53]. Low self-esteem has been found to be a factor in the development of a range of disorders, such as anxiety, depression, and eating disorders [54], leading also to self-harming and suicidal behaviors. The COMET therapy was developed to teach clients to retrieve positive and functional self-referent instead of negative information in situations that normally trigger negative thoughts and emotions linked to low self-esteem. The current Zen Garden was developed based on this therapy as a mobile app and adopts the aspects of playful interaction and progression. In this gamification, users are encouraged to plant, grow, and collect positive self-referent resources of themselves (eg, stories, photos, songs) and revisit the plants to strengthen memories of their positive qualities (Figure 12).

Figure 12. Screenshots of the Zen Garden.



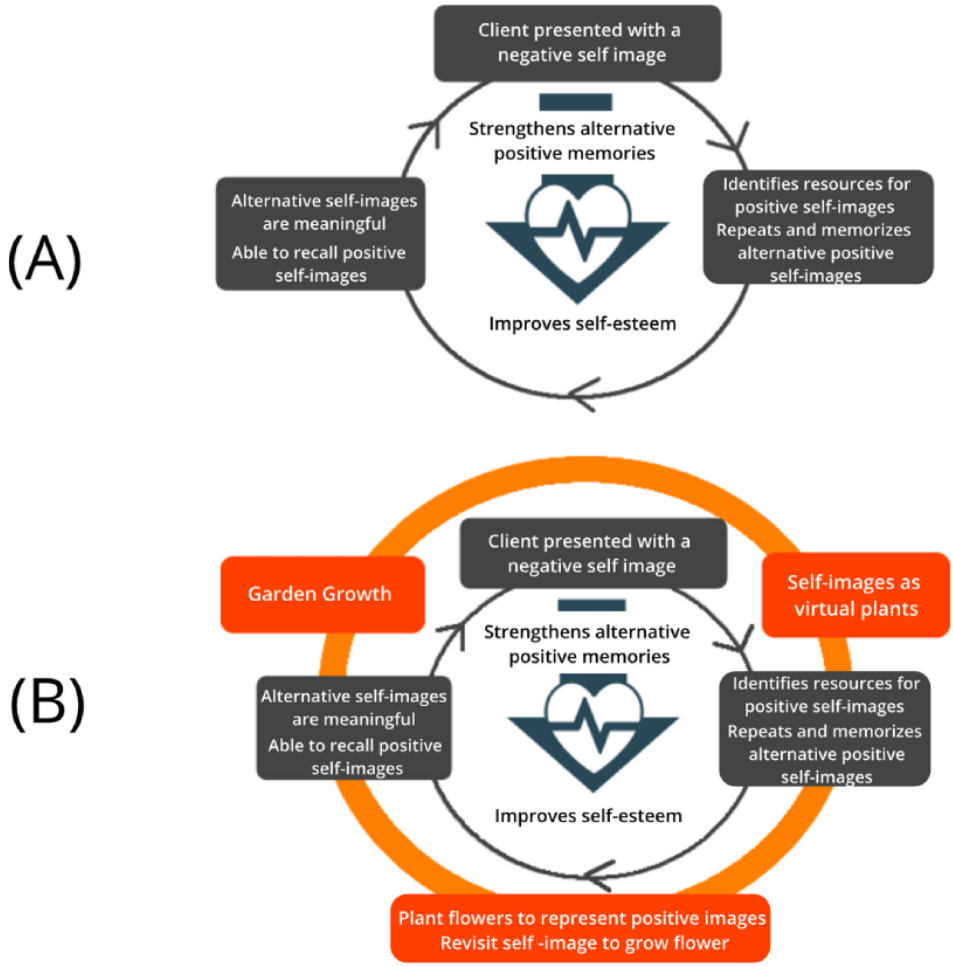
XSL•FO
RenderX

The Gamification Strategy and Core-Game Loop Design of the Zen Garden

In the original therapy, clients are asked to adjust dysfunctional negative self-views that impact their self-esteem (eg, clients thinking "I am worthless"; *sample*). They first identify positive personal characteristics that are inconsistent with their negative self-images (being honest, helpful, etc.) and describe examples of actions or situations in their daily life that are illustrative for those characteristics (eg, I helped a friend yesterday). These new positive characteristics of themselves are made more emotionally salient by imagining them repeatedly. In a step-by-step fashion, these positive images are further supported by adding a self-confident body posture and facial expression, positive self-statements, and self-selected positive music (*action*). Overall, the COMET training is congruent with Brewin's [55] competitive memory retrieval account of the working mechanism of cognitive behavioral therapy [55]. By making functional positive self-opinions more emotionally salient, clients will be able to better retrieve them from long-term memory, overruling the retrievability of the original dysfunctional negative self-opinions (*principles*), which would result in the improvement of self-esteem.

During the gamification of the COMET therapy, early game concept testing with a group of clients who completed the original therapy revealed that they tended to react negatively to game experiences related to challenge and competence, as failure in game tasks tended to reinforce the negative self-belief about themselves instead of motivating them to retry and complete the task. Therefore, in the gamification of the COMET therapy, we focused on more passive playful experiences, such as playful interaction and narration. The rationale for this design was to provide players with a sense of achievement without evoking fear of failure. We eventually adopted the concept of a virtual garden, deployed on a mobile device, where counteracting positive characteristics are represented as virtual plants in a 3D garden. The visual style of the garden was based on a peaceful Zen garden concept, with nature sounds such as that of water flowing incorporated into the garden to provide players with a feeling of serenity to help reduce anxiety (*presentation*). The garden is divided into different areas, each representing a negative self-belief, and users plant and grow flowers in each area by adding positive information about themselves. Text-based stories, images, and music representing real-life examples of those positive beliefs are added to each plant to help them grow (*game logic*). As users add more positive beliefs, the garden becomes more visually beautiful (*feedback*). Figure 13 summarizes the core-game loop and gamification design of the COMET application.

Figure 13. The core-game loop and gamification design of Comet Garden.



XSL•FO
RenderX

The Therapy Game World of the Zen Garden App

The gamification process of the COMET therapy relies on elements of playful interaction and the feeling of achievement from collecting and growing positive memories. The *rules*, *content*, and *structure* of the original therapy remain largely unchanged (Figure 14). Similar to the original therapy, players are asked to identify negative self-beliefs and then come up with alternative positive representations about themselves. Afterward, they add evidence for such new self-beliefs in their daily life and reflect on such evidence to strengthen their alternative positive self-beliefs (*rules and content*). The task given to participants each week was consistent with the original therapy (*structure*). More specifically, when used as part of a blended therapy activity, players would collaborate with the therapist during the first week to identify negative self-images, plant their first flower (representing a contradicting positive self-image/personal characteristics), and add self-referencing stories. In the following 4 weeks, players would add more stories (week 2), photos (week 3), and songs (week 4), which represent real-life examples that contradict their negative self-belief and review them daily through the Zen Garden app (similar assignments are given to clients during each week in the original therapy). Afterward, clients would share their garden with the therapist who would provide feedback about the evidence added and later clients would further review their contradicting positive self-images to strengthen them (week 5). However, due to the garden metaphor and technological limitations, the amount of resources that could be added for each flower was fixed (up to a maximum of 5), limiting the number of positive self-beliefs players could add, thus impacting the *performance space* of the therapy.

Figure 14. Therapy game world analysis of the Zen Garden gamification.

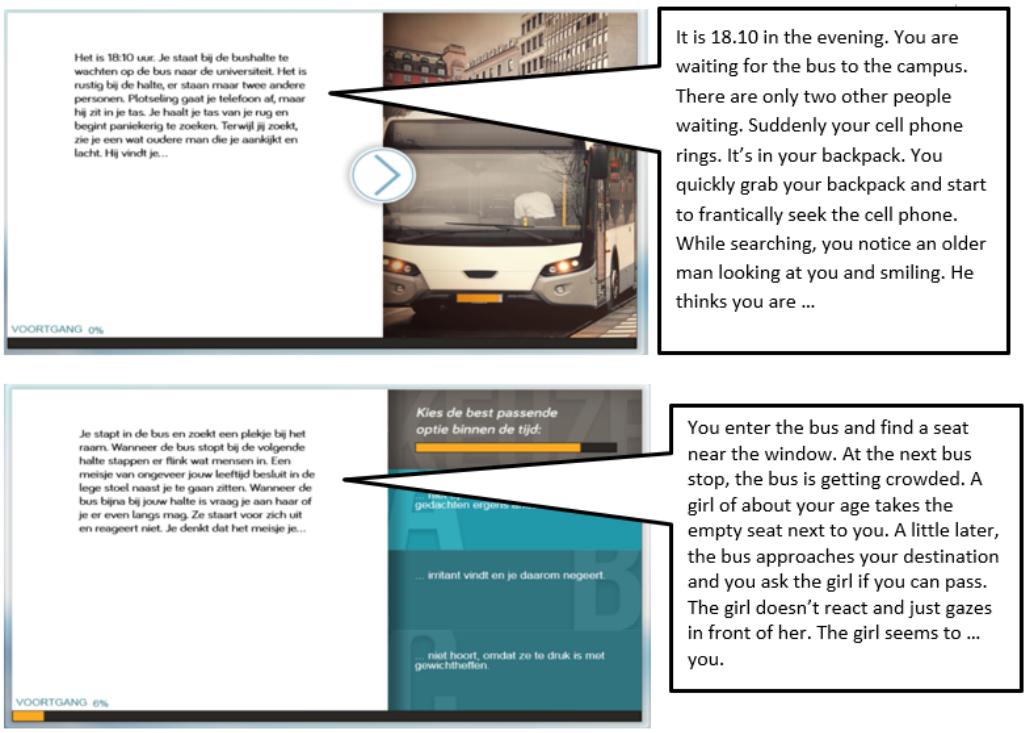
Één Klein Probleempje (A Tiny Problem): Social Anxiety Gamification

Purpose

The Één klein probleempje app is a narrative-based gamification of the cognitive bias modification of interpretations (CBM-I) training to address anxiety problems [56]. People with anxiety problems are generally inclined to interpret ambiguous information in a negative and threatening way, which in turn reinforces and exacerbates their anxiety symptoms. Interpretation bias training interventions have been designed to change this maladaptive information processing mechanism by training individuals to make more benign interpretations of ambiguous situations. However, this type of intervention is generally perceived as repetitive and tedious as clients are requested to complete multiple training sessions [57]. The Één klein probleempje was developed as a gamification of the CBM-I training by adopting the element of narration to enhance training. In this gamification, the fragmented scenarios in the original training are transformed into an engaging and connected story. Branching sections were added to the story which allowed users to decide on how the story unfolds. As users navigate through the story, they are encouraged to interpret the various ambiguous situations in a more positive way. Figure 15 shows a screenshot of the gamification.

XSL•FO
RenderX

Figure 15. Screenshots of the "A Tiny Problem" gamification.

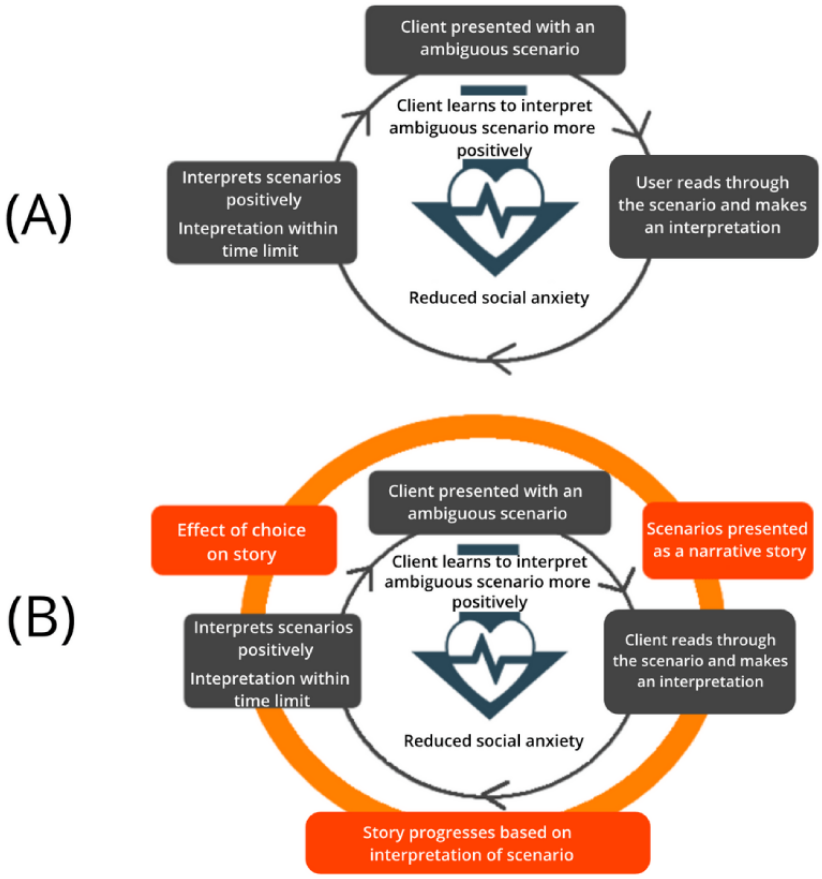


The Gamification Strategy and Core-Game Loop Design of the Één Klein Probleempje

In the original CBM-I training, clients are first presented with an ambiguous scenario ending with a word fragment, which can be solved with either a positive or negative outcome. Clients need to read the text and resolve the word fragment in a meaningful fashion (*action*). An example scenario is the following: *You've finished writing the answer to the second question in your exam. You take a small break, looking at what's left. You then realize that the questions left are more difficult than you had anticipated. Checking the watch, you decide you've planned your time w_ll [well]*. A subsequent question relating to the interpretation (eg, *Will you have time to complete the exam?*) is then presented, and users need to provide an answer (*yes* or *no*; *Scenario*). To be successful, the client would need to resolve the word fragment and related scenario in a positive manner and within a time limit (*rules*) [58]. Repeated training helps clients learn to interpret ambiguous situations in daily life in a more benevolent manner, thus breaking the dysfunctional thinking patterns underlying anxiety (Figure 16A).

XSL•FO
RenderX

Figure 16. The core-game loop and gamification design of the small problems.



Because of the key role of text-based scenarios in the original training, we decided to use the element of narration to gamify the CBM-I training task. The different scenarios in the original therapy were tied together and presented to users as an adventure story (*presentation*). The aim was to engross users through the narration of the story, as they would be motivated to complete the training task and resolve ambiguous situations to find out how the story ends. To progress along the story, users need to provide a positive interpretation of each story scenario by selecting the correct option from a multiple choice list presented in a timely manner (*game logic*). A negative interpretation will lead users back to the same choice point and users would need to make the correct choice (ie, the positive interpretation) to continue with the story (*feedback*; Figure 16B).

The Therapy Game World of Één Klein Probleempje

In Één klein probleempje, the scenarios from the original CBM-I training were converted into a storyline, thus changing the original *content* of the training. In addition, the *structure* of the original training was modified (Figure 17). Instead of each scenario being presented in a random, unrelated order, they are presented in a logical order based on how the story unfolds. Through discussions with cognitive training experts, we decided that the content and structure of the scenarios were flexible enough to be adapted without severely impacting the effectiveness of the targeted training. In addition, instead of filling in a word fragment to complete the scenario, users would choose how the scenario unfolds by selecting 1 of 4 optional outcomes. This was done to provide a more natural user experience as users go through the story, thus impacting the *performance space* of the original therapy. The overall *rules* of the therapy were kept the same, as users are still forced to choose a positive outcome for the scenario to succeed in the therapy.

Table 2 provides a summary of the therapy loop of the 4 gamification case studies and Table 3 provides a summary of the overall gamification design of the 4 case studies.

XSL•FO
RenderX

Figure 17. Therapy game world analysis of the Small Problems gamification.

Table 2. A summary of the therapy loop of the four gamification case studies.

Gamification	Existing therapeutic activity	Challenge of existing activity	Therapeutic activity loop			Desired specific behavioral or cognitive change	Desired overall well-being objective
			Sample	Action	Principles		
ReadySet-Goals	Goal setting within cognitive behavioral therapy.	It is difficult to motivate clients to set and accomplish goals related to cognitive behavioral therapy.	Users presented with a task beneficial to their therapy to carry out in daily practice.	Users carry out the actions required to complete the tasks.	The task must be completed successfully and within the set time limit.	*Behavioral change:* Users are able to set and accomplish goals that are beneficial to their therapy.	Accomplishing goals (abstinence goals, skill attainment goals, etc.) leads to better therapeutic outcome.
Addiction Beater	Cognitive bias modification training used in alcohol addiction treatment.	High dropout rate. Boredom due to repetitiveness, particularly with youth in prevention studies.	Users are presented with target stimuli (alcohol related or nonalcohol related).	Users decide to "push"/"pull" or inhibit their response toward a target stimulus which is explicitly alcohol related or associated with alcohol.	Users must respond in time and accurately.	*Cognitive change:* Users learn to avoid and inhibit automatic responses toward alcohol.	Lower relapse rate or reduced alcohol use.
Zen Garden	Competitive Memory Training used in the treatment of low self-esteem.	The e-training version has problems related to low retention.	Users identify a negative self-image.	Users identify alternative positive self-images and rehearses them.	Users must be able to recall positive self-images.	*Cognitive change:* Strengthens positive self-images as opposed to negative ones.	Improvements in self-esteem.
Één klein probleempje	Cognitive bias modification of interpretations training for anxiety.	Boredom during training, repetitiveness, and fragmentation of scenarios.	Users presented with an ambiguous scenario.	Users read through the scenario and fill in a word fragment, which can end positively or negatively.	Users interpret the scenario positively and within the time limit.	*Cognitive change:* Learns to interpret ambiguous information in a benevolent manner.	Reduce social anxiety.

XSL•FO
RenderX

Table 3. A summary of the overall gamification design of the 4 gamification case studies.

Gamification	Game concept	Core-game loop elements			Structural game elements
		Presentation design	Game logic design	Feedback design	
ReadySetGoals	A risk-taking concept where users place wagers on their set goals.	Goal setting is presented through the metaphor of climbing a mountain.	Players receive points proportionate to the risk taken (the amount of wager placed and the set time limit) and the difficulty of the task. Players progress further up the mountain for each accomplished goal.	Successful task completion and risk taking are rewarded with points. Pictures of accomplished tasks are placed on the mountain path to provide a sense of achievement.	Player progression (the number of points accumulated increases player levels); gameplay progression (easy goals are placed at the bottom of the mountain and more difficult goals are placed at the top).
Addiction Beater	A music rhythm game concept where users must respond based on the beat of the music.	Alcohol- and nonalcohol-related stimuli are presented based on the beat rhythm of the music.	Users have to react to the stimuli (to the content or other features) as close as possible on the beat of the music.	Users receive points for each correct response and more points for consecutive correct responses (combos).	Player progression (players accumulate experience points for each correct response and level up); gameplay progression (players unlock more difficult songs after completing easier ones); social competition (players see their performance [response time, accuracy] in relation to others).
Zen Garden	A playful Zen garden where users can plant positive memories into flowers.	Players' "self" is represented through a garden metaphor: negative self-beliefs are represented with zones within the garden and positive self-images are symbolized through flowers planted in the garden.	Users need to plant flowers to grow their garden and their positive self-images.	The more flowers are planted, the more the garden grows.	Narrative guidance (an avatar is used to guide the participants throughout the therapy and give them tasks).
Één klein probleempje	An adventure story–based gamification.	Each ambiguous scenario is presented as an element of a larger narrative story.	Users interpret the ambiguous scenarios and the narrative story responds to the players' interpretations.	The results of the player interpretations are shown through narrative feedback.	Story progression (the story progresses as players interpret the ambiguous scenarios).

Discussion

Overall, the 4 case studies were presented to help illustrate how our proposed Dual-Loop model and the game therapy world concept could be put into practice when designing gamification for mental health care. In each of the gamifications, different design elements in the core-game loop (presentation, game logic, etc.) were selected to correspond with the inherent characteristics of the components in therapeutic activity as analyzed through the Dual-Game Loop model. The therapy game world was then used to analyze the impact of the selected game design elements on the underlining therapeutic activity. The key reflections and observations from our use of the framework to design the gamification in the case studies are highlighted in Textbox 2.

XSL•FO
RenderX

Textbox 2. Key reflections and observations.

Presentation design

In gamifications designed based on computerized cognitive training therapies, such as Addiction Beater and Één klein probleempje, the presentation was generally designed based on the limitations and characteristic of the stimulus used in the original computerized training as they tended to be predetermined. For instance, in Addiction Beater, it was unclear whether changing the visual size or color in which the alcoholic/nonalcoholic images were presented would impact the therapy effectiveness [43] and thus the presentation was enhanced mainly through audio. In Één klein probleempje, a narrative element fitted well as this allowed the text-based scenarios to be presented in an engaging manner without severely effecting the original content of the scenarios used for the cognitive bias modification training. Such considerations to presentation design are reflected in cognitive training gamifications developed in previous studies as well [22,46]. For cognitive behavioral therapies that were deployed as blended therapies (The ReadySetGoals and the Zen Garden), however, the presentation tended to be more abstract and dependent on the user (ie, goals to set or positive self-images). In such cases, a metaphor system was used to highlight the achievements of the user in the therapeutic activity (ie, presenting completed goals as steps on a mountain and positive self-beliefs that users added as gardens), acting as a form of virtual trophy, a commonly used gamification strategy to reinforce progress or good behavior from users [59].

Game logic design

In the Addiction Beater and Één klein probleempje gamifications, the therapy principles generally contained user interaction rules which were fixed (ie, respond to the stimulus within 500 ms) [43,44] and it is unknown whether modifying them would have an impact on the therapeutic effectiveness. In such cases, similar to the presentation design, rather than risk changing the original principles in the therapy, we added a simple game play mechanism upon the original user action in the therapy to enhance the gamified user experience (ie, while users still had to respond within 500 ms to the stimulus, this was matched to the beat of the music in Addiction Beater to create challenge). For cognitive behavioral therapies deployed in a blended format, completely new game logic and interaction mechanism were devised (the wager mechanism for the ReadySetGoals, etc.). The rules involved in these therapies were often more subjective and flexible (ie, in the goal setting activity, users set their own deadline and the success and failure of the goal are decided in collaboration with the therapist), giving more freedom in the game logic design. In a sense, the game logic design constraints and strategies employed for these gamifications are similar to those found on other gamifications where user interaction is primarily driven by the behavior of users in the real world (such as those designed to encourage behavioral change [60] or facilitate the self-management of a disease [61]).

Feedback

The feedback was generally designed as a way to quantify the achievements of the users within the therapeutic activities, with points or experience points being awarded for performance in ReadySetGoals and Addiction Beater. This numerical quantification was necessary to allow challenge-based structural game element, such as competition or player progression, to be employed (which coincided well with gamification strategies designed to treat explicit mental health disorders such as addiction [62]). For gamifications designed to treat implicit disorders, such as Één klein probleempje and Zen Garden, we generally used a more abstract representation of player achievement, such as the growth of flowers in the garden or narrative progression in the story, as our design strategy in these gamifications had less emphasis on challenge.

Structural game elements

The structural game elements used in the design tended to be strongly influenced by the nature of disorder that the therapy seeks to address. Overall, the element of *progression* was commonly used as the structural game element in most of the gamifications which were cited as case studies. We felt that this element was appropriate as it helped emphasize the achievements and success of players in their therapeutic activity, replicating a commonly used strategy of highlighting past successes [60]. The *game leveling system* in ReadySetGoals and Addiction Beater was designed to enhance the experience of challenge which we felt appealed to people having substance abuse problems, who tended to favor sensation-seeking experiences, as such users also tended to be sensitive toward rewards (see [10]). The *player leveling system* was also used to increase the feeling of achievement from their successes in the activities. In Zen Garden, however, we felt that progression could add pressure for users to perform, especially for people with low self-esteem who were the main target audience of this gamification. Therefore, no performance-based progression element was used in this gamification.

To our knowledge this is the first paper in which the demands, restrictions, and possibilities of interdisciplinary health/game research are combined in a framework aimed to help designers integrate game design elements with therapeutic content in the mental health care context. Previous practice-based gamification approaches that have been proposed in the literature tend not to focus on a specific domain [63,64], or are designed for use in other areas such as business and education [65]. However, our method was developed specifically to support the design process of mental health care gamification, with the key processes used in a therapeutic activity forming the foundation of the concepts and models proposed in our framework. The Dual-Loop model serves as a practical tool to help designers deconstruct the processes used within a therapeutic activity and allow them to best determine the appropriate design strategy to create engaging experiences around each therapeutic component. The game therapy world concept provides a basic way for designers to analyze the potential impact of their gamification design on the therapeutic activity, an essential process which can be difficult to carry out in mental health gamification [21]. Moreover, our proposed method highlights how the game and therapeutic elements could be balanced in the development of a gamified intervention.

Conclusion

In this paper, we propose a framework for the gamification of mental health care therapies. First, we analyzed existing mental health care gamifications and propose the concept of the game therapy world to illustrate how game elements could be integrated into therapeutic activities. We highlighted 4 different components of a game therapy world: the *performance space*, *rules*, *content*, and *structure*, which form a key part of a therapeutic activity that can be enhanced using game elements.

In addition, to aid developers in the process of designing the gamifications and allow them to better analyze the effect of their design choice on the 4 aforementioned components, we proposed the Dual-Loop Design model. This model consists of

XSL•FO
RenderX

a core-game loop encompassing a therapeutic activity loop. The role of the designer is to consider how "presentation design," "game logic design," and "feedback design" could be used to enhance the "sample", "user action," and "therapy principle" components in the therapeutic activity loop. Further *structural game elements* such as player progression could then be added to encourage sustained engagement across the loops.

To illustrate how our framework could be used in practice, we provided 4 case studies of gamifications we developed and analyzed their design process through our proposed framework.

It should be noted, however, that there are several limitations to the framework proposed in this study. First, the framework itself was developed based on a limited class of intervention types and mental disorders. Whether the framework would be generalizable to other types of mental health intervention activities (eg, preventative health care or posttreatment activities) or psychological disorders remains unclear. In addition, the framework aims to support the design process of mental health gamifications formed from integrated therapy game worlds and not those formed from separated game and therapy worlds (such as in the "Integrating Game Elements Into Therapeutic Activity Through a Game Therapy Worlds Concept" section). Finally, the framework is conceptual in nature and specifying the tools and methods (prototype testing approaches, etc.) used in the design and development of the gamification elements (presentation, game logic, etc.) is beyond the scope of our framework.

In our future work, we would further investigate the general applicability of our models and framework toward gamifications in other domains, starting from the more general areas of health care (eg, gamification in preventative and posttreatment care). In addition, we would examine the tools and methods which could be used to support the design of gamifications based on our proposed Dual-Loop model and evaluate their effectiveness. In particular, we would investigate the various formative and summative evaluation approaches which could be effective in evaluating the gamification prototypes generated during different stages of design and conceptualization.

Conflicts of Interest

None declared.

References

1. 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]
2. Hervas R, Ruiz-Carrasco D, Mondejar T, Bravo J. Gamification mechanics for behavioral change: a systematic review and proposed taxonomy. 2017 May Presented at: 11th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 23, 2017; Mount Sinai, NY p. a-404. [doi: 10.1145/3154862.3154939]
3. Theng Y, Lee JW, Patinadan PV, Foo SS. The Use of Videogames, Gamification, and Virtual Environments in the Self-Management of Diabetes: A Systematic Review of Evidence. Games Health J 2015 Oct;4(5):352-361. [doi: 10.1089/g4h.2014.0114] [Medline: 26287926]
4. Beale IL, Kato PM, Marin-Bowling VM, Guthrie N, Cole SW. Improvement in cancer-related knowledge following use of a psychoeducational video game for adolescents and young adults with cancer. J Adolesc Health 2007 Sep;41(3):263-270. [doi: 10.1016/j.jadohealth.2007.04.006] [Medline: 17707296]
5. Brown M, O'Neill N, van Woerden H, Eslambolchilar P, Jones M, John A. Gamification and Adherence to Web-Based Mental Health Interventions: A Systematic Review. JMIR Ment Health 2016 Aug 24;3(3):e39 [FREE Full text] [doi: 10.2196/mental.5710] [Medline: 27558893]
6. Richards D, Caldwell PH. Handbook of Research on Holistic Perspectives in Gamification for Clinical Practice. Hershey, PA: IGI Global; 2016:80-111.
7. Kazantzis N, Whittington C, Dattilio F. Meta-analysis of homework effects in cognitive and behavioral therapy: A replication and extension. Clinical Psychology: Science and Practice 2010 Jun;17(2):144-156. [doi: 10.1111/j.1468-2850.2010.01204.x]
8. Kazantzis N, Shinkfield G. Conceptualizing Patient Barriers to Nonadherence with Homework Assignments. Cognitive and Behavioral Practice 2007 Aug;14(3):317-324. [doi: 10.1016/j.cbpra.2006.08.003]
9. Bul KC, Franken IH, Van der Oord S, Kato PM, Danckaerts M, Vreeke LJ, et al. Development and User Satisfaction of "Plan-It Commander," a Serious Game for Children with ADHD. Games Health J 2015 Dec;4(6):502-512. [doi: 10.1089/g4h.2015.0021] [Medline: 26325247]
10. van Dooren MM, Siriaraya P, Visch V, Spijkerman R, Bijkerk L. Reflections on the design, implementation, and adoption of a gamified eHealth application in youth mental healthcare. Entertainment Computing 2019 Aug;31:100305. [doi: 10.1016/j.entcom.2019.100305]
11. Nicholson S. A recipe for meaningful gamification. In: Gamification in Education and Business. Cham, Switzerland: Springer; 2015:1-20. [doi: 10.1007/978-3-319-10208-5_1]
12. 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]

XSL•FO
RenderX

13. Boendermaker WJ, Sanchez Maceiras S, Boffo M, Wiers RW. Attentional Bias Modification With Serious Game Elements: Evaluating the Shots Game. JMIR Serious Games 2016 Dec 06;4(2):e20 [FREE Full text] [doi: 10.2196/games.6464] [Medline: 27923780]
14. 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]
15. Brezinka V. Treasure Hunt - a serious game to support psychotherapeutic treatment of children. Stud Health Technol Inform 2008;136:71-76. [Medline: 18487710]
16. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MFG. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012 Apr 18;344(apr18 3):e2598-e2598 [FREE Full text] [doi: 10.1136/bmj.e2598] [Medline: 22517917]
17. Purgato A, Gamberini L. Playsafety: Virtual environments as a persuasive tool to contrast risky behaviors in youth. In: Cyberpsychology & Behaviour. New Rochelle, NY: Mary Ann Liebert; Aug 1, 2005:320-321.
18. Wouters P, van der Spek ED, van Oostendorp H. Current practices in serious game research: A review from a learning outcomes perspective. In: Games-Based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices. Hershey, PA: IGI Global; 2009:232-250.
19. Cheng VWS, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in Apps and Technologies for Improving Mental Health and Well-Being: Systematic Review. JMIR Ment Health 2019 Jun 26;6(6):e13717 [FREE Full text] [doi: 10.2196/13717] [Medline: 31244479]
20. Visch V, Vegt H, Anderlessen H, van der Kooij K. Persuasive game design: a model and its definitions. New York, NY: ACM; 2013 Presented at: CHI 2013: Workshop Designing Gamification; Apr 27-May 2, 2013; Paris, France URL: https://www.researchgate.net/publication/263561793_Persuasive_Game_Design_A_model_and_its_definitions
21. van der Kooij K, Hoogendoorn E, Spijkerman R, Visch V. Validation of Serious Games. IJSG 2015 Sep 18;2(3):63-75. [doi: 10.17083/ijsg.v2i3.75]
22. Zielhorst T, van den Brule D, Visch V, Melles M, van Tienhoven S, Sinkbaek H, et al. Using a digital game for training desirable behavior in cognitive-behavioral therapy of burnout syndrome: a controlled study. Cyberpsychol Behav Soc Netw 2015 Feb;18(2):101-111. [doi: 10.1089/cyber.2013.0690] [Medline: 25684611]
23. Lewis ZH, Swartz MC, Lyons EJ. What's the Point?: A Review of Reward Systems Implemented in Gamification Interventions. Games Health J 2016 Apr;5(2):93-99. [doi: 10.1089/g4h.2015.0078] [Medline: 26812253]
24. Fizek S. Why fun matters: in search of emergent playful experiences. In: Rethinking Gamification. Lüneburg, Germany: Meson Press; 2014:273-287.
25. Ferrara J. Games for Persuasion. Games and Culture 2013 Aug 27;8(4):289-304. [doi: 10.1177/1555412013496891]
26. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining? gamification? : ACM; 2011 Presented at: The 15th International Academic MindTrek Conference: Envisioning Future Media Environments; 2011 Sep 28; Tampere Finland. [doi: 10.1145/2181037.2181040]
27. Huizinga J. Homo Ludens: A Study of the Play-Element in Culture. Oxfordshire, UK: Routledge; 2014.
28. Green MC, Brock TC. The role of transportation in the persuasiveness of public narratives. Journal of Personality and Social Psychology 2000 Nov;79(5):701-721. [doi: 10.1037/0022-3514.79.5.701]
29. van der Kooij K, Hoogendoorn E, Spijkerman R, Visch V. Changamoto: Design and Validation of a Therapy Adherence Game. In: Games for Health 2014. Heidelberg, Germany: Springer; 2014:144-149.
30. Hunicke R, LeBlanc M, Zubek R. MDA: A formal approach to game design and game research. 2004 Presented at: 2004 AAAI Workshop on Challenges in Game AI; July 25, 2004; San Jose, CA p. 1722-1726.
31. Duke RD. A Paradigm for Game Design. Simulation & Games 2016 Aug 04;11(3):364-377. [doi: 10.1177/104687818001100308]
32. Schell J. The Art of Game Design: A book of lenses. Boca Raton, FL: CRC press; 2008.
33. Klinger E, Bouchard S, Légeron P, Roy S, Lauer F, Chemin I, et al. Virtual reality therapy versus cognitive behavior therapy for social phobia: a preliminary controlled study. Cyberpsychol Behav 2005 Feb;8(1):76-88. [doi: 10.1089/cpb.2005.8.76] [Medline: 15738695]
34. Gamberini L, Breda L, Grassi A. VIDEODOPE: applying persuasive technology to improve awareness of drugs abuse effects. 2007 Presented at: 2007 International Conference on Virtual Reality; July 22, 2007; Beijing, China p. 633-641. [doi: 10.1007/978-3-540-73335-5_68]
35. Klisch Y, Bowling KG, Miller LM, Ramos MA. The Impact of Science Education Games on Prescription Drug Abuse Attitudes among Teens: A Case Study. J Drug Educ 2014 Aug 01;43(3):255-275. [doi: 10.2190/de.43.3.d]
36. Hamari J, Koivisto J, Sarsa H. Does gamification work? A literature review of empirical studies on gamification. 2014 Presented at: 47th Hawaii International Conference on System Sciences; January 6, 2014; Waikoloa, HI p. 3025-3034. [doi: 10.1109/hicss.2014.377]
37. Blagys MD, Hilsenroth MJ. Distinctive activities of cognitive–behavioral therapy. Clinical Psychology Review 2002 Jun;22(5):671-706. [doi: 10.1016/s0272-7358(01)00117-9]

XSL•FO
RenderX

38. Kazantzis N, Ronan KR. Can between-session (homework) activities be considered a common factor in psychotherapy? Journal of Psychotherapy Integration 2006;16(2):115-127. [doi: 10.1037/1053-0479.16.2.115]
39. Clark DM, Ehlers A, Hackmann A, McManus F, Fennell M, Grey N, et al. Cognitive therapy versus exposure and applied relaxation in social phobia: A randomized controlled trial. Journal of Consulting and Clinical Psychology 2006 Jun;74(3):568-578. [doi: 10.1037/0022-006x.74.3.568]
40. Beard C, Amir N. A multi-session interpretation modification program: changes in interpretation and social anxiety symptoms. Behav Res Ther 2008 Oct;46(10):1135-1141 [FREE Full text] [doi: 10.1016/j.brat.2008.05.012] [Medline: 18675400]
41. Wiers RW, Gladwin TE, Hofmann W, Salemink E, Ridderinkhof KR. Cognitive Bias Modification and Cognitive Control Training in Addiction and Related Psychopathology. Clinical Psychological Science 2013 Jan 09;1(2):192-212. [doi: 10.1177/2167702612466547]
42. Smith JE, Meyers RJ, Miller WR. The Community Reinforcement Approach to the Treatment of Substance Use Disorders. American Journal on Addictions 2001 Jan 1;10:51-59. [doi: 10.1080/10550490190942515]
43. Houben K, Havermans R, Nederkoorn C, Jansen A. Beer à no-go: learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. Addiction 2012 Jul;107(7):1280-1287. [doi: 10.1111/j.1360-0443.2012.03827.x] [Medline: 22296168]
44. Salemink E, Kindt M, Rienties H, van den Hout M. Internet-based cognitive bias modification of interpretations in patients with anxiety disorders: a randomised controlled trial. J Behav Ther Exp Psychiatry 2014 Mar;45(1):186-195. [doi: 10.1016/j.jbtep.2013.10.005] [Medline: 24177146]
45. Wu Y, Kankanhalli A, Huang K. Gamification in fitness apps: how do leaderboards influence exercise? 2015 Presented at: Thirty Sixth International Conference on Information Systems; December 13-16, 2015; Fort Worth, TX.
46. Boendermaker WJ, Boffo M, Wiers RW. Exploring Elements of Fun to Motivate Youth to Do Cognitive Bias Modification. Games for Health Journal 2015 Dec;4(6):434-443. [doi: 10.1089/g4h.2015.0053] [Medline: 26421349]
47. Francillette Y, Gouaich A, Hocine N, Pons J. A gameplay loops formal language. 2012 Presented at: 2012 17th International Conference on Computer Games (CGAMES); July 30 to August 1, 2012; Louisville, KY p. 94-101. [doi: 10.1109/cgames.2012.6314558]
48. Lieberoth A, Møller M, Marin A. Deep and Shallow Gamification in Marketing: Thin Evidence and the forgotten Powers of Really Good Games. In: Engaging Consumers Through Branded Entertainment and Convergent Media. Hershey, PA: IGI Global; 2015:110-126. [doi: 10.4018/978-1-4666-8342-6.ch006]
49. Belcher AM, Volkow ND, Moeller FG, Ferré S. Personality traits and vulnerability or resilience to substance use disorders. Trends Cogn Sci 2014 Apr;18(4):211-217 [FREE Full text] [doi: 10.1016/j.tics.2014.01.010] [Medline: 24612993]
50. Wiers RW, Eberl C, Rinck M, Becker ES, Lindenmeyer J. Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. Psychol Sci 2011 Apr 09;22(4):490-497 [FREE Full text] [doi: 10.1177/0956797611400615] [Medline: 21389338]
51. Boffo M, Zerhouni O, Gronau QF, van Beek RJJ, Nikolaou K, Marsman M, et al. Cognitive Bias Modification for Behavior Change in Alcohol and Smoking Addiction: Bayesian Meta-Analysis of Individual Participant Data. Neuropsychol Rev 2019 Mar 14;29(1):52-78 [FREE Full text] [doi: 10.1007/s11065-018-9386-4] [Medline: 30644025]
52. Lemmens JS, Hendriks SJ. Addictive Online Games: Examining the Relationship Between Game Genres and Internet Gaming Disorder. Cyberpsychol Behav Soc Netw 2016 Apr;19(4):270-276. [doi: 10.1089/cyber.2015.0415] [Medline: 26959285]
53. Korrelboom K, Marissen M, van Assendelft T. Competitive Memory Training (COMET) for Low Self-Esteem in Patients with Personality Disorders: A Randomized Effectiveness Study. Behav. Cogn. Psychother 2010 Sep 01;39(1):1-19. [doi: 10.1017/s1352465810000469]
54. Sowislo JF, Orth U. Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. Psychol Bull 2013 Jan;139(1):213-240. [doi: 10.1037/a0028931] [Medline: 22730921]
55. Brewin CR. Understanding cognitive behaviour therapy: A retrieval competition account. Behav Res Ther 2006 Jun;44(6):765-784. [doi: 10.1016/j.brat.2006.02.005] [Medline: 16620779]
56. Salemink E, van den Hout M, Kindt M. Effects of positive interpretive bias modification in highly anxious individuals. J Anxiety Disord 2009 Jun;23(5):676-683. [doi: 10.1016/j.janxdis.2009.02.006] [Medline: 19272750]
57. Beard C, Weisberg RB, Primack J. Socially Anxious Primary Care Patients' Attitudes Toward Cognitive Bias Modification (CBM): A Qualitative Study. Behav. Cogn. Psychother 2011 Nov 30;40(5):618-633. [doi: 10.1017/s1352465811000671]
58. Salemink E, van den Hout M, Kindt M. Trained interpretive bias: validity and effects on anxiety. J Behav Ther Exp Psychiatry 2007 Jun;38(2):212-224. [doi: 10.1016/j.jbtep.2006.10.010] [Medline: 17141735]
59. Miller AS, Cafazzo JA, Seto E. A game plan: Gamification design principles in mHealth applications for chronic disease management. Health Informatics J 2016 Jun 01;22(2):184-193 [FREE Full text] [doi: 10.1177/1460458214537511] [Medline: 24986104]
60. 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]

XSL•FO
RenderX

61. Priesterroth L, Grammes J, Holtz K, Reinwarth A, Kubiak T. Gamification and Behavior Change Techniques in Diabetes Self-Management Apps. J Diabetes Sci Technol 2019 Sep 14;13(5):954-958 [FREE Full text] [doi: 10.1177/1932296818822998] [Medline: 30762420]
62. van Dooren M, Spijkerman R, Goossens R, Hendriks V, Visch V. PLEX as input and evaluation tool in persuasive game design: pilot study. 2014 Presented at: The first ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play; October 19, 2014; New York, Ny p. 449-450. [doi: 10.1145/2658537.2661314]
63. Morschheuser B, Werder K, Hamari J, Abe J. How to gamify? Development of a method for gamification. 2017 Presented at: The 50th Annual Hawaii International Conference on System Sciences (HICSS); January 4, 2017; Hilton Waikoloa Village, HI p. 4-7. [doi: 10.24251/hicss.2017.155]
64. Marache-Francisco C, Brangier E. From The Consideration of Gamification To Its Practical Implementation. 2013 Presented at: The Sixth International Conference on Advances in Human-Oriented and Personalized Mechanisms, Technologies, and Services; October 27-November 1, 2013; Venice, Italy p. 126-131.
65. Urh M, Vukovic G, Jereb E, Pintar R. The Model for Introduction of Gamification into E-learning in Higher Education. Procedia - Social and Behavioral Sciences 2015 Jul;197:388-397. [doi: 10.1016/j.sbspro.2015.07.154]

Abbreviations

CBM: cognitive bias modification
CBM-I: cognitive bias modification of interpretations
COMET: Competitive Memory Training

Edited by N Zary; submitted 15.02.21; peer-reviewed by X Xu, R Nuijten, M Kapitány-Fövény, F Ghezelbash, F Ghezelbash; comments to author 16.04.21; revised version received 17.06.21; accepted 19.06.21; published 01.12.21.

Please cite as:
Siriaraya P, Visch V, Boffo M, Spijkerman R, Wiers R, Korrelboom K, Hendriks V, Salemink E, van Dooren M, Bas M, Goossens R
Game Design in Mental Health Care: Case Study–Based Framework for Integrating Game Design Into Therapeutic Content
JMIR Serious Games 2021;9(4):e27953
URL: https://games.jmir.org/2021/4/e27953
doi:10.2196/27953
PMID:34855611

©Panote Siriaraya, Valentijn Visch, Marilisa Boffo, Renske Spijkerman, Reinout Wiers, Kees Korrelboom, Vincent Hendriks, Elske Salemink, Marierose van Dooren, Michael Bas, Richard Goossens. Originally published in JMIR Serious Games (https://games.jmir.org), 01.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Serious Games Are Not Serious Enough for Myoelectric Prosthetics

Christian Alexander Garske[1], MSc; Matthew Dyson[1], PhD; Sigrid Dupan[2], PhD; Graham Morgan[3], PhD; Kianoush Nazarpour[2], PhD

[1]Intelligent Sensing Laboratory, School of Engineering, Newcastle University, Newcastle upon Tyne, United Kingdom
[2]Edinburgh Neuroprosthetics Laboratory, School of Informatics, University of Edinburgh, Edinburgh, United Kingdom
[3]Networked and Ubiquitous Systems Engineering Group, School of Computing, Newcastle University, Newcastle upon Tyne, United Kingdom

Corresponding Author:
Christian Alexander Garske, MSc
Intelligent Sensing Laboratory
School of Engineering
Newcastle University
Merz Court
Newcastle upon Tyne, NE1 7RU
United Kingdom
Phone: 44 191 20 86682
Email: c.a.garske2@ncl.ac.uk

Abstract

Serious games show a lot of potential for use in movement rehabilitation (eg, after a stroke, injury to the spinal cord, or limb loss). However, the nature of this research leads to diversity both in the background of the researchers and in the approaches of their investigation. Our close examination and categorization of virtual training software for upper limb prosthetic rehabilitation found that researchers typically followed one of two broad approaches: (1) focusing on the game design aspects to increase engagement and muscle training and (2) concentrating on an accurate representation of prosthetic training tasks, to induce task-specific skill transfer. Previous studies indicate muscle training alone does not lead to improved prosthetic control without a transfer-enabling task structure. However, the literature shows a recent surge in the number of game-based prosthetic training tools, which focus on engagement without heeding the importance of skill transfer. This influx appears to have been strongly influenced by the availability of both software and hardware, specifically the launch of a commercially available acquisition device and freely available high-profile game development engines. In this Viewpoint, we share our perspective on the current trends and progress of serious games for prosthetic training.

(JMIR Serious Games 2021;9(4):e28079) doi:10.2196/28079

KEYWORDS

rehabilitation; serious games; engagement; transfer; upper limb; arm prosthesis; virtual training; virtual games

Background

Adherence of patients to interventions (eg, home exercises) remains a key challenge in rehabilitation medicine [1]. Patients complain that exercises often feel tedious and tiring and that progress, if any, is slow and incremental [1]. Delivering virtual training in the form of games can help overcome issues related to nonadherence (or noncompliance) of patients to their exercise regimen [1]. The use of serious games has been recommended to motivate patients in performing their prescribed exercises consistently and completely [2-4].

The stroke rehabilitation literature includes a large number of publications that use serious games. Koutisiana et al [5] identified 96 publications between the years 1999 and 2019. The serious games used in stroke rehabilitation are showing significant benefits for the users, most notably an increased number of repetitions performed, which is a prime goal for this kind of rehabilitation [6]. Supported by this academic evidence, rehabilitation programs like Rehability (Imaginary srl), which has grown out of the Rehab@Home project [7], are being incorporated in clinical practice.

Although serious games have found their way into a multitude of areas of everyday life, industry, and research, including prosthetic training [8], academic results supporting the efficacy of serious games in myoelectric prosthetic training are scarce, if not nonexistent. Using games in virtual rehabilitation has

XSL•FO
RenderX

been a part of research for 30 years [9], but they have only gained proper traction in the field in the last decade. This rise coincided with the commercialization of a range of game-related technologies (eg, motion tracking cameras and game controllers with inertial measurement sensors [6]). In this paper, we will offer our perspective on the efficacy of virtual training in general and serious games specifically for myoelectric prosthetics training.

Current research claims that the use of serious games in myoelectric prosthetics training has promise to improve training. Examples include faster learning [10], reduction of fatigue and irritation while training [11], and increased muscle control [12]. In addition, serious games can offer a faster route to myoelectric training after limb loss [11], as a game would likely not rely on socket fitting or full wound closure. Furthermore, it can make the training more enjoyable and engaging [11], as well as affordable and accessible for the home environment [10]. It also has the potential to assist the user with their body image [11], decrease phantom limb pain [11], and let the user feel more in charge of their own rehabilitation [8,10], while at the same time make it feel less like rehabilitation [8].

The prevailing view is that this combination of positive effects has the potential to significantly add to the existing prosthetic training and lead to a reduction in prosthesis abandonment, which has been linked to a lack of motivation and engagement [13] and poor training [10]. The performance of virtual prosthetic training at home can also offer benefits to the therapists. As a supplement to existing training regimes, it can offer an objective measure of how diligently the patient is doing their exercises at home and of their improvements [10]. It also has the potential to decrease rehabilitation times and the time necessary for each patient, thereby reducing the workload for therapists [8].

We investigated papers that included any virtual training or assessment for upper limb prosthesis control using myoelectric signals as input. The included papers were identified during investigation of the literature and has been augmented with systematic searches in multiple databases, including PubMed, Web of Science, and Google Scholar. This led to the inclusion of 55 journal articles and conference papers, with a total of 59 different virtual training programs. CAG classified these programs into two categories, namely serious games and simulators (Table 1), according to Narayanasamy et al [14]. Both training simulators and serious games are interactive simulations in a virtual environment with the purpose of skill development. Simulators often duplicate real-world scenarios, require standard operational procedures, are not designed for entertainment, have no secondary purpose, and usually do not have an obvious final state. Conversely, a serious game is set in a fictitious scenario, provides various challenges, allows for entertainment, and allows the user to develop gameplay patterns while trying to achieve game-specific goals. This can include an end state. Therefore, some of the programs are classified as "simulators," even when the authors identified them as "games." The programs were further classified by the type of task the user was given, the type of control scheme the program used, and the input and output devices that were used. This more detailed table can be found in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Categorization of the virtual training programs.

Names	Publications
Serious games	
Air-Guitar Hero (rhythm game)	[15]
WiiEMG (sports game)	[16]
Sonic Racing (racing game)	[17]
MyoBox (dexterity game)	[18,19]
Flappy Bird (sidescroller)[a]	[20]
Space Invaders (fixed shooter)[a]	[20]
MyoBeatz (rhythm game)[b]	[21]
Falling of Momo (vertical scroller)[b]	[22-25]
Volcanic Crush (reaction game)[a,b]	[10]
Dino Sprint (endless runner)[a,b]	[10]
Dino Feast (dexterity game)[a,b]	[10]
Space ARMada (fixed shooter)	[11]
SuperTuxKart (racing game)	[2,12,26,27]
Step Mania 5 (rhythm game)	[2,12,26,27]
Pospos (dexterity game)	[2,12,26,27]
Who nose?/Nose Picker (simple game)[a]	[28,29]
Smash Bro/Bash and Debris (sidescroller)[a]	[28,29]
Sushi Slap (action game)[a]	[28,29]
Crazy Meteor (multidirectional shooter)[a]	[28,29]
Dog Jump/Beeline Border Collie (sidescroller)	[28,29]
Breakout-EMG (arcade game)	[30]
Training Game Prototype[a]	[31]
Dino Claw (dexterity game)[a,b]	[10]
Training, TAC[c] test, and Crossbow Game[a]	[32]
UpBeat (rhythm game)[a,b]	[4]
Rhythm Game[a,b]	[13]
Crate Whacker (tech demo)[b]	[33]
Race the Sun (endless runner)[b]	[33]
Fruit Ninja (dexterity game)[b]	[33]
Kaiju Carnage (action game)[b]	[33]
Simulators	
UVa Neuromuscular Training System	[34,35]
Commercial software PAULA	[36]
Virtual training	[36]
Virtual training environment	[37]
Mixed reality training[b]	[38]
Virtual box and beans test[b]	[39]

XSL•FO
RenderX

Names	Publications
Virtual box and blocks test[a,b]	[40]
Virtual rehabilitation training tool[a,b]	[41]
VITA: Virtual Therapy Arm[a,b]	[42]
Exploration[a,b]	[43]
AR prosthesis simulator	[44]
Virtual training system	[45-47]
Training system	[48]
Catching simulator	[49]
Performance assessment	[50]
Catching simulator Prosthesis Gripper	[19]
MSMS (Musculoskeletal Modelling Software)	[51,52]
Prosthesis simulator	[53]
VR[d] testing environment	[54]
Virtual simulation	[55]
VR evaluation environment	[56]
Virtual reality environment System	[57]
AR[e] training system[a]	[58,59]
Myoelectric training tool	[60]
Training environment	[61]
Virtual prosthesis	[62]
Virtual model	[63]
Training platform	[64]
Manus VR Training Platform	[65]
Dual-arm EMG[f] signal control training system	[66]
Myoelectric Control Evaluation and Trainer System	[67]

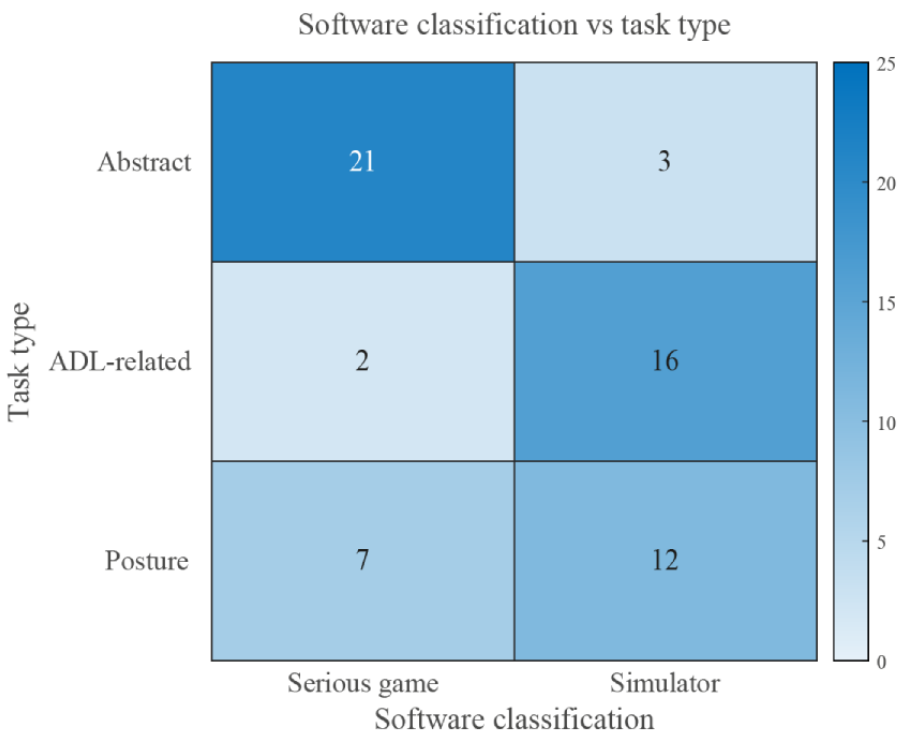
[a]Developed using the Unity engine.
[b]Uses the Myo Gesture Control Armband.
[c]TAC: Target Achievement Control.
[d]VR: Virtual Reality.
[e]AR: Augmented Reality.
[f]EMG: electromyography.

Different Approaches

The categorization of the publications in this field and the software presented therein has shown a significant split of the approaches of researchers into roughly two groups, as can be seen in Figure 1. This divide is most noticeable with regard to whether the software is classified as a game or a simulator and which type of tasks are implemented. The first approach focuses on the engagement and motivation of the user and seems to have grown in popularity in recent years. Researchers develop serious games that often have an explicit or implicit focus on game design elements to keep the user engaged in the game and therefore in the rehabilitation or training. The majority of these myogames (21/30) incorporate abstract tasks not resembling a real-life scenario. These games attempt to train the user in the use of a myoelectric prosthesis by focusing on different aspects of muscle control, including proportional control, independent control, and others. Only two games feature a task that is somewhat activities of daily living–related, both consisting of variants of a pick-and-place task, one stationary [10] and one moving in a 3D environment [31]. In a further seven games, the user is tasked with reproducing specific postures in two rhythm games [4,13], a virtual reality crossbow game [32], and four open-source games [33].

XSL•FO
RenderX

Figure 1. Heatmap of the authors' software classification against the performed tasks: serious games using abstract tasks [2,10-12,15-30], tasks related to ADL [10,31], and posture reproduction tasks [4,13,32,33]; as well as simulators using abstract tasks [34-36], ADL-related tasks [19,37-54], and posture reproduction tasks [56-67]. ADL: activities of daily living.



Publications introducing and assessing these games show that they are engaging and enjoyable to the participants. Some studies involving people with limb difference showed their willingness to use them in a home environment [8,21,27]. With regard to skill acquisition, a general increase in in-game performance is shown for a number of these games [12,25,26]. However, it was rarely tested whether these myogames increase prosthetic ability or speed up the learning process of acquiring that skill. The one research group that tested for an increase in prosthetic ability did not find evidence of a significant increase following the playing of a myogame for different control schemes [19,30].

The second approach focuses on skill transfer and therefore involves the development and investigation of simulators that mostly show the user a representation of the real world and require the performance of activities of daily living–related or posture reproduction tasks. Only two training programs that were classified as simulators used abstract tasks; these tasks were embedded in a sterile software environment and lacked distinctive game traits. In this type of research, the focus is on the effectiveness of the skill transfer from the virtual training to actual prosthetic ability. The prescribed tasks can involve recreations or tasks inspired by tests used in the assessment of prosthetic ability, like the Southampton Hand Assessment Procedure (SHAP) test [68], the Target Achievement Control (TAC) test [69], and others. The focus on task specificity for learning prosthetic skills seems like a promising approach as the results of one study indicated that skill transfer occurred. Performing a virtual task resembling the control of a prosthetic hand led to an increase in prosthetic ability [49] as opposed to when the task was to play a classic arcade game [30]. The task specificity therefore seems to have an influence on the effectiveness of virtual training; however, further research must be done to substantiate this.

The effectiveness of virtual training in increasing prosthetic ability is without doubt one of the necessary requirements for any adoption into clinical rehabilitation; however, consensus on a universal measure of effectiveness is not available. A criticism of the myogames in the game-focused group is that they work on the implied assumption that an improvement in skill performing any myoelectric task will lead to an improvement in prosthetic skill [49]. Although it has been shown that the user increases their skill in different aspects on the muscular level [10,25], it is not clear whether that influences the way or speed at which a person might acquire prosthetic skill.

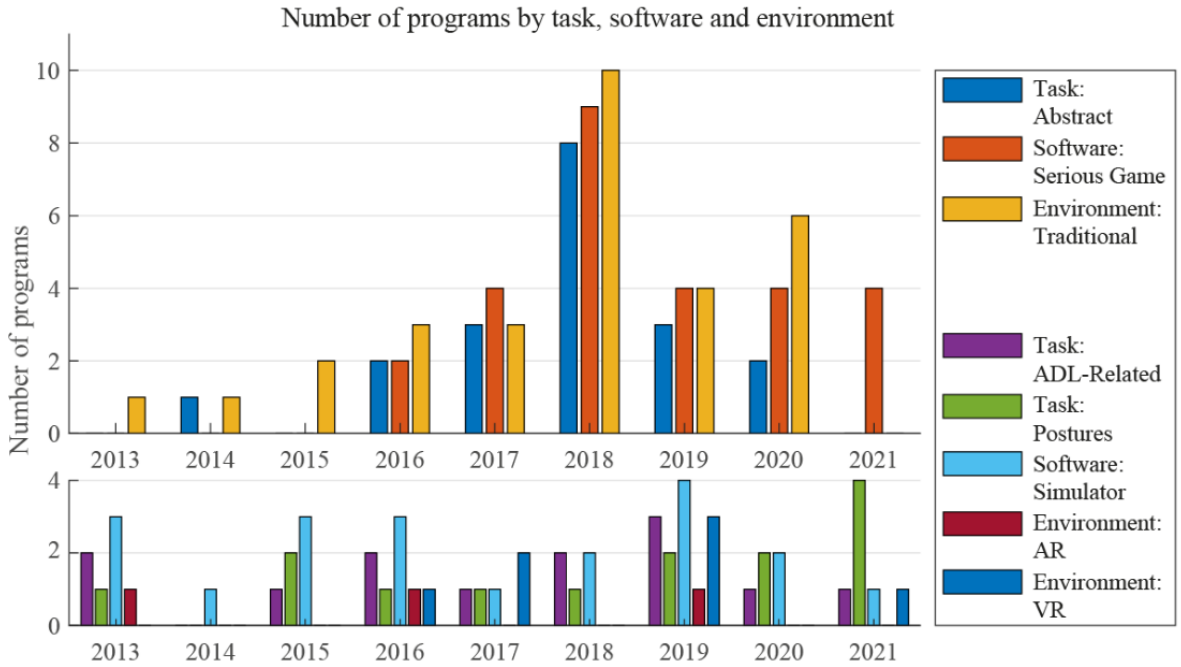
Other Influences

Figure 2 shows another interesting development regarding the first research approach. It clearly features a joint spike in more recent years in the development of serious games incorporating an abstract task and presented in a nonimmersive environment using traditional media. The development of simulators and software using other task types and environment experiences has remained comparatively steady over the same time frame. The start of this spike in publications coincides with the release of the Myo Gesture Control Armband (Thalmic Labs), a dry surface electrode armband, on the commercial market in the year 2015 [70]. The spike in publications started to decrease when the company stopped selling this product in 2018. However, even though it is no longer sold, the Myo armband is still in widespread use in research as there is currently no

XSL•FO
RenderX

commercial alternative available. In total, 30% of the programs analyzed use the Myo armband and the most recent work using it was published in 2021 [61]. The uptake in development has likely been further boosted due to a few professional commercial game development engines being free for personal and low-profit use alongside the provision of an application programming interface (API) for the Myo Armband; there are numerous open-source examples of the use of this API in these engines. The main example of one of these engines is the Unity game engine (Unity Technologies) made freely available in the year 2009 [71]; this game engine has been used in at least one-third of the published programs.

Figure 2. Number of training software introduced by task, software, and environment type. ADL: activities of daily living; AR: augmented reality; VR: virtual reality.



These two factors indicate that the development of serious games intended for prosthetic training was strongly influenced by the emergence of readily available software and hardware technology. However, the enthusiastic embrace of the newly available technology tended toward research exploring the engagement aspects of game design. This is likely due to the low barrier to entry of this approach as there are a multitude of resources for game development available and the study of engagement does not require the involvement of people with limb difference. Investigation of these serious games confirmed that people are more willing to engage in learning a task if it is an intrinsically enjoyable and motivating experience [2]. Such research has also shown that with these games, participants are able to quickly master fine control of their muscles [10,25]. However, this research often tacitly assumed an efficacy in skill transfer by this virtual muscle training, which has yet to be substantiated. As such, it is not clear whether this increase in motor control would lead to enhanced prosthesis control and which types of games might be more conducive to learning how to use a prosthetic device. Therefore, at this point in time, serious games are not serious enough to train upper limb prosthesis use effectively.

Where Do We Go From Here?

In the research targeting other conditions, such as stroke rehabilitation, the main target is to get the user to move their respective body part more to regain a substantial degree of control over it. The reason for the strong focus on the engagement and motivation of users to increase repetitions of a movement is therefore clear. However, using a myoelectric upper limb prosthesis requires the user to acquire a completely new set of skills. This can mean to either retrain or newly train muscles and their associated uses, depending on whether the limb difference is acquired or congenital. Therefore, a necessary requirement for a serious game in this field to be considered for clinical adoption would be evidence of a benefit to prosthetic ability (ie, evidence that the skill learned in the game transfers to the use of an actual myoelectric prosthesis). So far this kind of skill transfer has only been shown for software that we classified as simulators. It is hypothesized that the task specificity of the actions performed virtually allows the transfer to the real world to occur [49].

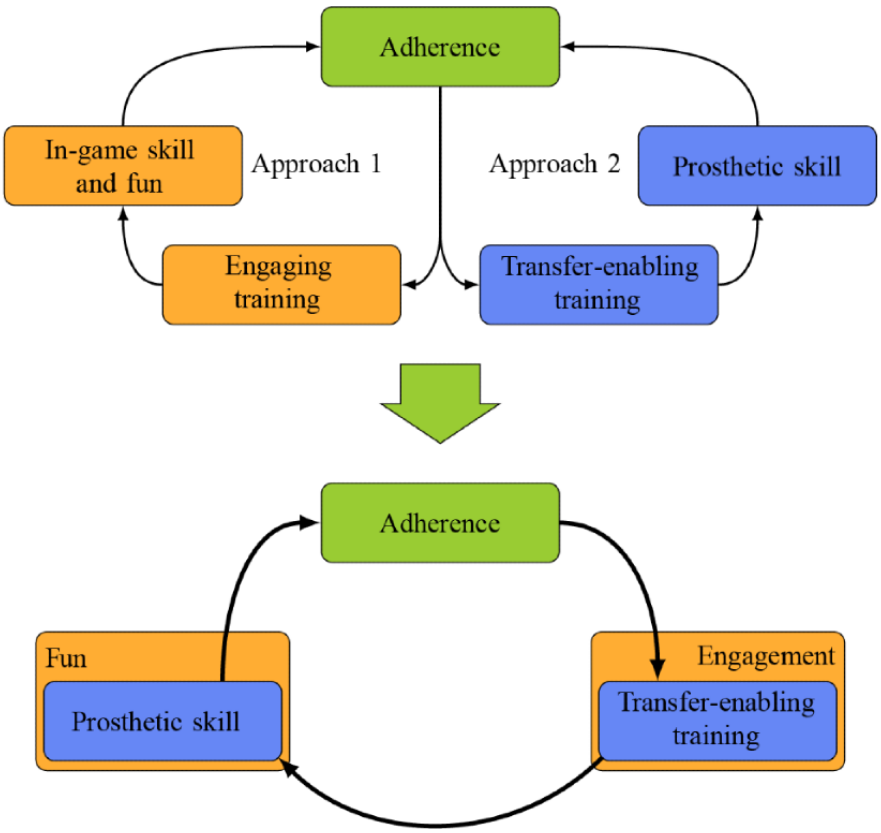
Research in this field needs to establish viable paths for transfer to occur before focusing on the topic of engaging and motivating the target user group. Serious games intended for prosthetic training need to show their benefit for prosthetic ability, be it direct or indirect. Hence, a sensible approach for the development of such a serious game could be to first demonstrate which types of tasks allow transfer at all and then to develop the engaging and motivating game structure around it. As with other conditions, researchers employ the attractiveness of games to actively engage users; however, the clinical benefit cannot be neglected or compromised. The two different approaches in this field encourage separate habit loops when they should merge and form a single loop more beneficial for the user, as shown in Figure 3. Engagement should not be viewed separately but in conjunction with transfer-enabling

XSL•FO
RenderX

training to enhance the habit formation of the user to adhere to the training.

Therefore, the research should establish one or multiple tasks that either directly enable skill transfer to prosthetic use or show evidence of supporting the acquisition of prosthetic skill. Built on these tasks, an engaging and motivating platform should be implemented, which should enable users to increase their prosthetic skill while having fun. The positive reinforcement of the skill increase combined with the fun experienced while training should have a positive effect on adherence to the training and therefore on the long-term success of the intervention.

Figure 3. Diagram of the existent and recommended habit loops.



Furthermore, the effect of the introduction of myogames on therapists' workloads should be determined; this is highly dependent on the nature of the training and whether there is a need for the direct involvement of the therapist, which could potentially result in a similar or even larger workload [72]. As this factor depends strongly on the design of the program, it emphasizes the importance of smart development including the input and feedback of all parties involved, including clinicians and therapists, to lead to a product that benefits everyone.

In conclusion, research on prosthetic training has confirmed that myoelectric skills can be acquired with serious games. However, for the development of a viable serious game intended for prosthetic training, it is necessary to validate the "serious" part of the game, namely the tasks that would allow for skill transfer. Serious games for prosthetic training can only be expected to yield fruitful results beyond engagement when they incorporate tasks that are found to facilitate prosthetic skill. We recommend that the research community investigates which types of myogame tasks might facilitate transfer, as the only existing results at the time of writing this paper indicate a lack of effectiveness [19,30,73]. This lack does not necessarily hold true for all tasks that are not related to activities of daily living, however, and ignoring abstract tasks entirely would exclude a wide range of possible avenues for prosthetic game development.

It would be beneficial to be more accurate regarding the terminology used in the field and, if the term "game" is used, to specify the incorporated game design elements explicitly. More long-term and ideally home-based experiments are needed to conclusively test for any prosthetic skill transfer that might occur with the consistent use of prosthetic gaming devices. Even though previous studies indicate that no change in prosthetic ability occurs after training with a myogame [19,30,73], these only tested the effect of comparatively short training sessions with able-bodied people or very small groups of prosthesis users. It should also be tested whether prosthetic gaming has the potential to support traditional prosthetic training by allowing for supplementary practice sessions between visits to the prosthetist.

XSL•FO
RenderX

Acknowledgments

This work was supported by the Leverhulme Doctoral Scholarship Programme in Behaviour Informatics (DS-2017-015), the National Institute of Health Research (NIHR) via Devices 4 Dignity Starworks STWK-006, and the Engineering and Physical Sciences Research Council (EPSRC) via grant EP/R004242/1.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed categorization of the virtual training programs.
[PDF File (Adobe PDF File), 39 KB - games_v9i4e28079_app1.pdf]

References

1. Lohse K, Shirzad N, Verster A, Hodges N, Van der Loos HFM. Video games and rehabilitation: using design principles to enhance engagement in physical therapy. J Neurol Phys Ther 2013 Dec;37(4):166-175. [doi: 10.1097/NPT.0000000000000017] [Medline: 24232363]
2. Prahm C, Kayali F, Vujaklija I, Sturma A, Aszmann O. Increasing motivation, effort and performance through game-based rehabilitation for upper limb myoelectric prosthesis control. In: 2017 International Conference on Virtual Rehabilitation (ICVR). 2017 Presented at: International Conference on Virtual Rehabilitation (ICVR); June 19-22, 2017; Montreal, QC, Canada. [doi: 10.1109/ICVR.2017.8007517]
3. Tatla SK, Shirzad N, Lohse KR, Virji-Babul N, Hoens AM, Holsti L, et al. Therapists' perceptions of social media and video game technologies in upper limb rehabilitation. JMIR Serious Games 2015;3(1):e2 [FREE Full text] [doi: 10.2196/games.3401] [Medline: 25759148]
4. Melero M, Hou A, Cheng E, Tayade A, Lee SC, Unberath M, et al. Upbeat: Augmented Reality-Guided Dancing for Prosthetic Rehabilitation of Upper Limb Amputees. J Healthc Eng 2019;2019:2163705. [doi: 10.1155/2019/2163705] [Medline: 31015903]
5. Koutsiana E, Ladakis I, Fotopoulos D, Chytas A, Kilintzis V, Chouvarda I. Serious Gaming Technology in Upper Extremity Rehabilitation: Scoping Review. JMIR Serious Games 2020 Dec 11;8(4):e19071. [doi: 10.2196/19071] [Medline: 33306029]
6. Bonnechère B. Serious games in physical rehabilitation: from theory to practice. Heidelberg, Germany: Springer International Publishing; 2018.
7. Pannese L, Bo G, Lawo M, Gabrielli S. The Rehab@Home project: Engaging game-based home rehabilitation for improved quality of life. In: Proceedings of the SEGAMED Conference. 2013 Presented at: SEGAMED Conference; 2013; Nice, France.
8. Garske CA, Dyson M, Dupan S, Nazarpour K. Perception of Game-Based Rehabilitation in Upper Limb Prosthetic Training: Survey of Users and Researchers. JMIR Serious Games 2021 Feb 01;9(1):e23710. [doi: 10.2196/23710] [Medline: 33522975]
9. Lovely DF, Stocker D, Scott RN. A computer-aided myoelectric training system for young upper limb amputees. Journal of Microcomputer Applications 1990 Jul;13(3):245-259. [doi: 10.1016/0745-7138(90)90026-4]
10. Winslow BD, Ruble M, Huber Z. Mobile, Game-Based Training for Myoelectric Prosthesis Control. Front Bioeng Biotechnol 2018;6:94. [doi: 10.3389/fbioe.2018.00094] [Medline: 30050900]
11. Anderson F, Bischof WF. Augmented reality improves myoelectric prosthesis training. International Journal on Disability and Human Development 2014;13(3):349-354. [doi: 10.1515/ijdhd-2014-0327]
12. Prahm C, Kayali F, Sturma A, Aszmann O. PlayBionic: Game-Based Interventions to Encourage Patient Engagement and Performance in Prosthetic Motor Rehabilitation. PM R 2018 Nov;10(11):1252-1260. [doi: 10.1016/j.pmrj.2018.09.027] [Medline: 30503232]
13. Bessa D, Rodrigues NF, Oliveira E, Kolbenschag J, Prahm C. Designing a serious game for myoelectric prosthesis control. In: 2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH). 2020 Presented at: 2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH); August 12-14, 2020; Vancouver, BC, Canada. [doi: 10.1109/segah49190.2020.9201812]
14. Narayanasamy V, Wong K, Fung C, Rai S. Distinguishing games and simulation games from simulators. Comput Entertain 2006;4(2):9. [doi: 10.1145/1129006.1129021]
15. Armiger RS, Vogelstein RJ. Air-Guitar Hero: A real-time video game interface for training and evaluation of dexterous upper-extremity neuroprosthetic control algorithms. In: 2008 IEEE Biomedical Circuits and Systems Conference. 2008 Presented at: 2008 IEEE Biomedical Circuits and Systems Conference; November 20-22, 2008; Baltimore, MD, USA p. 121-124. [doi: 10.1109/biocas.2008.4696889]
16. Oppenheim H, Armiger RS, Vogelstein RJ. WiiEMG: A real-time environment for control of the Wii with surface electromyography. In: Proceedings of 2010 IEEE International Symposium on Circuits and Systems. 2010 Presented at:

XSL•FO
RenderX

2010 IEEE International Symposium on Circuits and Systems; May 30-June 2, 2010; Paris, France p. 957-960. [doi: 10.1109/iscas.2010.5537390]
17. Martinez-Luna C, Kelly C, Rozell B, Lambrecht S, Horowitz A, Farrell T. A myoelectric video game training pilot study: changes in control signal properties. In: MEC20 Symposium Proceedings. 2020 Presented at: Myoelectric Controls and Upper Limb Prosthetics Symposium; July 2020; Fredericton, NB, Canada.
18. Kristoffersen MB, Franzke AW, van der Sluis CK, Murgia A, Bongers RM. Serious gaming to generate separated and consistent EMG patterns in pattern-recognition prosthesis control. Biomedical Signal Processing and Control 2020 Sep;62:102140. [doi: 10.1016/j.bspc.2020.102140]
19. Kristoffersen MB, Franzke AW, Bongers RM, Wand M, Murgia A, van der Sluis CK. User training for machine learning controlled upper limb prostheses: a serious game approach. J Neuroeng Rehabil 2021 Feb 12;18(1):32. [doi: 10.1186/s12984-021-00831-5] [Medline: 33579326]
20. Radhakrishnan M, Smailagic A, French B, Siewiorek D, Balan R. Design and assessment of myoelectric games for prosthesis training of upper limb amputees. In: 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). 2019 Presented at: 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops); March 11-15, 2019; Kyoto, Japan p. 151-157. [doi: 10.1109/percomw.2019.8730824]
21. Prahm C, Kayali F, Aszmann O. MyoBeatz: using music and rhythm to improve prosthetic control in a mobile game for health. In: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH). 2019 Presented at: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); August 5-7, 2019; Kyoto, Japan URL: https://doi.org/10.1109/SeGAH.2019.8882432 [doi: 10.1109/segah.2019.8882432]
22. Tabor A, Kienzle A, Smith C, Watson A, Wuertz J, Hanna D. The Falling of Momo: a myoelectric controlled game to support research in prosthesis training. In: CHI PLAY Companion '16: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. 2016 Presented at: CHI PLAY '16: The annual symposium on Computer-Human Interaction in Play; October 2016; Austin, TX, USA p. 71-77. [doi: 10.1145/2968120.2971806]
23. Tabor A, Bateman S, Scheme E. Game-based myoelectric training. In: CHI PLAY Companion '16: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. 2016 Presented at: CHI PLAY '16: The annual symposium on Computer-Human Interaction in Play; October 2016; Austin, TX, USA p. 299-306. [doi: 10.1145/2968120.2987731]
24. Tabor A, Bateman S, Scheme E, Flatla D, Gerling K. Designing game-based myoelectric prosthesis training. In: CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17: CHI Conference on Human Factors in Computing Systems; May 2017; Denver, CO, USA p. 1352-1363. [doi: 10.1145/3025453.3025676]
25. Tabor A, Bateman S, Scheme E. Evaluation of Myoelectric Control Learning Using Multi-Session Game-Based Training. IEEE Trans Neural Syst Rehabil Eng 2018 Sep;26(9):1680-1689. [doi: 10.1109/TNSRE.2018.2855561] [Medline: 30010580]
26. Prahm C, Vujaklija I, Kayali F, Purgathofer P, Aszmann OC. Game-Based Rehabilitation for Myoelectric Prosthesis Control. JMIR Serious Games 2017 Feb 09;5(1):e3 [FREE Full text] [doi: 10.2196/games.6026] [Medline: 28183689]
27. Prahm C, Kayali F, Sturma A, Aszmann O. Recommendations for games to increase patient motivation during upper limb amputee rehabilitation. In: Converging Clinical and Engineering Research on Neurorehabilitation II. 2017 Presented at: 3rd International Conference on NeuroRehabilitation (ICNR2016); October 18-21, 2016; Segovia, Spain p. 1161. [doi: 10.1007/978-3-319-46669-9_188]
28. Smith P, Dombrowski M, Buyssens R, Barclay P. Usability testing games for prosthetic training. In: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH). 2018 Presented at: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH); May 16-18, 2018; Vienna, Austria. [doi: 10.1109/segah.2018.8401376]
29. Smith PA, Dombrowski M, Buyssens R, Barclay P. The Impact of a Custom Electromyograph (EMG) Controller on Player Enjoyment of Games Designed to Teach the Use of Prosthetic Arms. Comput Game J 2018 May 2;7(2):131-147. [doi: 10.1007/s40869-018-0060-0]
30. van Dijk L, van der Sluis CK, van Dijk HW, Bongers RM. Learning an EMG Controlled Game: Task-Specific Adaptations and Transfer. PLoS One 2016;11(8):e0160817. [doi: 10.1371/journal.pone.0160817] [Medline: 27556154]
31. Dyson M, Olsen J, Nazarpour K. A home-based myoelectric training system for children. In: MEC20 Symposium Proceedings. 2020 Presented at: Myoelectric Controls and Upper Limb Prosthetics Symposium; July 2020; Fredericton, NB, Canada.
32. Woodward RB, Hargrove LJ. Adapting myoelectric control in real-time using a virtual environment. J Neuroeng Rehabil 2019 Jan 16;16(1):11. [doi: 10.1186/s12984-019-0480-5] [Medline: 30651109]
33. Hashim N, Abd Razak NA, Gholizadeh H, Abu Osman NA. Video Game-Based Rehabilitation Approach for Individuals Who Have Undergone Upper Limb Amputation: Case-Control Study. JMIR Serious Games 2021 Feb 04;9(1):e17017 [FREE Full text] [doi: 10.2196/17017] [Medline: 33538698]

XSL•FO
RenderX

34. De La Rosa R, Alonso A, De La Rosa S, Abásolo D. Myo-Pong: A neuromuscular game for the UVa-neuromuscular training system platform. In: 2008 Virtual Rehabilitation. 2008 Presented at: International Workshop on Virtual Rehabilitation; August 25-27, 2008; Vancouver, BC, Canada. [doi: 10.1109/icvr.2008.4625124]
35. De La Rosa R, De La Rosa S, Alonso A, Val L. The UVa-neuromuscular training system platform. In: IWANN '09: Proceedings of the 10th International Work-Conference on Artificial Neural Networks: Part II: Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living. 2009 Presented at: 10th International Work-Conference on Artificial Neural Networks: Part II: Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living; June 10-12, 2009; Salamanca, Spain p. 863-869. [doi: 10.1007/978-3-642-02481-8_131]
36. Terlaak B, Bouwsema H, Van Der Sluis CK, Bongers RM. Virtual Training of the Myosignal. PLoS One 2015 Sep;10(9):e0137161. [doi: 10.1371/journal.pone.0137161] [Medline: 26351838]
37. Cavalcante R, Gaballa A, Cabibihan J, Soares A, Lamounier E. The importance of sensory feedback to enhance embodiment during virtual training of myoelectric prostheses users. In: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). 2021 Presented at: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); March 27-April 1, 2021; Lisbon, Portugal. [doi: 10.1109/vrw52623.2021.00161]
38. Sharma A, Hunt CL, Maheshwari A, Osborn L, Lévay G, Kaliki RR, et al. A mixed-reality training environment for upper limb prosthesis control. In: 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS). 2018 Presented at: 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS); October 17-19, 2018; Cleveland, OH, USA. [doi: 10.1109/biocas.2018.8584739]
39. Prahm C, Schulz A, Paaben B, Schoisswohl J, Kaniusas E, Dorffner G, et al. Counteracting Electrode Shifts in Upper-Limb Prosthesis Control via Transfer Learning. IEEE Trans Neural Syst Rehabil Eng 2019 May;27(5):956-962. [doi: 10.1109/tnsre.2019.2907200]
40. Hashim NA, Razak NAA, Osman NAA. Comparison of Conventional and Virtual Reality Box and Blocks Tests in Upper Limb Amputees: A Case-Control Study. IEEE Access 2021;9:76983-76990. [doi: 10.1109/ACCESS.2021.3072988]
41. Dhawan D, Barlow M, Lakshika E. Prosthetic rehabilitation training in Virtual Reality. In: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH). 2019 Presented at: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); August 5-7, 2019; Kyoto, Japan. [doi: 10.1109/segah.2019.8882455]
42. Nissler C, Nowak M, Connan M, Büttner S, Vogel J, Kossyk I, et al. VITA - an everyday virtual reality setup for prosthetics and upper-limb rehabilitation. J Neural Eng 2019 Apr;16(2):026039. [doi: 10.1088/1741-2552/aaf35f] [Medline: 30864550]
43. Phelan I, Arden M, Garcia C, Roast C. Exploring virtual reality and prosthetic training. In: 2015 IEEE Virtual Reality (VR). 2015 Presented at: IEEE Annual International Symposium Virtual Reality; March 23-27, 2015; Arles, France. [doi: 10.1109/vr.2015.7223441]
44. Lamounier EA, Lopes K, Cardoso A, Soares AB. Using Augmented Reality Techniques to Simulate Myoelectric Upper Limb Prostheses. J Bioengineer & Biomedical Sci 2013;S1:1-6. [doi: 10.4172/2155-9538.s1-010]
45. Nakamura G, Shibanoki T, Shima K, Kurita Y, Hasegawa M, Otsuka A, et al. A training system for the MyoBock hand in a virtual reality environment. In: 2013 IEEE Biomedical Circuits and Systems Conference (BioCAS). 2013 Presented at: 2013 IEEE Biomedical Circuits and Systems Conference (BioCAS); October 31-November 2, 2013; Rotterdam, Netherlands p. 61-64. [doi: 10.1109/biocas.2013.6679640]
46. Nakamura G, Shibanoki T, Mizobe F, Masuda A, Honda Y, Chin T, et al. A high-fidelity virtual training system for myoelectric prostheses using an immersive HMD. In: i-CREATe 2016: Proceedings of the international Convention on Rehabilitation Engineering & Assistive Technology. 2016 Presented at: i-CREATe 2016: international Convention on Rehabilitation Engineering & Assistive Technology; July 2016; Singapore p. 1-4.
47. Nakamura G, Shibanoki T, Kurita Y, Honda Y, Masuda A, Mizobe F, et al. A virtual myoelectric prosthesis training system capable of providing instructions on hand operations. International Journal of Advanced Robotic Systems 2017 Sep 19;14(5):172988141772845. [doi: 10.1177/1729881417728452]
48. Takeuchi T, Wada T, Mukobaru M, Doi S. A training system for myoelectric prosthetic hand in virtual environment. 2007 IEEE/ICME International Conference on Complex Medical Engineering 2007:1351-1356. [doi: 10.1109/ICCME.2007.4381964] [Medline: 22630358]
49. van Dijk L, van der Sluis CK, van Dijk HW, Bongers RM. Task-Oriented Gaming for Transfer to Prosthesis Use. IEEE Trans Neural Syst Rehabil Eng 2016 Dec;24(12):1384-1394. [doi: 10.1109/TNSRE.2015.2502424] [Medline: 26625419]
50. Hargrove L, Losier Y, Lock B, Englehart K, Hudgins B. A real-time pattern recognition based myoelectric control usability study implemented in a virtual environment. In: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2007 Presented at: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 22-26, 2007; Lyon, France p. 4842-4845. [doi: 10.1109/IEMBS.2007.4353424]
51. Davoodi R, Loeb GE. MSMS software for VR simulations of neural prostheses and patient training and rehabilitation. Stud Health Technol Inform 2011;163:156-162. [Medline: 21335781]
52. Davoodi R, Loeb GE. Development of a Physics-Based Target Shooting Game to Train Amputee Users of Multijoint Upper Limb Prostheses. Presence: Teleoperators and Virtual Environments 2012;21(1):85-95. [doi: 10.1162/pres_a_00091]

XSL•FO
RenderX

53. Lambrecht JM, Pulliam CL, Kirsch RF. Virtual reality environment for simulating tasks with a myoelectric prosthesis: an assessment and training tool. J Prosthet Orthot 2011 Apr;23(2):89-94. [doi: 10.1097/JPO.0b013e318217a30c] [Medline: 23476108]
54. Blana D, Kyriacou T, Lambrecht JM, Chadwick EK. Feasibility of using combined EMG and kinematic signals for prosthesis control: A simulation study using a virtual reality environment. J Electromyogr Kinesiol 2016 Aug;29:21-27 [FREE Full text] [doi: 10.1016/j.jelekin.2015.06.010] [Medline: 26190031]
55. Soares A, Andrade A, Lamounier E, Carrijo R. The development of a virtual myoelectric prosthesis controlled by an EMG pattern recognition system based on neural networks. Journal of Intelligent Information Systems 2003;21:127-141. [doi: 10.1023/A:1024758415877]
56. Côté-Allard U, Gagnon-Turcotte G, Phinyomark A, Glette K, Scheme E, Laviolette F, et al. A Transferable Adaptive Domain Adversarial Neural Network for Virtual Reality Augmented EMG-Based Gesture Recognition. IEEE Trans Neural Syst Rehabil Eng 2021;29:546-555. [doi: 10.1109/tnsre.2021.3059741]
57. Resnik L, Etter K, Klinger SL, Kambe C. Using virtual reality environment to facilitate training with advanced upper-limb prosthesis. J Rehabil Res Dev 2011;48(6):707-718. [doi: 10.1682/jrrd.2010.07.0127] [Medline: 21938657]
58. Boschmann A, Dosen S, Werner A, Raies A, Farina D. A novel immersive augmented reality system for prosthesis training and assessment. In: 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI). 2016 Presented at: 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI); February 24-27, 2016; Las Vegas, NV, USA. [doi: 10.1109/bhi.2016.7455889]
59. Boschmann A, Neuhaus D, Vogt S, Kaltschmidt C, Platzner M, Dosen S. Immersive augmented reality system for the training of pattern classification control with a myoelectric prosthesis. J Neuroeng Rehabil 2021 Feb 04;18(1):25 [FREE Full text] [doi: 10.1186/s12984-021-00822-6] [Medline: 33541376]
60. Dawson MR, Fahimi F, Carey JP. The development of a myoelectric training tool for above-elbow amputees. Open Biomed Eng J 2012;6:5-15 [FREE Full text] [doi: 10.2174/1874230001206010005] [Medline: 22383905]
61. Al-Jumaily A, Olivares RA. Electromyogram (EMG) driven system based virtual reality for prosthetic and rehabilitation devices. In: iiWAS '09: Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services. 2009 Presented at: iiWAS '09: 11th International Conference on Information Integration and Web-based Applications & Services; December 2009; Kuala Lumpur, Malaysia p. 582-586. [doi: 10.1145/1806338.1806448]
62. Eriksson L, Sebelius F, Balkenius C. Neural control of a virtual prosthesis. In: ICANN 98: Proceedings of the 8th International Conference on Artificial Neural Networks. 1998 Presented at: International Conference on Artificial Neural Networks; September 2-4, 1998; Skövde, Sweden p. 905-910. [doi: 10.1007/978-1-4471-1599-1_141]
63. Muri F, Carbajal C, Echenique AM, Fernández H, López NM. Virtual reality upper limb model controlled by EMG signals. In: Journal of Physics: Conference Series 477. 2013 Presented at: 19th Argentinean Bioengineering Society Congress (SABI 2013); September 4-6, 2013; Tucumán, Argentina. [doi: 10.1088/1742-6596/477/1/012041]
64. Perry BN, Armiger RS, Yu KE, Alattar AA, Moran CW, Wolde M, et al. Virtual Integration Environment as an Advanced Prosthetic Limb Training Platform. Front Neurol 2018;9:785. [doi: 10.3389/fneur.2018.00785] [Medline: 30459696]
65. Pons J, Ceres R, Rocon E, Levin S, Markovitz I, Saro B, et al. Virtual reality training and EMG control of the MANUS hand prosthesis. Robotica 2005;23(3):311-317. [doi: 10.1017/s026357470400133x]
66. Shibanoki T, Nakamura G, Tsuji T, Hashimoto K, Chin T. A new approach for training on EMG-based prosthetic hand control. In: 2020 IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech). 2020 Presented at: 2020 IEEE 2nd Global Conference on Life Sciences and Technologies (LifeTech); March 10-12, 2020; Kyoto, Japan p. 307-308. [doi: 10.1109/LifeTech48969.2020.1570620346]
67. Dupont AC, Morin EL. A myoelectric control evaluation and trainer system. IEEE Trans Rehab Eng 1994;2(2):100-107. [doi: 10.1109/86.313151]
68. Light CM, Chappell PH, Kyberd PJ. Establishing a standardized clinical assessment tool of pathologic and prosthetic hand function: normative data, reliability, and validity. Arch Phys Med Rehabil 2002 Jun;83(6):776-783. [doi: 10.1053/apmr.2002.32737] [Medline: 12048655]
69. Simon AM, Hargrove LJ, Lock BA, Kuiken TA. Target Achievement Control Test: evaluating real-time myoelectric pattern-recognition control of multifunctional upper-limb prostheses. J Rehabil Res Dev 2011;48(6):619-627. [doi: 10.1682/jrrd.2010.08.0149] [Medline: 21938650]
70. Thalmic Labs Myo armband hits consumer release, for sale on Amazon. IT Business Canada. URL: https://www.itbusiness.ca/news/thalmic-labs-myo-armband-hits-consumer-release-for-sale-on-amazon/54056 [accessed 2021-02-16]
71. Unity 2.6 Released And Now Free!. Unity Technologies. URL: https://unity.com/our-company/newsroom/unity-2-6-released-and-now-free [accessed 2021-02-16]
72. Almeida J, Nunes F. The Practical Work of Ensuring Effective Use of Serious Games in a Rehabilitation Clinic: A Qualitative Study. JMIR Rehabil Assist Technol 2020 Feb 21;7(1):e15428. [doi: 10.2196/15428] [Medline: 32130177]
73. Heerschop A, van der Sluis CK, Otten E, Bongers RM. Performance among different types of myocontrolled tasks is not related. Hum Mov Sci 2020 Apr;70:102592 [FREE Full text] [doi: 10.1016/j.humov.2020.102592] [Medline: 32217210]

XSL•FO
RenderX

Abbreviations

EMG: electromyography
SHAP: Southampton Hand Assessment Procedure
TAC: Target Achievement Control

Edited by N Zary; submitted 19.02.21; peer-reviewed by R Armiger, E Donoso Brown; comments to author 16.04.21; revised version received 09.06.21; accepted 25.08.21; published 08.11.21.

Please cite as:
Garske CA, Dyson M, Dupan S, Morgan G, Nazarpour K
Serious Games Are Not Serious Enough for Myoelectric Prosthetics
JMIR Serious Games 2021;9(4):e28079
URL: https://games.jmir.org/2021/4/e28079
doi:10.2196/28079
PMID:34747715

©Christian Alexander Garske, Matthew Dyson, Sigrid Dupan, Graham Morgan, Kianoush Nazarpour. Originally published in JMIR Serious Games (https://games.jmir.org), 08.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Long-term Effectiveness and Adoption of a Cellphone Augmented Reality System on Patients with Stroke: Randomized Controlled Trial

Chong Li[1*], MS; Xinyu Song[2*], PhD; Shugeng Chen[1], PhD; Chuankai Wang[1], MS; Jieying He[1], MS; Yongli Zhang[1], MS; Shuo Xu[1], MS; Zhijie Yan[1], MS; Jie Jia[1,3,4*], PhD; Peter Shull[2*], PhD

[1]Department of Rehabilitation Medicine, Huashan Hospital, Fudan University, Shanghai, China

[2]The State Key Laboratory of Mechanical System and Vibration, Shanghai Jiao Tong University, Shanghai, China

[3]National Clinical Research Center for Aging and Medicine, Huashan Hospital, Fudan University, Shanghai, China

[4]National Center for Neurological Disorders, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Jie Jia, PhD
Department of Rehabilitation Medicine
Huashan Hospital
Fudan University
12 Wulumuqi Zhong Rd
Jing'an District
Shanghai, 200040
China
Phone: 86 136 1172 2357
Email: shannonjj@126.com

Abstract

Background: A serious game–based cellphone augmented reality system (CARS) was developed for rehabilitation of stroke survivors, which is portable, convenient, and suitable for self-training.

Objective: This study aims to examine the effectiveness of CARS in improving upper limb motor function and cognitive function of stroke survivors via conducting a long-term randomized controlled trial and analyze the patient's acceptance of the proposed system.

Methods: A double-blind randomized controlled trial was performed with 30 poststroke, subacute phase patients. All patients in both the experimental group (n=15) and the control group (n=15) performed a 1-hour session of therapy each day, 5 days per week for 2 weeks. Patients in the experimental group received 30 minutes of rehabilitation training with CARS and 30 minutes of conventional occupational therapy (OT) each session, while patients in the control group received conventional OT for the full 1 hour each session. The Fugl-Meyer Assessment of Upper Extremity (FMA-UE) subscale, Action Research Arm Test (ARAT), manual muscle test and Brunnstrom stage were used to assess motor function; the Mini-Mental State Examination, Add VS Sub, and Stroop Game were used to assess cognitive function; and the Barthel index was used to assess activities of daily living before and after the 2-week treatment period. In addition, the User Satisfaction Evaluation Questionnaire was used to reflect the patients' adoption of the system in the experimental group after the final intervention.

Results: All the assessment scores of the experimental group and control group were significantly improved after intervention. After the intervention. The experimental group's FMA-UE and ARAT scores increased by 11.47 and 5.86, respectively, and were both significantly higher than the increase of the control group. Similarly, the score of the Add VS Sub and Stroop Game in the experimental group increased by 7.53 and 6.83, respectively, after the intervention, which also represented a higher increase than that in the control group. The evaluation of the adoption of this system had 3 sub-dimensions. In terms of accessibility, the patients reported a mean score of 4.27 (SD 0.704) for the enjoyment of their experience with the system, a mean 4.33 (SD 0.816) for success in using the system, and a mean 4.67 (SD 0.617) for the ability to control the system. In terms of comfort, the patients reported a mean 4.40 (SD 0.737) for the clarity of information provided by the system and a mean 4.40 (SD 0.632) for comfort. In terms of acceptability, the patients reported a mean 4.27 (SD 0.884) for usefulness in their rehabilitation and a mean 4.67 (0.617) in agreeing that CARS is a suitable tool for home-based rehabilitation.

XSL•FO
RenderX

Conclusions: The rehabilitation based on combined CARS and conventional OT was more effective in improving both upper limb motor function and cognitive function than was conventional OT. Due to the low cost and ease of use, CARS is also potentially suitable for home-based rehabilitation.

Trial Registration: Chinese Clinical Trial Registry ChiCTR1800017568; https://tinyurl.com/xbkkyfyz

(JMIR Serious Games 2021;9(4):e30184) doi:10.2196/30184

KEYWORDS

stroke; augmented reality; serious game; upper limb motor function; cognitive function; home-based rehabilitation

Introduction

Stroke is the leading cause of mortality and permanent disability in adults worldwide [1,2]. Upper limb impairment is one of the common consequences after stroke, with up to 80% of survivors experiencing upper limb paresis after stroke onset [3,4]. Furthermore, the decline of cognitive state is present in more than 60% of stroke survivors, typically relevant to degradation of concentration, executive function, and comprehension [5]. Upper limb impairment and decreased cognitive state after a stroke can hamper the activities of the upper limb, which leads to greater dependence in the performance of activities of daily living (ADLs) [6].

The primary mechanism of functional recovery after stroke is synaptic reorganization and neurological functional recovery [7,8], and the bimodal balance-recovery model that links interhemispheric balancing of the brain is the foundation for upper limb function [9]. Studies indicate that high-intensity, repeatable, task-oriented, and feedback rehabilitation training can promote brain remodeling and functional restoration, and enhance the normal bimodal balance-recovery model [10-13]. Clinically, the improvement of upper limb function and cognitive state mainly depends on the treatment of occupational therapists. However, issues such as cost and access may limit the dose of one-on-one rehabilitation exercise with an occupational therapist [14,15]. In addition, patients may have difficulties in keeping concentration when they are doing intensive and repetitive training. Furthermore, stroke survivors still need long-term rehabilitation after being discharged from the hospital. At present, home-based rehabilitation for stroke survivors is mainly based on the exercise prescription provided by therapists. However, patients often abandon rehabilitation because of the boring and repetitive training mode. Additionally, in the progress of rehabilitation, patients can obtain better rehabilitation effects only if they have a high degree of participation. However, treatment methods for self-oriented or home-based rehabilitation are rare. Therefore, convenient rehabilitation systems that can enhance stoke survivors' enthusiasm and allow them to perform self-guided home-based rehabilitation are needed.

To address these issues, augmented reality (AR) technology has been introduced to the field of stroke rehabilitation. The AR system is a useful new technology that blends virtual objects with real scenes in real time [16]. An increasing number of studies report promising results of its application to stroke rehabilitation [17-19]. At present, the AR systems commonly used in stroke rehabilitation can be divided into conventional AR systems, mirror-based AR systems, and cellphone-based AR systems. However, existing AR systems are not suitable for self-oriented or home-based rehabilitation. Conventional AR systems can enhance patients' motivation, but such systems usually need guidance from therapists [20,21]. Mirror-based AR systems combine AR technology with mirror therapy. Patients can use AR technology for 3D mirror-image treatment, which can effectively promote the recovery of upper limb function. However, this kind of system often requires the patient to wear a head-mounted display or a large device, so patients experience discomfort or difficulty when using it [22]. AR systems based on cellphones have only been used in stroke rehabilitation in recent years. This kind of system is cheaper and more convenient than are other AR technology systems, but few articles have reported on this kind of system for self-guided or home-based rehabilitation. Therefore, AR rehabilitation systems that are suitable for self-oriented home-based treatment are rare and needed.

Based on a review of the relevant literature, we have developed a serious game-based cellphone augmented reality rehabilitation system (CARS). In a previous pilot study performed in a clinical setting, we found that CARS motivated individuals with stroke to perform task-oriented games (eg, Pyramid Reach) during a 30-minute intervention. Most patients who used CARS also reported that the exercise was more motivating than conventional occupational therapy (OT) [23]. An important question is whether or not self-guided exercise with CARS is feasible and improves upper limb function and cognitive state compared with conventional OT in a long-term intervention. In addition, whether this system is suitable for home-based rehabilitation is unknown.

To study these questions, a double-blind randomized controlled trial that compared combined CARS and conventional OT rehabilitation with conventional OT alone was performed. The objective of this paper is twofold: first, to study the long-term effectiveness of the system in improving upper limb function and cognitive state in survivors of subacute stroke; and second, to determine the acceptance of this intervention for home-based rehabilitation of poststroke survivors. We hypothesized that the participants using the combined CARS and conventional OT rehabilitation would receive at least equivalent results to those using conventional OT. We also hypothesized that CARS would be a suitable option for home-based rehabilitation.

Methods

Study Design

This was a multicenter, double-blind, 2-group randomized controlled trial comparing combined CARS and conventional

XSL•FO
RenderX

OT rehabilitation with matched control conventional OT in patients with upper limb dysfunction in the subacute phase of stroke. The study protocol was approved by the Institutional Review Board of Huashan Hospital, Fudan University (no. KY2018-248) and registered at the Chinese Clinical Trial Registry (ChiCTR1800017568) (Multimedia Appendix 1).

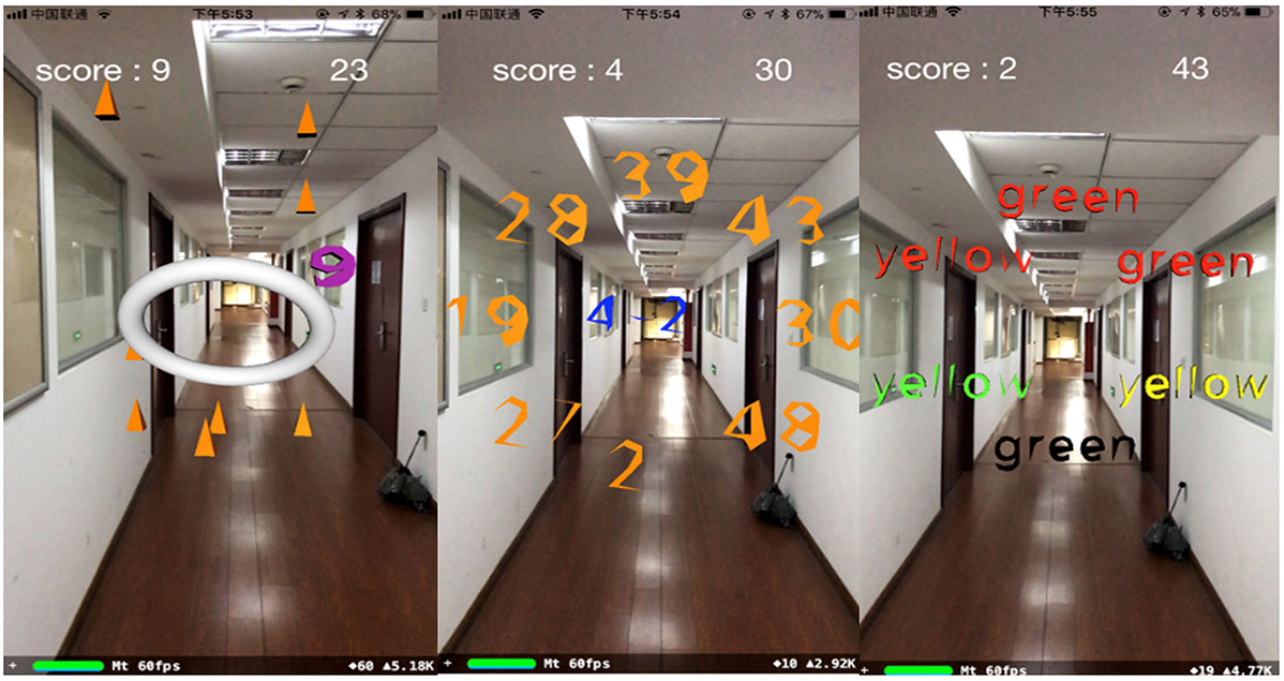
Serious Game-Based CARS

CARS was developed based on the ARKit toolbox run on an iPhone XR cellphone (Apple Inc) with an iOS 13 operating system. Stroke survivors may not be able to pick up the phone due to the weakness of their affected side. Therefore, a fingerless glove cellphone case was designed, which could easily fix the cellphone to the patient's affected hand. Patients could move their affected upper limb to use the cellphone and interact with 3D virtual targets generated on the cellphone screen (Figure 1).

Figure 1. Hardware for the cellphone augmented reality rehabilitation system. (Left) Phone: iPhone XR, iOS 13. (Center) Cellphone cases and gloves. (Right) Method of wearing.

Three serious games were developed based on CARS for improving both motor function and cognition function of stroke survivors (Figure 2) [23]. The first game is Pyramid Reach, which was designed to enhance upper limb function and concentration of stroke survivors. In this game, patients are expected to move the cellphone to touch the virtual pyramid targets. This is the basis of an engaging game to help patients reach out for virtual targets in a certain period of time. The second game is Add VS Sub (AVS), which was intended to strengthen both physical and cognitive ability. In this game, patients need to calculate the formula generated in the center of the screen and touch the correct answer in the set of numbers that appear around the formula. The third game, Stroop Game (SG), aims to train both motor and cognitive function. SG is based on the Stroop effect which is our tendency to experience difficulty naming a physical color when it is used to spell the name of a different color. These 3 serious games have the potential to keep patients engaged and motivated even though they are performing continuous and repetitive movements.

Figure 2. Three augmented reality–based serious games for rehabilitation of upper limb motor function and cognitive function. (Left) Pyramid Reach. (Center) Add VS Sub. (Right) Stroop Game.



XSL•FO
RenderX

The proposed system can record the scores of each round and the trajectory of user's upper limb distal during the training, which enables clinicians to track the change in the range of motion of the patient's affected side and the progress of the patient's recovery. This function can help clinicians give further diagnosis and rehabilitation guidance to outpatients.

Participants

Thirty participants were recruited in the study between August 2020 and March 2021. The inclusion criteria were the following: age ≥20 and <70 years; first incidence of a stroke with unilateral hemiparesis; chronicity ≥ 7 and < 180 days, Mini-Mental State Examination (MMSE) score ≥20, Brunnstrom stage for upper limb ≥3, ability to give informed consent and operate a mobile phone, and the visual and mental ability to actively participate in the protocol. Exclusion criteria were the following: history of epilepsy orthopedic alteration or pain syndrome of the upper limb, severe aphasia or other psychiatric illnesses that limit the ability to participate or give consent, visual disturbance such as visuospatial neglect, and poor sitting balance.

PASS software (NCSS Statistical Software) was used to calculate the required sample size. The Fugl-Meyer Assessment for Upper Extremity (FMA-UE) subscale was considered significant when the change value was more than 5 points [24]. To satisfy an α level of .05 and a power of 0.95, a minimum of 12 patients was required in each of the 2 groups. Assuming a dropout rate of 20%, we aimed to include 15 patients in each of the 2 groups.

Study Procedures and Interventions

All participants were invited for an initial assessment to confirm that they met the inclusion criteria. An independent researcher not involved in the study created a blocked randomization sequence using a computerized program (Microsoft Excel). Block randomization ensured equal numbers of participants for the experimental and the control group. Allocation assignments were placed in sequentially numbered, opaque, and sealed envelopes by an offsite officer not involved in the study. Patients were not blind regarding the intervention received. Once the participant passed the screening process and completed the baseline assessment, an independent researcher would open an envelope and reveal the group allocation. In addition, outcome evaluators were blinded to the group assignment.

After giving informed consent, eligible participants were allocated to 1 of 2 groups. A previous study by Saposnik et al [25] had shown that eight 60-minute sessions of Wii gaming over 2 weeks in a group of subacute stroke survivors resulted in significant improvements. Therefore, participants were instructed to train for 10 hours, divided into 5 days per week for 2 weeks. In both groups, the total treatment dose was matched at 1 hour per day. Patients allocated to the experimental group were treated for 30 minutes using CARS and 30 minutes of conventional OT per day. The 30 minutes CARS therapy consisted of 15 trails, 5 trials per game. The patients were monitored by therapists or their caretakers when they were using the system. If there was an adverse event, the therapist or clinician would examine the patient and deal with it promptly. In addition, the therapists or clinicians would record the adverse event. Patients in the control group received conventional OT for 1 hour each day. Conventional OT consisted of passive and active range of motion exercises, muscle strengthening, and functional tasks that matched CARS therapy. The occupational therapists selected relevant motor and cognitive training according to the patient's functional status.

For the experimental group, we provided 2 methods for patients to use the system (Figure 3). We encouraged patients to use their affected hand to play serious games. However, if the patients had upper limb motor function, they could use the unaffected side to assist the affected side to perform movements and interact with the targets generated from the serious games. Before the intervention, patients in the experimental group were instructed on how to use the serious game–based cellphone AR system. Patients moved their hands back and forth in front of them for 5 seconds to initialize the system and allow the cellphone to familiarize itself with the surrounding environment; they then performed a practice trial for each game to become familiar with the game function and operation.

XSL•FO
RenderX

Figure 3. Two training methods. (Left) The participant training individually with the affected side. (Right) The participant using the unaffected side to assist the affected side for training.

Outcome Measures

Motor Function Assessment

Four effective clinical-based assessments were selected and performed to evaluate the upper limb motor function of patients in both groups before and after the intervention. Primary outcomes were the FMA-UE and the Action Research Arm Test (ARAT). The FMA-UE can objectively measure arm impairment and the degree of muscle synergies present. Performance is rated on a 3-point ordinal scale from 0 to 2 with a maximum score of 66. A higher score indicates minimal or no impairment [26-28]. The internal consistency and validity are excellent [29,30]. The ARAT can objectively assess arm and hand function using observational methods. The ARAT is divided into 4 subtests of grasp, grip, pinch, and gross arm movement. Performance on each item is rated on a 4-point ordinal scale from 0 to 3 with a maximum score of 57, with a higher score indicating a better level of function [31]. The reliability and validity are excellent [32,33].

In addition, the manual muscle test (MMT) and Brunnstrom stage (BS) were selected as second motor outcomes. The MMT is a procedure for the evaluation of muscle strength, based on the effective performance of a movement in relation to the forces of gravity or manual resistance through the available range of motion. The grading of MMT ranges from 0 (no visible or palpable contraction) to 5 (full range of motion against gravity, maximal resistance) [34]. The observational cohort studies indicated good internal and external validity of the MMT [35]. BS is used by therapists to assess how well their patients are recovering. The stages of BS range from 1 (flaccid paralysis) to 6 (normal function) [36,37].

Cognition Function Assessment

To assess the change of cognitive function for patients in both the experimental and control groups, 1 clinical-based assessment and 2 cognition evaluations designed based on the proposed serious games were performed additionally before and after the intervention.

MMSE is a widely used test of cognitive function for stroke survivors and includes tests of orientation, attention, memory, language, and visual-spatial skills [38]. A higher score indicates a better level of cognitive function, with a maximum score of 30. The reliability and validity are excellent [39-41].

Game 2 (AVS) and game 3 (SG) can train the patient's ability of calculation and comprehension. The scores of the 2 games trained by patients each day were recorded by CARS, and these scores were derived and analyzed. We further designed a paper version of AVS and SG to test the change of cognitive ability in patients. The patients were asked to answer the questions from the paper version as quickly as possible within 1 minute. The number of correct answers was the score of the patients.

ADL Assessment

The Barthel index (BI) is used to measure performance in ADLs. Ten variables describing ADL and mobility are scored, with a higher number being a reflection of greater ability to function independently following hospital discharge. Scores of 0 to 20 indicate “total” dependency, 21 to 60 indicate “severe”

XSL•FO
RenderX

dependency, 61 to 90 indicate "moderate" dependency, and 91 to 99 indicate "slight" dependency [42-44]. The reliability and validity are good [45].

Questionnaire

Referring to the User Satisfaction Evaluation Questionnaire [46], we provided a 7-question questionnaire that assayed the patient's perceived engagement and the acceptability of their experience at the end of therapy. This questionnaire assessed the patient's feelings regarding the gaming experience, their perception of comfort, and their enjoyment of the game (1, strongly disagree, 2, disagree; 3, neutral; 4, agree; 5, strongly agree). We also asked an additional 2 questions about their previous experience and their suggestions about the system: (1) Do you experience any negative symptoms during or after gameplay? (2) What are some suggestions you have that you think could improve the system?

Statistical Analysis

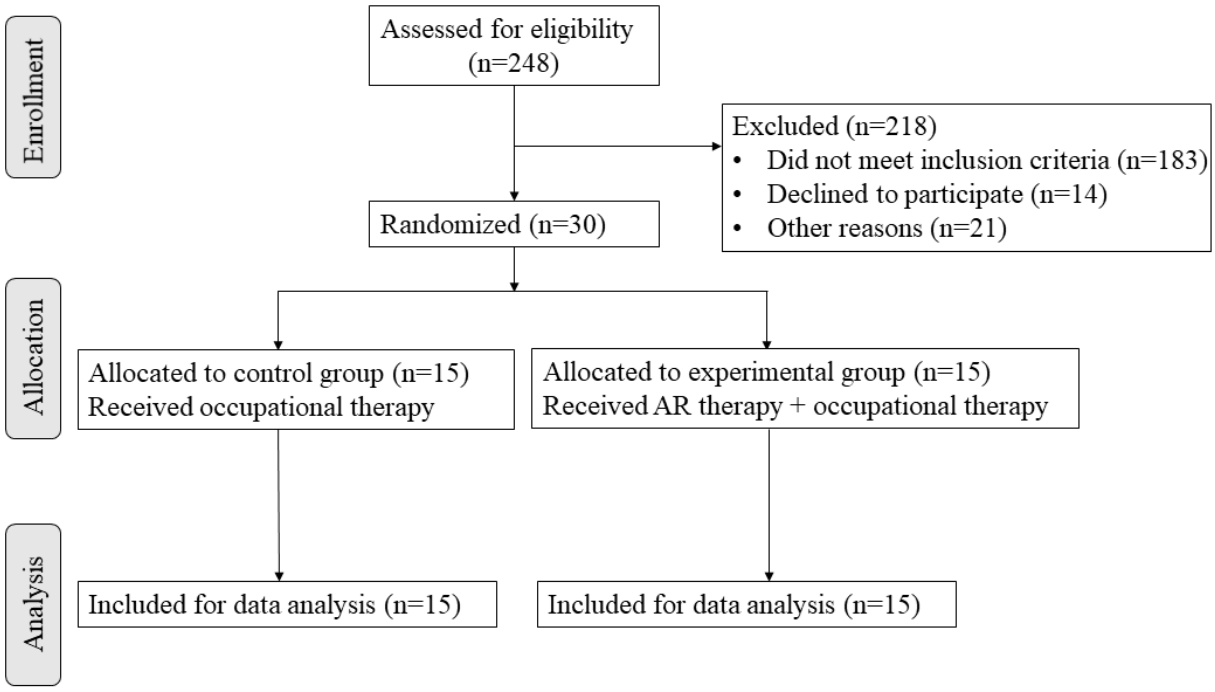
Statistical analysis was performed in SPSS version 26 (IBM Corp). Data were confirmed to have a normal distribution according to the Shapiro-Wilk normality test since the sample size was small. Two-tailed *t* tests (for continuous variables) and the chi-square test (for categorical variables) were used to compare baseline measures between the 2 treatment groups. Nonparametric tests were used if the data seemed to be nonnormally distributed. The Wilcoxon signed rank test was used for within-group analyses, and the Mann-Whitney test was used for between-group analyses. The level of statistical significance was set at a *P* value <.05. Intention to treat was applied, and missing data were replaced by the mean of the previous outcomes of the given patient.

Results

Participant Characteristics

The flowchart for participants is presented in Figure 4. Initially, 248 hospitalized patients in Huashan Hospital, Nanshi Hospital affiliated to Henan University, and The Third Rehabilitation Hospital of Shanghai were assessed for eligibility, but 218 patients did not meet the inclusion criteria. Finally, 30 patients were enrolled and randomly divided into 2 groups: 15 patients were allocated to the experimental group and 15 patients were allocated to the control group. None of the participants withdrew from the study.

Figure 4. Flowchart for participant selection and assignment. AR: augmented reality.



The patients' baseline demographic and clinical characteristics are shown in Table 1. The time from onset ranged from 9 to 160 days, with an average of 73.3 days. Furthermore, 63% (19/30) of the patients were men, and their age ranged from 26 to 75 years, with an average age of 53.07 years. No significant differences were observed between the groups regarding gender, age, duration after stroke onset, stroke type, or affected side. There were no significant differences in the baseline values of the motor outcomes between the 2 groups (FMA-UE: *P*=.56; ARAT: *P*=.68; MMT shoulder: *P*=.81; MMT elbow: *P*=.53; MMT wrist: *P*=.20). The baseline values of the cognitive outcomes (MMSE: *P*=.87; AVS: *P*=41; SG: *P*=.53) were well distributed with no significant between-group differences.

XSL•FO
RenderX

Table 1. Baseline demographic and clinical characteristics of the patients.

Variable	Experimental group (N=15)	Control group (N=15)	*t* or *z* value	*P* value
Age (years), median (IQR)	62 (24)	57 (32)	0.062[a]	.96
Gender, n (%)			0.144[b]	.70
Male	10 (66)	9 (60)		
Female	5 (33)	6 (40)		
Time from onset (days), mean (SD)	78.2 (40)	69.2 (51)	0.530[c]	0.60
Affected side, n (%)			0.133[b]	.71
Left	8 (53)	7 (46)		
Right	7 (46)	8 (53)		
MMSE[d], median (IQR)	26 (4)	25 (3)	0.168[a]	.87
FMA-UE[e], median (IQR)	30 (18)	25 (23)	0.583[a]	.56
ARAT[f], median (IQR)	14 (17)	12 (22)	0.417[a]	.68
BS-U[g], median (IQR)	3 (1)	3 (1)	0.024[a]	.99
BS-H[g], median (IQR)	3 (2)	4 (3)	0.085[a]	.93
MMT[h] shoulder, median (IQR)	3 (0)	3 (0)	0.338[a]	.81
MMT elbow, median (IQR)	3 (0)	3 (1)	0.898[a]	.53
MMT (wrist), median (IQR)	1 (3)	3 (1)	1.346[a]	.20
BI[i], mean (SD)	64.67 (12)	63 (13)	0.354[c]	.72
AVS[j], mean (SD)	17.8 (4)	19.47 (6)	0.841[c]	.41
SG[k], mean (SD)	11.67 (5)	13 (6)	0.625[c]	.53

[a]Wilcoxon rank sum test.
[b]Chi-square test.
[c]Two-tailed *t* test.
[d]MMSE: Mini-Mental State Examination.
[e]FMA-UE: Fugl-Meyer Assessment of the Upper Extremity.
[f]ARAT: Action Research Arm Test.
[g]BS: Brunnstrom stage (U: upper extremity; H: hand).
[h]MMT: manual muscle test.
[i]BI: Barthel index.
[j]AVS: Add VS Sub.
[k]SG: Stroop Game.

Comparison of Motor Function

After the intervention, both groups showed significant improvement in the FMA-UE, ARAT, BS, and MMT scores over time (Table 2). The experimental group's score of FMA-UE and ARAT increased by 11.47 and 5.86, respectively, after intervention, and were both significantly higher than the corresponding increase ih the control group (Figure 5). Additionally, a considerable change in Brunnstrom stage-hand (mean 0.60) and MMT wrist (mean 1.07) in the experimental group was found compared to the control group. Other than this, there were no between-group differences in terms of Brunnstrom stage-upper extremity, MMT shoulder, MMT elbow, or BI.

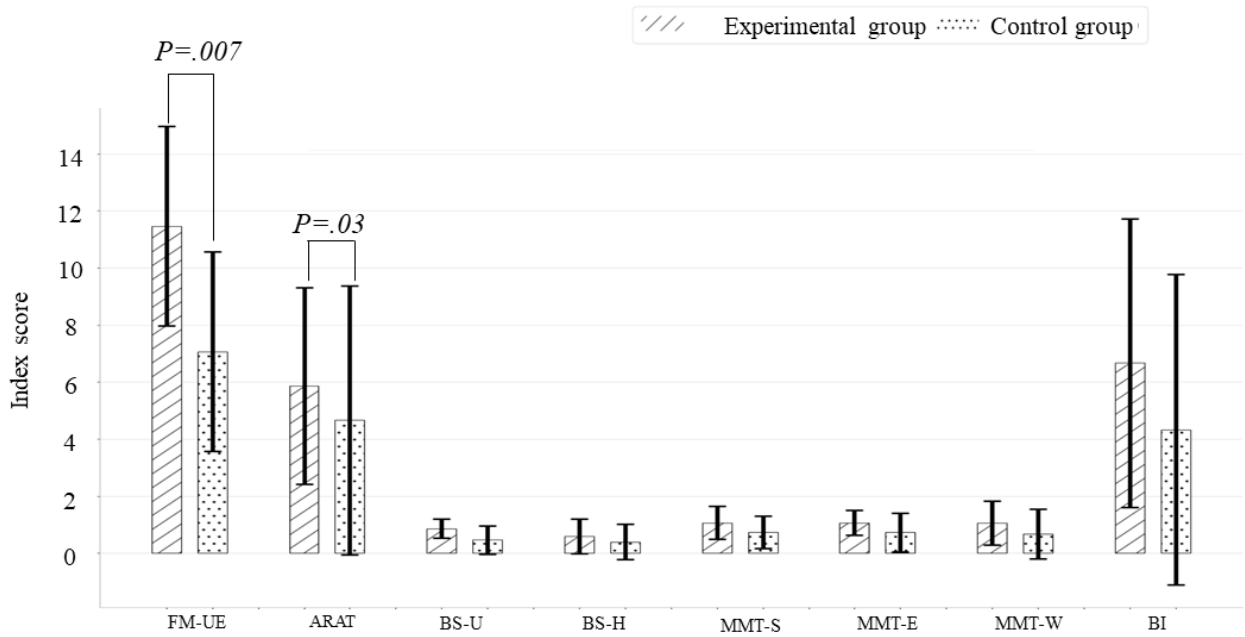
XSL•FO
RenderX

Table 2. Comparison of outcomes in the experimental group and control group.

Outcomes	Experimental group (N=15)			Control group (N=15)		
	Pretest	Posttest	*P* value	Pretest	Posttest	*P* value
FMA-UE[a], median (IQR)	30 (18)	41 (21)	.001	25 (23)	33 (26)	.001
ARAT[b], median (IQR)	14 (17)	21 (23)	.001	12 (22)	16 (24)	.001
MMSE[c], median (IQR)	26 (4)	27 (2)	.005	25 (5)	25 (4)	.004
BS-U[d], median (IQR)	3 (1)	4 (1)	< .001	3 (1)	4 (2)	.008
BS-H[d], median (IQR)	3 (2)	4 (3)	.007	4 (3)	4 (3)	.034
MMT[e] shoulder, median (IQR)	3 (0)	4 (1)	.001	3 (0)	4 (0)	.002
MMT elbow, median (IQR)	3 (0)	4 (0)	< .001	3 (1)	4 (0)	.005
MMT wrist, median (IQR)	1 (3)	3 (2)	.001	3 (1)	3 (2)	.014
BI[f], median (IQR)	65 (25)	75 (15)	.002	60 (25)	65 (20)	.01
AVS[g], median (IQR)	17 (4)	25 (5)	.001	18 (10)	23 (10)	.001
SG[h], median (IQR)	10 (6)	18 (10)	.001	12 (7)	15 (8)	.001

[a]FMA-UE: Fugl-Meyer Assessment of the Upper Extremity.
[b]ARAT: Action Research Arm Test.
[c]MMSE: Mini-Mental State Examination.
[d]BS: Brunnstrom stage (U: upper extremity; H: hand).
[e]MMT: manual muscle test.
[f]BI: Barthel index.
[g]AVS: Add VS Sub.
[h]SG: Stroop Game.

Figure 5. Longitudinal changes in motor outcomes with the experimental group showing significantly greater improvements than the control group in FMA-UE and ARAT. ARAT: Action Research Arm Test; BI: Barthel index; BS: Brunnstrom stage (U: upper extremity; H: hand); FMA-UE: Fugl-Meyer Assessment of the Upper Extremity; MMT: manual muscle test (S: shoulder; E: elbow; W: wrist).

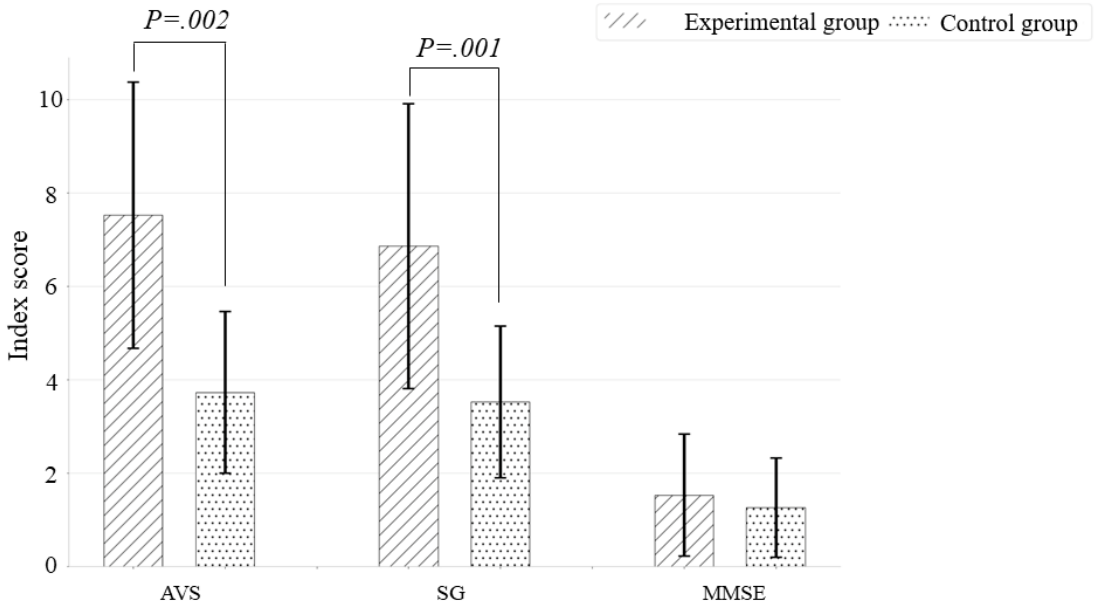


Comparison of Cognitive Function

Significant group interaction effects were observed for AVS and SG scores (Table 2). The average AVS and SG score in the experimental group increased to 7.53 and 6.83, respectively, after intervention, which was statistically significant compared to the increase in the control group (Figure 6). However, no significant between-group difference was observed in the MMSE.

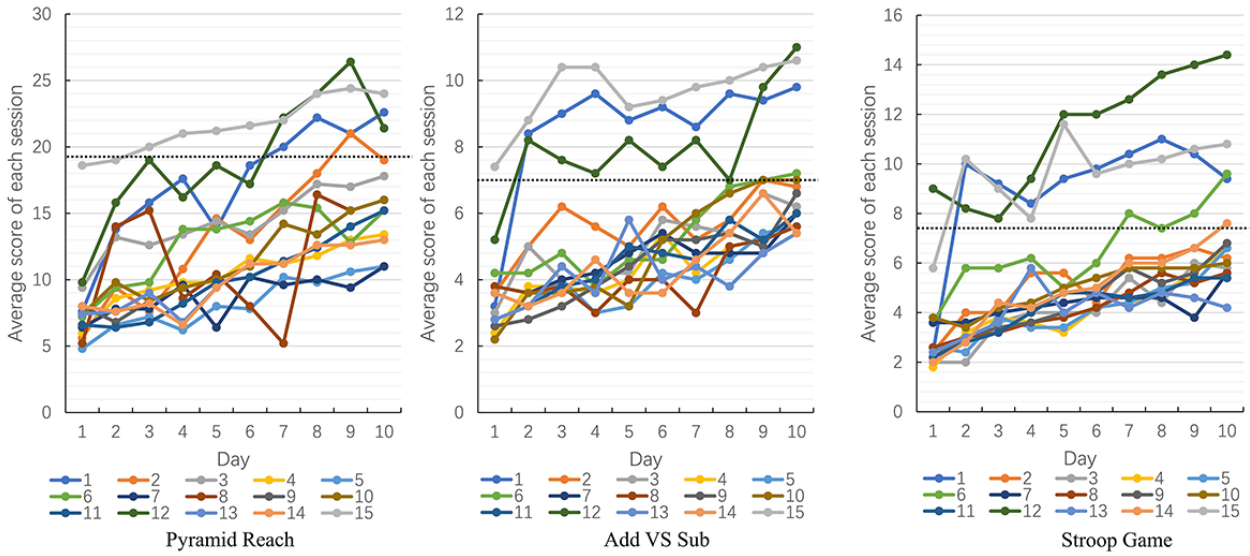
XSL•FO
RenderX

Figure 6. Longitudinal changes in Add VS Sub (AVS), Stroop Game (SG), and Mini-Mental State Examination (MMSE) with the experimental group showing significantly greater improvements than the control group in AVS and SG. AVS: Add VS Sub; MMSE: Mini-Mental State Examination; SG: Stroop Game.



We derived all patients' training scores recorded on the mobile phone. After processing and analyzing the scores, we created a line chart of each patient (Figure 7). Overall, the final mean score of 3 serious games for each patient improved to some extent compared to the first mean score.

Figure 7. Line charts of 15 patients' daily scores for 3 serious games. The average score of 5 trails represents a session, and each color represents a patient.



Clinical Acceptability

A user satisfaction questionnaire was administered for patients at the end of their final intervention in the experimental group. The 5-point rating was used to assess responses to the questions (1, strongly disagree; 2, disagree; 3, neutral; 4, agree; 5, strongly agree). In terms of accessibility, the patients reported a mean score of 4.27 (SD 0.704) for the enjoyment of their experience with the system, a mean 4.33 (SD 0.816) for success in using the system, and a mean 4.67 (SD 0.617) for the ability to control the system. In terms of comfort, the patients reported a mean 4.40 (SD 0.737) for the clarity of information provided by the system and a mean 4.40 (SD 0.632) for comfort. In terms of acceptability, the patients reported a mean 4.27 (SD 0.884) for usefulness in their rehabilitation (Table 3). All patients in the experimental group thought that CARS can be a good option for self-oriented or home-based rehabilitation.

Here, we report patients' suggestions for the system. Some patients felt that the cellphone screen was too small to see clearly. Some patients thought these 3 AR serious games should be set to a different level, so that they can be applied to patients in different stages.

XSL•FO
RenderX

Table 3. Results of the acceptability questionnaire[a].

Questions	Score, mean (SD)
Q1. Did you enjoy your experience with the system?	4.27 (0.704)
Q2. Were you successful using the system?	4.33 (0.816)
Q3. Were you able to control the system?	4.67 (0.617)
Q4. Is the information provided by the system clear?	4.40 (0.737)
Q5. Did you feel comfortable during your experience with the system?	4.40 (0.632)
Q6. Do you think that this system will be helpful for your rehabilitation?	4.27 (0.884)
Q7. Do you think this system can be used for home-based rehabilitation?	4.67 (0.617)

[a]The questionnaire includes 7 questions, each with a score of 1 to 5 (1, strongly disagree; 2, disagree; 3, neutral; 4, agree; 5, strongly agree).

Safety

No significant adverse events related to CARS occurred during the clinical research. Reported adverse events were pain and fatigue in the shoulder (n=3) and elbow (n=2), which were foreseen and already mentioned to the patients before therapy began. The use of this cellphone game-based AR system was safe and acceptable for all participants in the study.

Discussion

Principal Results

We have developed a serious game–based cellphone augmented reality rehabilitation system for upper limb recovery and improving cognitive function after stroke. Compared with conventional OT, combined CARS and conventional OT rehabilitation proved to be more effective in improving both upper limb function and cognitive function. As a new rehabilitation technique, CARS can replace part of traditional OT and be used as a supplementary treatment method in the hospital to decrease the consumption of medical resources and reduce the burden of patients. In addition, this system can be a suitable option for self-oriented or home-based rehabilitation.

Our findings showed that CARS effectively enhanced upper extremity recovery in patients with stroke. In the comparison between CARS and conventional OT, intervention using CARS showed greater improvement in the FMA-UE and ARAT scores. Comparable improvements between groups were shown in the BS, MMT, and BI. Our results showed that compared with the control group, the experiment group using CARS experienced effectively improved calculation and color-matching ability. In addition, patients in the experimental group completed the 2-week intervention using CARS without severe adverse effects, and they were satisfied with this system for self-oriented or home-based rehabilitation. We speculated that the therapeutic effectiveness of combined CARS and conventional OT rehabilitation would be equal to or greater than that of conventional OT alone, and the results confirmed our hypothesis. The results that CARS can effectively improve the upper limb function and cognitive state of patients may be related to the following factors. Studies conclude that serious games seem to be a safe rehabilitation modality for patients recovering from stroke [47]. The 3 AR serious games we designed can provide immediate feedback, enjoyment, high engagement, and task-oriented movement. Additionally, the AVS game and SG were specially designed for training calculation and color-matching ability. Moreover, visual and haptic feedback can enhance the patients' desire for interaction. Therefore, CARS can facilitate motor learning, improve cognition function, and increase rehabilitation motivation.

Although some studies have attempted to use AR-based rehabilitation systems for patients with stroke, few studies have included a control group. Assis et al [48] developed a system based on AR for upper-limb motor rehabilitation of stroke survivors. Two case studies were conducted to determine the clinical feasibility. Hoermann et al [49] presented a novel AR system in combination with a validated mirror therapy for stroke survivors. The study examined the therapeutic intervention in 5 patients. Kaneko et al [50] developed a novel approach using AR that could improve self-body cognition in stroke survivors. The study clarified the effect of the system by treating 11 patients. However, none of these studies included a control group. Additionally, most studies that compared AR therapy with conventional therapy allowed more therapy time in the experimental group [51,52]. Consequently, we are not sure whether these results indicate the AR system's greater effectiveness or are rather due to the effect of excess treatment time. Moreover, few studies have determined the acceptance for self-oriented home-based rehabilitation, and many studies have focused on only a single motor aspect rather than on multiple aspects. Therefore, we tried to design a multiaspect (motor function, cognitive function, and acceptance) randomized controlled trial with a matched intervention dose.

Compared with other systems, the proposed CARS we developed has the following advantages. First, our system requires the least amount of equipment and is also the cheapest and most convenient. CARS was developed for users with a mobile phone because almost everyone has one. The system is portable and very easy to use regardless of a person's location. Other AR systems often require independent 3D tracking systems, monitors, and interactive systems, and we integrate all 3 into a single cellphone [53,54]. Second, we use dual-task training to simultaneously train patients' motor and cognitive abilities. In addition, we adopted the training theory involving central-peripheral-central closed-loop training [55]. Patients start task training by activating the central nervous system of consciousness. When the patient completes the task, the peripheral cellphone gives the patient a short vibration stimulus,

XSL•FO
RenderX

which uploads and then strengthens the central control function of the patient.

Limitations

There are a few limitations in this study. First, although compensatory movements were restricted during the intervention, they were not controlled during the assessment, which could have influenced the performance in the scales and tests. Second, although the physical therapist who assessed the participants' condition did not know the protocol, the therapists who administered and controlled the intervention were not blind. Third, we use gloves of uniform size, but each patient had a different hand size, and thus it might have been difficult for some patients to wear the gloves. Fourth, none of the patients who participated in the study had any experience with AR rehabilitation. Therefore, they could not compare our games with other similar games. Finally, the sample of the study (N=30) can be considered small, which may limit the degree to which the results can be extrapolated.

Future Studies

In future research, we will further improve our devices, which will involve developing an Android version to reduce installation costs, preparing different specifications of gloves for patients with different hand sizes, and increasing the variety and difficulty settings of the game. In addition, future work will include more engaging serious games to increase the variety of therapy solutions and adaptability to patient abilities so that a therapist or patient can match the degree of challenge necessary to keep the rehabilitation advancing. In the trial design, large-sample, controlled, follow-up clinical studies will be conducted in the future to verify the long-term efficacy of the system for home-based rehabilitation.

Conclusions

At the behavioral level, there was additional benefit received from CARS. Combined CARS and conventional OT rehabilitation was more effective in improving both upper limb function and cognitive function compared with conventional OT alone. The results of our study indicated that the proposed CARS can replace the one-on-one conventional OT delivered by an occupational therapist and that the system can be used as an assistant therapeutic tool in the hospital. In addition, CARS is convenient, low-cost, and user-friendly, which indicates that this system is also suitable for home-based rehabilitation. Future studies with a longer intervention time and a follow-up of patients for home-based rehabilitation are needed to explore the effectiveness of the system.

Acknowledgments

The authors wish to thank all the therapists and patients for their participation in this study. This work was supported by the Key National Research and Development Program (#2018YFC2002300), the National Nature Innovation Research Group Project (#82021002), the National Nature Integration Project (#91948302), and the National Natural Science Foundation of China (#51950410602).

Authors' Contributions

CL and XS conceived the study design. JH, CW, YZ, ZY, and SX recruited the patients and performed the assessments. CL and SC completed the data collection. XS developed the CARS system. CL and XS performed the data analysis and drafted the manuscript. JJ and PS were involved in the interpretation of the results and provided feedback on the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consolidated Standards of Reporting Trials (CONSORT) eHealth checklist (v 1.6.1).
[PDF File (Adobe PDF File), 359 KB - games_v9i4e30184_app1.pdf]

References

1. Campbell BCV, Khatri P. Stroke. Lancet 2020 Jul 11;396(10244):129-142. [doi: 10.1016/S0140-6736(20)31179-X] [Medline: 32653056]
2. GBD 2016 Stroke Collaborators. Global, regional, and national burden of stroke, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 May;18(5):439-458 [FREE Full text] [doi: 10.1016/S1474-4422(19)30034-1] [Medline: 30871944]
3. Coupar F, Pollock A, Rowe P, Weir C, Langhorne P. Predictors of upper limb recovery after stroke: a systematic review and meta-analysis. Clin Rehabil 2012 Apr;26(4):291-313. [doi: 10.1177/0269215511420305] [Medline: 22023891]
4. Nakayama H, Jørgensen HS, Raaschou HO, Olsen TS. Recovery of upper extremity function in stroke patients: the Copenhagen Stroke Study. Arch Phys Med Rehabil 1994 Apr;75(4):394-398. [Medline: 8172497]
5. Kalaria RN, Akinyemi R, Ihara M. Stroke injury, cognitive impairment and vascular dementia. Biochim Biophys Acta 2016 May;1862(5):915-925 [FREE Full text] [doi: 10.1016/j.bbadis.2016.01.015] [Medline: 26806700]

XSL•FO
RenderX

6. Nichols-Larsen DS, Clark PC, Zeringue A, Greenspan A, Blanton S. Factors influencing stroke survivors' quality of life during subacute recovery. Stroke 2005 Jul;36(7):1480-1484. [doi: 10.1161/01.STR.0000170706.13595.4f] [Medline: 15947263]
7. Sampaio-Baptista C, Sanders Z, Johansen-Berg H. Structural plasticity in adulthood with motor learning and stroke rehabilitation. Annu Rev Neurosci 2018 Jul 08;41:25-40. [doi: 10.1146/annurev-neuro-080317-062015] [Medline: 29490196]
8. Murphy TH, Corbett D. Plasticity during stroke recovery: from synapse to behaviour. Nat Rev Neurosci 2009 Dec;10(12):861-872. [doi: 10.1038/nrn2735] [Medline: 19888284]
9. Di Pino G, Pellegrino G, Assenza G, Capone F, Ferreri F, Formica D, et al. Modulation of brain plasticity in stroke: a novel model for neurorehabilitation. Nat Rev Neurol 2014 Oct;10(10):597-608. [doi: 10.1038/nrneurol.2014.162] [Medline: 25201238]
10. Israely S, Leisman G, Carmeli E. Improvement in arm and hand function after a stroke with task-oriented training. BMJ Case Rep 2017 Mar 17;2017:219-256 [FREE Full text] [doi: 10.1136/bcr-2017-219250] [Medline: 28314812]
11. Hubbard IJ, Parsons MW, Neilson C, Carey LM. Task-specific training: evidence for and translation to clinical practice. Occup Ther Int 2009;16(3-4):175-189 [FREE Full text] [doi: 10.1002/oti.275] [Medline: 19504501]
12. Winstein CJ, Wolf SL, Dromerick AW, Lane CJ, Nelsen MA, Lewthwaite R, Interdisciplinary Comprehensive Arm Rehabilitation Evaluation (ICARE) Investigative Team. Effect of a task-oriented rehabilitation program on upper extremity recovery following motor stroke: the ICARE randomized clinical trial. JAMA 2016 Feb 09;315(6):571-581 [FREE Full text] [doi: 10.1001/jama.2016.0276] [Medline: 26864411]
13. Jeon B, Kim W, Park E. Effect of task-oriented training for people with stroke: a meta-analysis focused on repetitive or circuit training. Top Stroke Rehabil 2015 Feb;22(1):34-43. [doi: 10.1179/1074935714Z.0000000035] [Medline: 25776119]
14. Legg LA, Lewis SR, Schofield-Robinson OJ, Drummond A, Langhorne P. Occupational therapy for adults with problems in activities of daily living after stroke. Cochrane Database Syst Rev 2017 Jul 19;7:CD003585 [FREE Full text] [doi: 10.1002/14651858.CD003585.pub3] [Medline: 28721691]
15. Hoffmann T, Bennett S, Koh C, McKenna KT. Occupational therapy for cognitive impairment in stroke patients. Cochrane Database Syst Rev 2010 Sep 08(9):CD006430 [FREE Full text] [doi: 10.1002/14651858.CD006430.pub2] [Medline: 20824849]
16. Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B. Recent advances in augmented reality. IEEE Comput. Grap. Appl 2001;21(6):34-47. [doi: 10.1109/38.963459]
17. Held JPO, Yu K, Pyles C, Veerbeek JM, Bork F, Heining S, et al. Augmented reality-based rehabilitation of gait impairments: case report. JMIR Mhealth Uhealth 2020 May 26;8(5):e17804 [FREE Full text] [doi: 10.2196/17804] [Medline: 32452815]
18. LaPiana N, Duong A, Lee A, Alschitz L, Silva RML, Early J, et al. Acceptability of a mobile phone-based augmented reality game for rehabilitation of patients with upper limb deficits from stroke: case study. JMIR Rehabil Assist Technol 2020 Sep 02;7(2):e17822 [FREE Full text] [doi: 10.2196/17822] [Medline: 32876580]
19. Mousavi Hondori H, Khademi M, Dodakian L, Cramer SC, Lopes CV. A Spatial Augmented Reality rehab system for post-stroke hand rehabilitation. Stud Health Technol Inform 2013;184:279-285. [Medline: 23400171]
20. Colomer C, Llorens R, Noé E, Alcañiz M. Effect of a mixed reality-based intervention on arm, hand, and finger function on chronic stroke. J Neuroeng Rehabil 2016 May 11;13(1):45 [FREE Full text] [doi: 10.1186/s12984-016-0153-6] [Medline: 27169462]
21. Al-Issa H, Regenbrecht H, Hale L. Augmented reality applications in rehabilitation to improve physical outcomes. Physical Therapy Reviews 2013 Nov 12;17(1):16-28. [doi: 10.1179/1743288x11y.0000000051]
22. Da Gama AEF, Chaves TM, Figueiredo LS, Baltar A, Meng M, Navab N, et al. MirrARbilitation: A clinically-related gesture recognition interactive tool for an AR rehabilitation system. Comput Methods Programs Biomed 2016 Oct;135:105-114 [FREE Full text] [doi: 10.1016/j.cmpb.2016.07.014] [Medline: 27586484]
23. Song X, Ding L, Zhao J, Jia J, Shull P. Cellphone augmented reality game-based rehabilitation for improving motor function and mental state after stroke. May 2009:1-4. [doi: 10.1109/bsn.2019.8771093]
24. Gladstone DJ, Danells CJ, Black SE. The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. Neurorehabil Neural Repair 2002 Sep;16(3):232-240. [doi: 10.1177/154596802401105171] [Medline: 12234086]
25. Saposnik G, Teasell R, Mamdani M, Hall J, McIlroy W, Cheung D, et al. Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation: a pilot randomized clinical trial and proof of principle. Stroke 2010 Jul;41(7):1477-1484 [FREE Full text] [doi: 10.1161/STROKEAHA.110.584979] [Medline: 20508185]
26. Sullivan KJ, Tilson JK, Cen SY, Rose DK, Hershberg J, Correa A, et al. Fugl-Meyer assessment of sensorimotor function after stroke: standardized training procedure for clinical practice and clinical trials. Stroke 2011 Feb;42(2):427-432. [doi: 10.1161/STROKEAHA.110.592766] [Medline: 21164120]
27. Singer B, Garcia-Vega J. The Fugl-Meyer Upper Extremity Scale. J Physiother 2017 Jan;63(1):53 [FREE Full text] [doi: 10.1016/j.jphys.2016.08.010] [Medline: 27964964]
28. 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]

XSL•FO
RenderX

29. Hernández ED, Galeano CP, Barbosa NE, Forero SM, Nordin, Sunnerhagen KS, et al. Intra- and inter-rater reliability of Fugl-Meyer Assessment of Upper Extremity in stroke. J Rehabil Med 2019 Oct 04;51(9):652-659 [FREE Full text] [doi: 10.2340/16501977-2590] [Medline: 31448807]
30. Duncan PW, Propst M, Nelson SG. Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. Phys Ther 1983 Oct;63(10):1606-1610. [doi: 10.1093/ptj/63.10.1606] [Medline: 6622535]
31. Yozbatiran N, Der-Yeghiaian L, Cramer SC. A standardized approach to performing the action research arm test. Neurorehabil Neural Repair 2008;22(1):78-90. [doi: 10.1177/1545968307305353] [Medline: 17704352]
32. Hsieh CL, Hsueh IP, Chiang FM, Lin PH. Inter-rater reliability and validity of the action research arm test in stroke patients. Age Ageing 1998 Mar;27(2):107-113. [doi: 10.1093/ageing/27.2.107] [Medline: 16296669]
33. Nordin, Alt Murphy M, Danielsson A. Intra-rater and inter-rater reliability at the item level of the Action Research Arm Test for patients with stroke. J Rehabil Med 2014 Sep;46(8):738-745 [FREE Full text] [doi: 10.2340/16501977-1831] [Medline: 24953235]
34. GONNELLA C. The manual muscle test in the patient's evaluation and program for treatment. Phys Ther Rev 1954 Jan;34(1):16-18. [doi: 10.1093/ptj/34.1.16] [Medline: 13120371]
35. Cuthbert SC, Goodheart GJ. On the reliability and validity of manual muscle testing: a literature review. Chiropr Osteopat 2007 Mar 06;15:4 [FREE Full text] [doi: 10.1186/1746-1340-15-4] [Medline: 17341308]
36. Safaz I, Yilmaz B, Yaşar E, Alaca R. Brunnstrom recovery stage and motricity index for the evaluation of upper extremity in stroke: analysis for correlation and responsiveness. Int J Rehabil Res 2009 Sep;32(3):228-231. [doi: 10.1097/MRR.0b013e32832a62ad] [Medline: 19339892]
37. Shah SK, Harasymiw SJ, Stahl PL. Stroke rehabilitation: outcome based on Brunnstrom recovery stages. The Occupational Therapy Journal of Research 2016 Aug 24;6(6):365-376. [doi: 10.1177/153944928600600604]
38. Katzman R, Zhang MY, Ouang-Ya-Qu, Wang ZY, Liu WT, Yu E, et al. A Chinese version of the Mini-Mental State Examination; impact of illiteracy in a Shanghai dementia survey. J Clin Epidemiol 1988;41(10):971-978. [doi: 10.1016/0895-4356(88)90034-0] [Medline: 3193141]
39. Li H, Jia J, Yang Z. Mini-Mental State Examination in elderly Chinese: a population-based normative study. J Alzheimers Dis 2016 May 07;53(2):487-496. [doi: 10.3233/JAD-160119] [Medline: 27163822]
40. Boban M, Malojčić B, Mimica N, Vuković S, Zrilić I, Hof PR, et al. The reliability and validity of the mini-mental state examination in the elderly Croatian population. Dement Geriatr Cogn Disord 2012;33(6):385-392 [FREE Full text] [doi: 10.1159/000339596] [Medline: 22814030]
41. Rakusa M, Granda G, Kogoj A, Mlakar J, Vodusek DB. Mini-Mental State Examination: standardization and validation for the elderly Slovenian population. Eur J Neurol 2006 Feb;13(2):141-145. [doi: 10.1111/j.1468-1331.2006.01185.x] [Medline: 16490044]
42. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989;42(8):703-709. [Medline: 2760661]
43. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
44. Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988;10(2):61-63. [Medline: 3403500]
45. Quinn TJ, Langhorne P, Stott DJ. Barthel index for stroke trials: development, properties, and application. Stroke 2011 Apr;42(4):1146-1151. [doi: 10.1161/STROKEAHA.110.598540] [Medline: 21372310]
46. Gil-Gómez JA, Manzano-Hernández P, Albiol-Pérez S, Aula-Valero C, Gil-Gómez H, Lozano-Quilis J. USEQ: a short questionnaire for satisfaction evaluation of virtual rehabilitation systems. Sensors (Basel) 2017 Jul 07;17(7):1-12 [FREE Full text] [doi: 10.3390/s17071589] [Medline: 28686174]
47. Mubin O, Alnajjar F, Al Mahmud A, Jishtu N, Alsinglawi B. Exploring serious games for stroke rehabilitation: a scoping review. Disabil Rehabil Assist Technol 2020 Jun 08:1-7. [doi: 10.1080/17483107.2020.1768309] [Medline: 32508187]
48. Assis GAD, Corrêa AGD, Martins MBR, Pedrozo WG, Lopes RDD. An augmented reality system for upper-limb post-stroke motor rehabilitation: a feasibility study. Disabil Rehabil Assist Technol 2016 Aug;11(6):521-528. [doi: 10.3109/17483107.2014.979330] [Medline: 25367103]
49. Hoermann S, Ferreira Dos Santos L, Morkisch N, Jettkowski K, Sillis M, Devan H, et al. Computerised mirror therapy with Augmented Reflection Technology for early stroke rehabilitation: clinical feasibility and integration as an adjunct therapy. Disabil Rehabil 2017 Jul;39(15):1503-1514. [doi: 10.1080/09638288.2017.1291765] [Medline: 28478736]
50. Kaneko F, Shindo K, Yoneta M, Okawada M, Akaboshi K, Liu M. A case series clinical trial of a novel approach using augmented reality that inspires self-body cognition in patients with stroke: effects on motor function and resting-state brain functional connectivity. Front Syst Neurosci 2019;13:76 [FREE Full text] [doi: 10.3389/fnsys.2019.00076] [Medline: 31920571]
51. Trojan J, Diers M, Fuchs X, Bach F, Bekrater-Bodmann R, Foell J, et al. An augmented reality home-training system based on the mirror training and imagery approach. Behav Res Methods 2014 Sep;46(3):634-640 [FREE Full text] [doi: 10.3758/s13428-013-0412-4] [Medline: 24338625]

XSL•FO
RenderX

52. Chinthammit W, Merritt T, Pedersen S, Williams A, Visentin D, Rowe R, et al. Ghostman: augmented reality application for telerehabilitation and remote instruction of a novel motor skill. Biomed Res Int 2014;2014:646347 [FREE Full text] [doi: 10.1155/2014/646347] [Medline: 24829910]
53. Gorman C, Gustafsson L. The use of augmented reality for rehabilitation after stroke: a narrative review. Disabil Rehabil Assist Technol 2020 Jul 14:1-9. [doi: 10.1080/17483107.2020.1791264] [Medline: 32663112]
54. Domínguez-Téllez P, Moral-Muñoz JA, Salazar A, Casado-Fernández E, Lucena-Antón D. Game-based virtual reality interventions to improve upper limb motor function and quality of life after stroke: systematic review and meta-analysis. Games Health J 2020 Feb;9(1):1-10. [doi: 10.1089/g4h.2019.0043] [Medline: 32027185]
55. Sitaram R, Ros T, Stoeckel L, Haller S, Scharnowski F, Lewis-Peacock J, et al. Closed-loop brain training: the science of neurofeedback. Nat Rev Neurosci 2017 Feb;18(2):86-100. [doi: 10.1038/nrn.2016.164] [Medline: 28003656]

Abbreviations

ADL: activity of daily living
ARAT: Action Research Arm Test
AVS: Add VS Sub
BI: Barthel index
BS: Brunnstrom stage
CARS: cellphone augmented reality system
FMA-UE: Fugl-Meyer Assessment Upper Extremity
MMSE: Mini-Mental State Examination
MMT: manual muscle test
SG: Stroop Game

Edited by N Zary; submitted 04.05.21; peer-reviewed by J He, R Ciorap, Z Xu, B Xia; comments to author 28.05.21; revised version received 15.06.21; accepted 10.09.21; published 23.11.21.

Please cite as:
Li C, Song X, Chen S, Wang C, He J, Zhang Y, Xu S, Yan Z, Jia J, Shull P
Long-term Effectiveness and Adoption of a Cellphone Augmented Reality System on Patients with Stroke: Randomized Controlled Trial
JMIR Serious Games 2021;9(4):e30184
URL: https://games.jmir.org/2021/4/e30184
doi:10.2196/30184
PMID:34817390

©Chong Li, Xinyu Song, Shugeng Chen, Chuankai Wang, Jieying He, Yongli Zhang, Shuo Xu, Zhijie Yan, Jie Jia, Peter Shull. Originally published in JMIR Serious Games (https://games.jmir.org), 23.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Augmented Reality in Physical Therapy: Systematic Review and Meta-analysis

Maria Jesus Vinolo Gil[1,2,3], PhD; Gloria Gonzalez-Medina[1,3], PhD; David Lucena-Anton[1], PhD; Veronica Perez-Cabezas[1], PhD; María Del Carmen Ruiz-Molinero[1,3], PhD; Rocío Martín-Valero[4], PhD

[1]Department of Nursing and Physical Therapy, University of Cadiz, Cadiz, Spain

[2]Clinical Management Unit Rehabilitation Intercentre-Interlevel, University Hospitals of Puerto Real and Cadiz, Cadiz Bay-La Janda Health District, Cadiz, Spain

[3]Institute for Biomedical Research and Innovation of Cádiz, Cadiz, Spain

[4]Department of Physical Therapy, University of Malaga, Malaga, Spain

Corresponding Author:
Gloria Gonzalez-Medina, PhD
Department of Nursing and Physical Therapy
University of Cadiz
Avda Ana de Viya 52
Cadiz, 11009
Spain
Phone: 34 670609656
Email: gloriagonzalez.medina@uca.es

Abstract

Background: Augmented reality (AR) is a rapidly expanding technology; it comprises the generation of new images from digital information in the real physical environment of a person, which simulates an environment where the artificial and real are mixed. The use of AR in physiotherapy has shown benefits in certain areas of patient health. However, these benefits have not been studied as a whole.

Objective: This study aims to ascertain the current scientific evidence on AR therapy as a complement to physiotherapy and to determine the areas in which it has been used the most and which variables and methods have been most effective.

Methods: A systematic review registered in PROSPERO (International Prospective Register of Systematic Reviews) was conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta - Analyses) recommendations. The search was conducted from July to August 2021 in the PubMed, PEDro, Web of Science, Scopus, and Cochrane Library scientific databases using the keywords *augmented reality*, *physiotherapy*, *physical therapy*, *exercise therapy*, *rehabilitation*, *physical medicine*, *fitness*, and *occupational therapy*. The methodological quality was evaluated using the PEDro scale and the Scottish Intercollegiate Guidelines Network scale to determine the degree of recommendation. The Cochrane Collaboration tool was used to evaluate the risk of bias.

Results: In total, 11 articles were included in the systematic review. Of the 11 articles, 4 (36%) contributed information to the meta-analysis. Overall, 64% (7/11) obtained a good level of evidence, and most had a B degree of recommendation of evidence. A total of 308 participants were analyzed. Favorable results were found for the Berg Balance Scale (standardized mean change 0.473, 95% CI −0.0877 to 1.0338; z=1.65; P=.10) and the Timed Up and Go test (standardized mean change −1.211, 95% CI −3.2005 to 0.7768; z=−1.194; P=.23).

Conclusions: AR, in combination with conventional therapy, has been used for the treatment of balance and fall prevention in geriatrics, lower and upper limb functionality in stroke, pain in phantom pain syndrome, and turning in place in patients with Parkinson disease with freezing of gait. AR is effective for the improvement of balance; however, given the small size of the samples and the high heterogeneity of the studies, the results were not conclusive. Future studies using larger sample sizes and with greater homogeneity in terms of the devices used and the frequency and intensity of the interventions are needed.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD42020180766; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=180766

(JMIR Serious Games 2021;9(4):e30985) doi:10.2196/30985

XSL•FO
RenderX

KEYWORDS

augmented reality; physical therapy; rehabilitation; functionality

Introduction

Background

New technologies are rapidly emerging in our society to streamline, optimize, and perfect some of the activities we perform in our day-to-day lives [1]. Among them is augmented reality (AR), which comprises generating new images from digital information in the real physical environment of a person, simulating an environment where the artificial and real are mixed [2]. AR must be differentiated from virtual reality (VR), in which additional data such as sound, text, or video are introduced, giving rise to multimedia virtual environments. AR is derived from VR but blends these virtual environments with real ones, enhancing the interaction with real life [3].

AR is currently being used in different fields such as advertising [4], psychology [5], medicine [6,7], and physiotherapy [8]. In physiotherapy, it has been developed mainly for motor and cognitive rehabilitation, which is considered a new method of intervention. AR can be used as a working tool and to complement the treatment conducted by the physiotherapist, as it generates safe environments that are similar to the patient's real environment [9]. Rehabilitation using AR has shown better results than repetitive movements practiced alone as AR allows better orientation of the exercises toward objectives with greater patient motivation and is enjoyable to use [10].

AR technologies have significant advantages: they provide new experiences to patients during physiotherapy sessions, increasing engagement and improving physical outcomes [11]; they can create interesting opportunities to provide low-cost physiotherapy at home [12,13]; and the physiotherapist can perform and evaluate different outcomes using these tools with data analysis [14]. Although lack of technological maturity and access to devices are their weaknesses [15], various types of interfaces are emerging to ensure user interaction with the AR rehabilitation environment, including wearable smart sensors, sensors embedded in the environment, and mobile devices that improve accessibility to this type of technology [16]. Despite these possible benefits, there are few studies on AR used in physiotherapy, unlike VR, which has been studied in more pathologies, mostly of the neurological type, such as stroke [17-19], cerebral palsy [20], multiple sclerosis [21,22], Parkinson disease [23,24], spinal cord injury [25,26], and chronic pain [27].

Of the few investigations that have been conducted on AR, most were performed on healthy people with the aim of determining strategies that could later be used in the clinic [28]. Interest in studying AR has also grown in certain areas, such as the kinematic analysis of gait parameters [28,29], the functionality of the upper limb [30,31], or the early diagnosis of breast cancer–related lymphedema [32]. Positive results on balance and mobility have also been achieved when using dance with AR devices, with high adherence [33]. However, no improvement was found in the use of AR for the performance of daily living tasks in patients with Alzheimer disease [4].

Recently, a protocol for interactive AR-based telerehabilitation in patients with adhesive capsulitis was published [34] and another was published about people with hereditary spastic paraplegia, in which gait adaptability training was treated with a treadmill equipped with AR [35]; however, their results have not yet been published. In 2010, a review was conducted [36] in which most of the AR studies analyzed were in the prototype development phase and not yet ready for general practice, although they did show promising results.

Objective

In the given context and taking into account all the advantages that the use of this kind of tool could have in physiotherapy, this review aims to determine how progress has been made in this regard, with the objective of ascertaining the current scientific evidence on AR therapy as a complement to physiotherapy, determining in which areas it has been used the most and which variables and methods have been most effective.

Methods

A systematic review and meta-analysis was conducted and registered in PROSPERO (International Prospective Register of Systematic Reviews; CRD42020180766) using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta - Analyses) guidelines [37].

Search Strategy

A search of scientific evidence published from 2011 to August 2021 was conducted between July and August 2021 in the following scientific databases: PubMed, PEDro, Web of Science, Scopus, and Cochrane Library. In addition, gray literature (the TESEO database of doctoral theses in Spain, OpenGrey, and Grey Literature Database) and AR conference proceedings were searched. The keywords *augmented reality*, *physiotherapy*, *physical therapy*, *exercise therapy*, *rehabilitation*, *physical medicine*, *fitness*, and *occupational therapy* were used, combining them by means of the Boolean operators AND and OR in the different searches in English or Spanish.

Criteria for Considering Studies

The criterion that was taken into account for selecting the articles was clinical trials published in indexed scientific databases. The selected intervention was AR used with patients aged >18 years with some pathology of the musculoskeletal system of neurological or physical origin that was subsidiary to improvement in any physical measure analyzed in an objective and standardized way. Duplicate studies, qualitative trials, case reports, single-subject studies, reviews, meta-analyses, studies conducted on healthy individuals, and studies using VR were excluded.

Study Selection and Data Extraction Process

After performing the search, potentially relevant articles were identified after reading the title and abstract and eliminating duplicates. All studies identified in the searches were assessed for inclusion by 2 independent reviewers (MJVG and GGM).

XSL•FO
RenderX

Any disagreements were resolved through discussions to reach a consensus. The following information was extracted from each included article: authors, year of publication, study population, type of intervention, number of participants, mean age, frequency of sessions per week, time of each session, total duration of the intervention, outcome measures, measurement instrument, and results obtained.

Assessment of the Methodological Quality

To assess the quality of the trials used for the review, we used the PEDro scale [38], which comprises 11 items related to the domains of selection, performance, attribution bases, and information. A study with a PEDro score of ≥6 was considered as evidence level 1 (6-8: good and 9-10: excellent), and a study with a score of ≤5 was considered as evidence level 2 (4-5: acceptable and <4: poor) [39]. The recommendation grades of the different studies were presented using the Scottish Intercollegiate Guidelines Network scale [40].

Risk of Bias Analysis

The risk of bias was calculated for each study using the Cochrane Collaboration tool [41], referring to the following types of bias: selection bias, performance bias, detection bias, attrition bias, reporting bias, and other bias. The risk of bias and study quality were calculated by 2 reviewers. In cases of doubt, the final decision was determined through discussion by including a third reviewer.

Statistical Analysis

The effect size measure was the difference in the standardized mean change with raw standardization (SMCR) between the intervention (AR) and control groups [42,43] for 2 dependent end points: standardized mean change of the Berg Balance Scale (BBS) and the Timed Up and Go (TUG) test, with improvement after treatment indicated by positive values in the BBS and negative values in the TUG. The difference in SMCR was estimated in such a way that a greater difference in the intervention group was indicated by positive values in the BBS and negative values in the TUG. Standardized mean differences, sampling variances, and covariances were estimated according to Gleser and Olkin [44]. A multivariate random effects model with restricted maximum likelihood estimation was used, allowing for a different effect depending on the outcome and adding random effects to each outcome within each study. The goodness of fit was evaluated using sensitivity analyses [45] and likelihood profile plots. Publication bias was evaluated using contour-enhanced funnel plots [46]. The analyses were performed using the metafor package (GNU General Public License Version 2) [47] of the R software (R Foundation for Statistical Computing) [48].

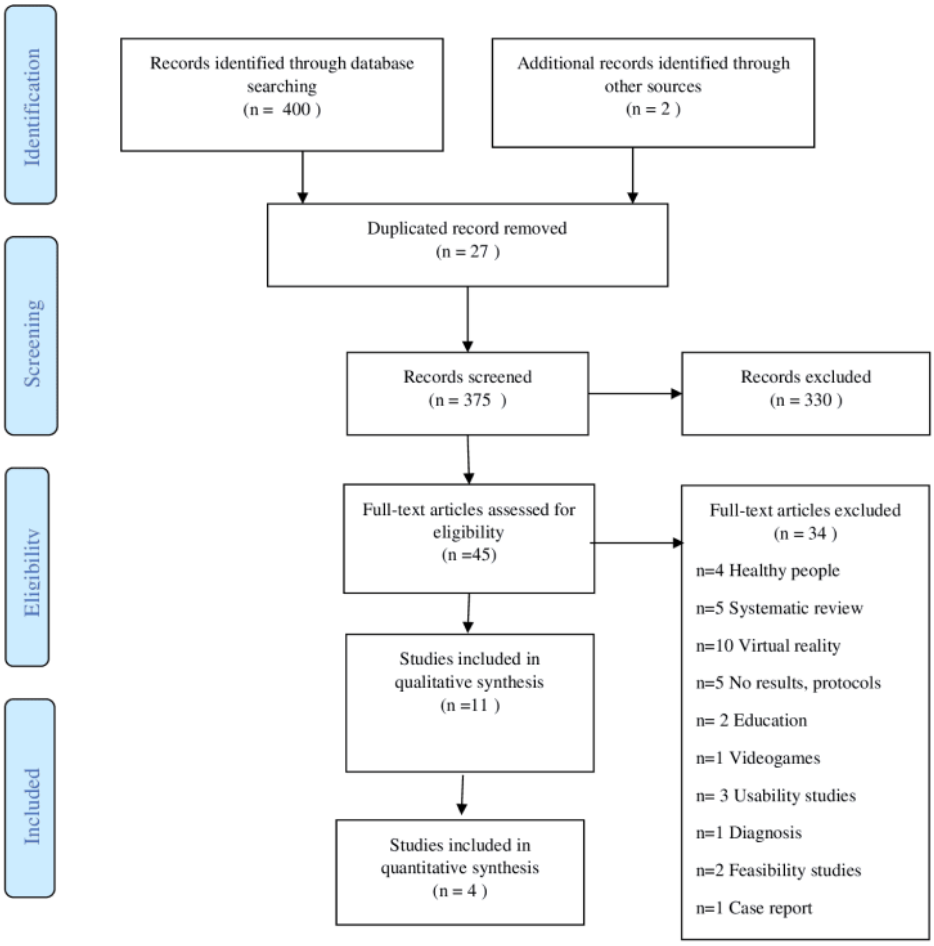
Results

Selection of Studies

The entire selection process during the corresponding phases is detailed in Figure 1.

XSL•FO
RenderX

Figure 1. Flow diagram. Graphical representation of the process of search and selection of studies.



Evaluation Outcomes

The sample size was variable, with the largest sample (75 patients) being in the study by Rothgangel et al [49] and the smallest (10 people) being in the study by Jung et al [50]. The included studies contained information on a total of 308 patients, of whom 89 (28.9%) had a stroke [50-52], 89 (28.9%) had amputations [49,53], 114 (37%) were geriatric patients [54-57], and 16 (5.2%) had Parkinson disease [58]. In terms of the age of the participants, the highest average was 76.4 in the study by Lee et al [55], and the lowest was 47.4 in the study by Kim et al [52]. It should be noted that of the 11 studies, 6 (55%) analyzed the effects of AR on the lower limb [49,50,52,54-56,58,59], and 2 (18%) did so for the upper limb [51,53]. The main characteristics of these studies are listed in Table 1.

XSL•FO
RenderX

Table 1. Main characteristics of the study interventions.

Study	Sample	Age (years), mean (SD)	Study population	Intervention	Frequency of treatment (times/week)	Session time (minutes)	Total time of the intervention	Measurement instrument	Outcome	Results
Colomer et al [51]	30	58.3 (10.1)	Stroke	IG[a]: reverse study—A-B-A; A: conventional physical therapy program; B: AR[b]	3-5	45	12 weeks	Wolf motor function test, box and block test, 9-hole plug test, and Intrinsic Motivation Inventory	Elbow flexion and extension, wrist flexion and extension, finger flexion and extension, and grabbing different objects	• Significant improvement in arm function and finger dexterity • High levels of interest, motivation, and enjoyment
Lee et al [52]	21; CG[c]: 11; IG: 11	Not specified	Stroke	CG: general physical therapy program; IG: general physical therapy program+AR-based postural control program	CG: 5; IG: 3	30	CG: 4 weeks; IG: 8 weeks	TUG[d], BBS[e], spatial-temporal parameters (GAITRite), and dynamometer	Gait, balance, and muscle strength	• Improvements in walking speed, balance and cadence, stride length, and stride length of paretic and nonparetic sides
Jung et al [50]	10; CG: 5; IG: 15	CG: 57.80 (10.23); IG: 58.40 (8.26)	Stroke	CG: FES[f]; IG: AR-based FES	3	20	4 weeks	Surface EMG[g] machine, electronic goniometer, and manual muscle tester	Muscle activation, ankle range of motion, and muscle strength	• Improved muscle activation in GCM[h]and TSA[i] • Improved muscle strength in dorsiflexion and plantar flexion
Kim et al [59]	28; group 1: 9; group 2: 10; group 3: 9	Group 1: 47.4 (8.4); group 2: 51.5 (12.9); group 3: 49.1 (11)	Stroke	Group 1: treadmill walking with EFS[j] and AR therapy; group 2: treadmill walking with EFS therapy; group 3: gait on treadmill walking	3	20	8 weeks	BBS and TUG	Muscle strength, balance, and gait	• Muscle strength increased significantly in groups 1 and 2. • Balance and gait showed significant improvements in all groups.

XSL•FO
RenderX

Study	Sample	Age (years), mean (SD)	Study population	Intervention	Frequency of treatment (times/week)	Session time (minutes)	Total time of the intervention	Measurement instrument	Outcome	Results
Rothgangel et al [49]	75; group 1: 25; group 2: 26; CG: 24	61.1 (14.2)	Lower limb amputation	Group 1: mirror therapy+ AR teleprocessing; group 2: mirror therapy+self-administered mirror therapy; CG: sensory-motor exercises	Not specified	30	Group 1: 10 weeks; group 2: 10 weeks; CG: 10 weeks	NRS[k] inventory of neuropathic pain symptoms, Patient-specific Functional Scale, EuroQol 5 Dimensions, Overall Perceived Effect Scale, and pain Self-Efficacy Questionnaire	Intensity, frequency, and duration of phantom pain	• AR had no additional effects compared with the other groups.
Ortiz-Catalán et al [53]	14	50.3 (13.9)	Upper limb amputation	IG: motor execution in AR, game series; use of a virtual member in different tasks	2	120	12 sessions	NRS pain rating index, Weighted Scale of Pain Distribution, and study-specific frequency scale for each session	Intensity, frequency, duration, and quality of phantom limb pain (upper)	• Clinical and statistical improvements in all phantom limb pain metrics
Lee et al [55]	30; group 1: 10; group 2: 10; group 3: 10	Women; group 1: 72.6 (2.67); group 2: 75.8 (5.47); group 3: 76.4 (5.54)	Older adults	Group 1: AR+Otago[l]; group 2: yoga; group 3: exercises at home	3	60	12 weeks	Strength of knee flexor, extensor, and ankle flexor muscles; footprint; static and dynamic load distribution; and MFS[m]	Muscle strength, balance, and risk of falling	• Group 1, group 2, and group 3 had improved strength. • The AR group improved significantly in balance and in the fall scale.
Yoo et al [56]	21; group 1: 10; group 2: 11	Women; group 1: 72.9 (3.41); group 2: 75.6 (5.57)	Older adults	Group 1: AR +Otago exercises; group 2: Otago exercises	3	50	12 weeks	Gait parameters, BBS, and FES-I[n]	Gait functionality, balance, and risk of falling	• Group 1 had significant differences in gait and balance parameters greater than group 2. • Group 1 had significant differences in fall prevention.
Ku et al [54]	36; CG: 18; IG: 18	CG: 65 (4.77); IG: 64.7 (7.27)	Older adults	CG: physical fitness program; IG: training with 3D-AR system	3	30	4 weeks		Lower limb balance and lower limb mobility	

XSL•FO
RenderX

Study	Sample	Age (years), mean (SD)	Study population	Intervention	Frequency of treatment (times/week)	Session time (minutes)	Total time of the intervention	Measurement instrument	Outcome	Results
								BBS, TUG, FAC[o], MBI[p], Fugl-Meyer lower limb subscale, Fugl-Meyer motor coordination, Fugl-Meyer motor score, and balance (Tetrax posturography)		• Improved stability index with interaction between BBS and TUG • Improvement in fall risk • Improvement of the posturographic index • Improved weight distribution index
Jeon et al [57]	27; CG: 13; IG: 14	CG: 72.71 (3.64); IG: 72.77 (3.79)	Older adults	CG: no exercise; IG: AR-based exercise	3	30	12 weeks	Stadiometer, BIA[q], hand dynamometer, SFT[r], and ESE[s]	Muscle mass, muscle function, physical performance, and exercise self-efficacy	• Improved ASM[t], SMI[u], gait speed, SFT in chair stand test, 2MST[v], and self-efficacy
Janssen et al [58]	16	Median: 69	Parkinson disease	Experimental condition: series of 180° turns with AR visual cues displayed through a HoloLens; 2 control conditions: with auditory cues and without any cues	1 session	N/A[w]	1 session	PTF[x], mean number, and duration of FOG[y] episodes; maximum medial COM[z] deviation, maximum head-pelvis separation, and time to maximum head-pelvis separation; cadence, peak angular velocity, stride time, coefficient of variation, step height, and turn time	FOG parameters, axial kinematics, scaling, and timing of turning	

XSL•FO
RenderX

Study	Sample	Age (years), mean (SD)	Study population	Intervention	Frequency of treatment (times/week)	Session time (minutes)	Total time of the intervention	Measurement instrument	Outcome	Results
										• AR visual cues did not reduce the PTF (*P*=.73) or the number (*P*=.73) and duration (*P*=.78) of FOG episodes, the peak angular velocity (visual vs uncued, *P*=.03; visual vs auditory, *P*=.02) and step height, and they increased the step height coefficient of variation and time to maximum head-pelvis separation. • All FOG parameters were higher with AR visual cues than with auditory cues (PTF, *P*=.01; number, *P*=.02; and duration, *P*=.007 of FOG episodes).

[a]IG: intervention group.
[b]AR: augmented reality.
[c]CG: control group.
[d]TUG: Timed Up and Go.
[e]BBS: Berg Balance Scale.
[f]FES: functional electric stimulation.
[g]EMG: electromyogram.
[h]GCM: medial and lateral gastrocnemius.
[i]TSA: tibialis anterior.
[j]EFS: electrical functional stimulation.
[k]NRS: Numerical Pain Rating Scale.
[l]Otago: Strength and Balance Training Program for Seniors.
[m]MFS: Morse Fall Scale.
[n]FES-I: Short Falls Efficacy Scale–International.
[o]FAC: functional ambulation category.
[p]MBI: Modified Barthel Index.

XSL•FO
RenderX

[q]BIA: Bioelectrical Impedance Analysis (Inbody 720, Biospace).

[r]SFT: senior fitness test.

[s]ESE: exercise self-efficacy.

[t]ASM: appendicular skeletal muscle mass.

[u]SMI: skeletal muscle index.

[v]2MST: 2-minute step test.

[w]N/A: not applicable.

[x]PTF: percentage time frozen.

[y]FOG: freezing of gait.

[z]COM: center of mass.

The AR systems used varied widely: projectors connected to computers with webcams where images were shown [51], virtual upper limbs [53], training videos [50,55], teletreatment using AR with tablets [49], projections with AR on treadmills [55,56,59] or on the ground [55,56], a head-mounted AR device used for holographic display of AR visual cues [58], and an AR-based exercise rehabilitation system [57] or a newer system such as the 3D-AR system, in which the participant's body movement was tracked, creating an AR environment that generated real images captured in videos with virtual images [54].

The intervention time ranged from 20 minutes [59] to 2 hours [53], although the most repeated chosen time was 30 minutes [49,52,54]. The most used frequency in the studies was 3 times per week [51,52,54,55,59] and 12 sessions [52-55].

The most widely used measurement scale in the selected studies was the BBS, used both for stroke [52,59] and in older adults [54,55]. Scales were also used to assess falls; in 1 trial, the Short Falls Efficacy Scale–International was used [56], and the Morse Fall Scale was used in another [55]. In both investigations, the target population was older adults. Another scale repeated in 3 of the selected articles was the TUG [52,54,59], which was applied to people with stroke and older adults. Regarding the studied population with amputations, many scales were used to assess pain; however, the only one in which both studies coincided was the Numerical Pain Rating Scale [49,53].

The results found regarding the interventions conducted in the field of physiotherapy were diverse in the different plots. In stroke, intensive and repetitive task-oriented exercises were used for upper limb functionality [51], postural control exercises [52], functional electrical stimulation [50], and treadmill [59]. Mirror therapy and sensorimotor exercises were used in the treatment of phantom limb pain [49]. In Parkinson disease, turns around the patient's axis were used [58]. Finally, in geriatrics, exercises from the Otago protocol were used [55,56], as well as yoga [55] and physical conditioning with strengthening and balance training [54,57].

Methodological Quality of the Included Studies

The results of the methodological quality assessment can be found in Table 2. After assessing the studies using the PEDro scale, it stands out that, of the 11 studies included in the review, 7 (64%) had high methodological quality (≥6 points), and the rest were acceptable. The scores obtained and the detailed characteristics of each study are shown in Table 2. Regarding the Scottish Intercollegiate Guidelines Network scale, most studies had a grade of B (Table 3).

XSL•FO
RenderX

Table 2. Evaluation of the methodological quality according to the PEDro scale[a].

Study	Item 1[b]	Item 2[c]	Item 3[d]	Item 4[e]	Item 5[f]	Item 6[g]	Item 7[h]	Item 8[i]	Item 9[j]	Item 10[k]	Item 11[l]	Score (out of 10)
Colomer et al [51]	1	0	0	1	0	0	0	1	1	0	1	4
Lee et al [52]	0	1	0	1	0	0	1	1	1	1	1	7
Jung et al [50]	1	1	1	1	0	0	0	1	0	1	0	6
Kim et al [59]	1	1	0	1	0	0	0	0	0	1	1	4
Rothgangel et al [49]	1	1	1	1	0	0	1	1	1	1	1	8
Ortiz-Catalán et al [53]	1	0	0	1	1	1	1	1	1	0	1	7
Lee et al [55]	1	1	0	1	0	0	0	0	0	1	1	4
Yoo et al [56]	0	1	0	1	0	0	0	1	0	1	1	5
Ku et al [54]	1	1	1	1	0	0	1	1	0	1	1	7
Jeon et al [57]	1	1	1	1	0	0	0	1	0	1	1	6
Janssen et al [58]	1	0	1	1	0	0	1	1	0	1	1	7

[a]1=yes and 0=no.
[b]Choice criteria specified; did not add up in the final computation.
[c]Random assignment.
[d]Covert assignment.
[e]Baseline similarity.
[f]Subject blinding.
[g]Therapist blinding.
[h]Evaluator blinding.
[i]Greater than 85% follow-up for at least 1 key outcome.
[j]Intention-to-treat analysis.
[k]Statistical comparison between groups for at least 1 key outcome.
[l]Point measures and variability for at least 1 key outcome.

Table 3. Grades of recommendation according to the Scottish Intercollegiate Guidelines Network scale.

Study	Grade of recommendation
Colomer et al [51]	B
Lee et al [52]	A
Jung et al [50]	B
Kim et al [59]	B
Rothgangel et al [49]	A
Ortiz-Catalán et al [53]	B
Lee et al [55]	B
Yoo et al [56]	B
Ku et al [54]	A
Jeon et al [57]	B
Janssen et al [58]	B

Risk of Bias

The results of the risk of bias can be observed in Table 4. It should be noted that 36% (4/11) of articles presented a low risk of selection bias, as they were randomized [49,52-54], although only 25% (1/4) of them also presented allocation concealment [54]. With respect to performance bias, none were at low risk. Regarding detection bias, 45% (5/11) of the articles included in the review were at low risk. In relation to dissertation bias, all of them were at low risk.

XSL•FO
RenderX

Table 4. Risk of bias.

Author	Criteria (risk)						
	1[a]	2[b]	3[c]	4[d]	5[e]	6[f]	7[g]
Colomer et al [51]	High	High	High	High	Low	Unclear	Unclear
Lee et al [52]	Low	High	High	Low	Low	Unclear	Unclear
Jung et al [50]	Low	Unclear	High	High	Low	Unclear	Unclear
Kim et al [59]	Unclear	High	High	High	Low	Unclear	Unclear
Rothgangel et al [49]	Low	High	High	Low	Low	Unclear	Unclear
Ortiz-Catalán et al [53]	Low	High	High	Low	Low	Unclear	Unclear
Lee et al [55]	Unclear	High	High	Unclear	Low	Unclear	Unclear
Yoo et al [56]	Unclear	High	High	Low	Low	Unclear	Unclear
Ku et al [54]	Low	Low	High	Low	Low	Unclear	Unclear
Jeon et al [57]	Low	Unclear	High	High	Low	Unclear	Unclear
Janssen et al [58]	High	High	High	High	Low	Unclear	Unclear

[a]Random sequence generation (selection bias).
[b]Allocation concealment (selection bias).
[c]Blinding of participants and personnel (performance bias).
[d]Blinding of outcome assessment (detection bias).
[e]Incomplete outcome data (attrition bias).
[f]Selective reporting (reporting bias).
[g]Other bias.

Study Groups Included in the Meta-analysis

In this meta-analysis, 36% (4/11) of studies were selected to evaluate the differences in mean changes in BBS and TUG scores. The power for detecting differences was low because of the reduced number of studies and small sample sizes. The data used for the meta-analysis are shown in Table 5. Descriptive data extracted from the selected studies are included in Multimedia Appendix 1 [49,51-56,59]. The (pooled) difference in standardized mean change was 0.473 (95% CI −0.0877 to 1.0338; z=1.65; P=.10) for the BBS and −1.211 (95% CI −3.2005 to 0.7768; z=−1.194; P=.23) for the TUG, both differences favoring the intervention group, although the null hypothesis cannot be rejected. The forest plot (Figure 2 [49-59]) showing the individual and pooled SMCR (with 95% CI), weights, and sample sizes of each study is shown in Figure 3 [49-59]. Substantial heterogeneity (test for residual heterogeneity: Q_5=45.82; P<.001) was present among the studies, with estimated variance components of 0.148 (95% CI 0.0001-2.2727; τ2BBS) and 3.098 (95% CI 0.5818-36.1115; τ2TUG). No identifiability problems for the variance components were found (Figure 4). The 2 outcomes showed a very high correlation (ρ=−0.99). The individual effect size was significant for both outcomes in the study by Ku et al [54], the study with the greater sample size. Nonetheless, sensitivity analysis showed that this study had higher standardized residuals and Cook distance values for the outcome TUG. The contour-enhanced funnel plot (Figure 5) seems to indicate the absence of publication bias (results should be considered with caution because of the small sample size).

XSL•FO
RenderX

Table 5. Data used for the meta-analysis.

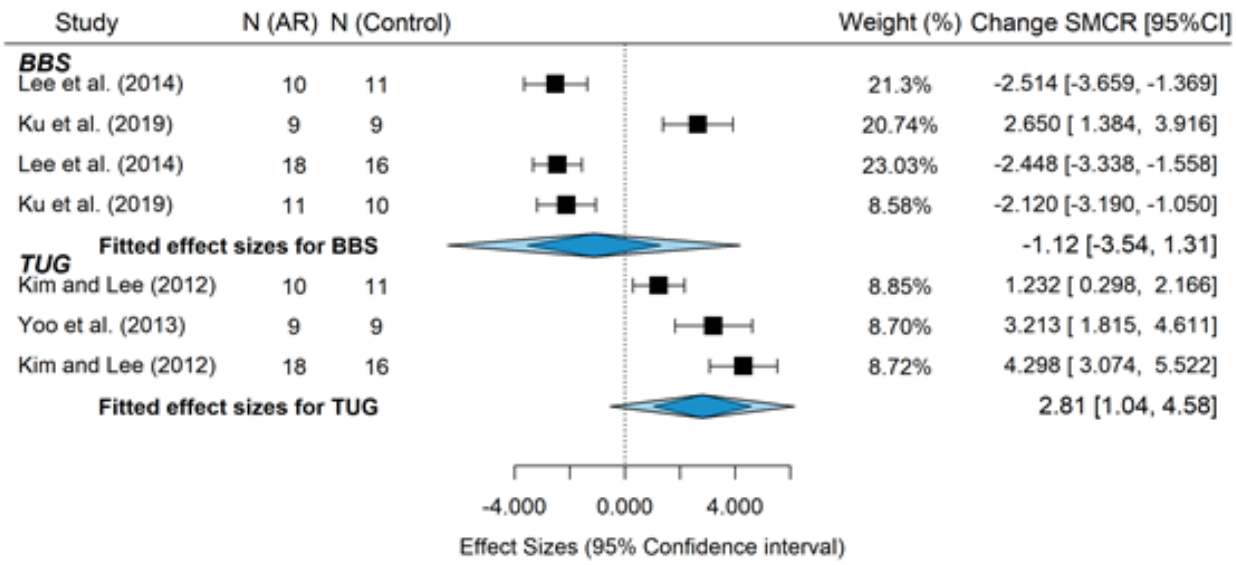
Study and outcome	SMC[a]		Correlations between pre- and postintervention means		SD prediction interval	Sample sizes		Pooled correlations between the 2 outcomes	Differences in SMC	Sampling variance	Sampling covariance
	Control	Intervention	Control	Intervention		Control	Intervention				
Lee et al [52]											
BBS[b]	0.2752	0.6691	0.5347075	0.8318219	1.3679	10	11	0.6482	0.2879	0.1929	0.1235
TUG[c]	−0.1982	−0.3870	0.4264936	0.7023536	2.3817	10	11	0.6482	−0.0792	0.1911	0.1235
Kim and Lee [59]											
BBS	1.3922	0.9171	0.9558409	0.8978624	1.7163	9	9	0.6482	−0.2768	0.2244	0.1452
TUG	−0.5768	−1.1695	0.7852951	0.9161890	1.6490	9	9	0.6482	−0.3594	0.2258	0.1452
Ku et al [54]											
BBS	0.3472	0.5847	0.7299132	0.9078490	0.2180	18	16	0.7069	1.0895	0.1355	0.0548
TUG	−0.2495	−0.6736	0.5644717	0.9050366	0.1187	18	16	0.7069	−3.5745	0.3060	0.0548
Yoo et al [56]											
BBS	0.7142	1.0501	0.8919225	0.9560710	0.9351	11	10	0.7069	0.3593	0.1940	0.1940

[a]SMC: standardized mean test.

[b]BBS: Berg Balance Scale.

[c]TUG: Timed Up and Go.

Figure 2. Forest plot. AR: augmented reality; SMCR: standardized mean change with raw standardization; BBS: Berg Balance Scale; TUG: Timed Up and Go.



XSL•FO
RenderX

Figure 3. Weights and sample size of each study. SMC: standardized mean change; AR: augmented reality; BBS: Berg Balance Scale; TUG: Timed Up and Go.

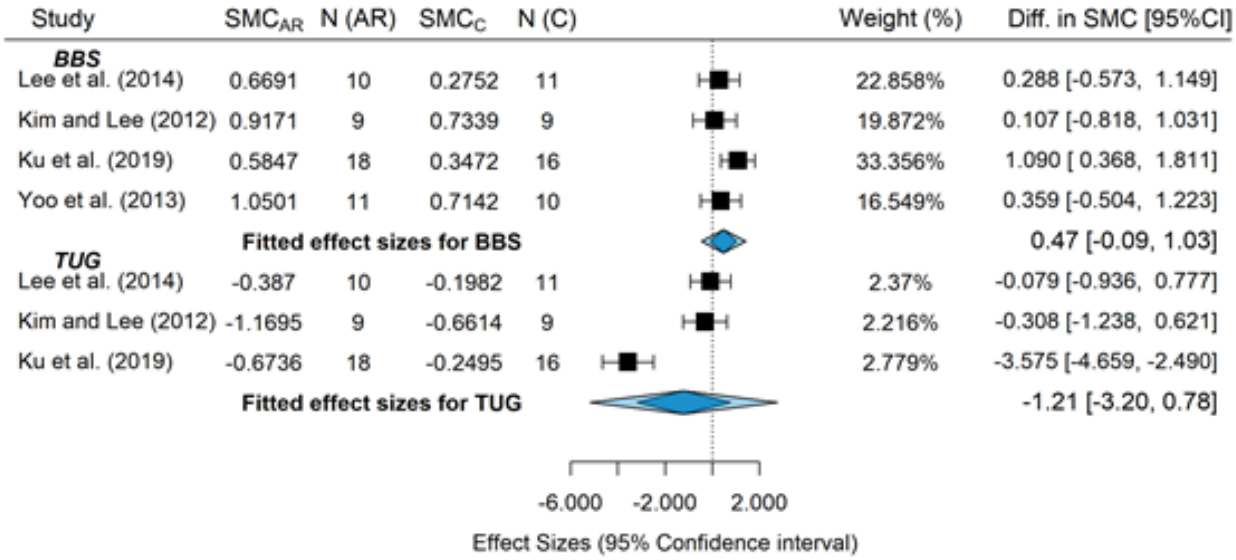
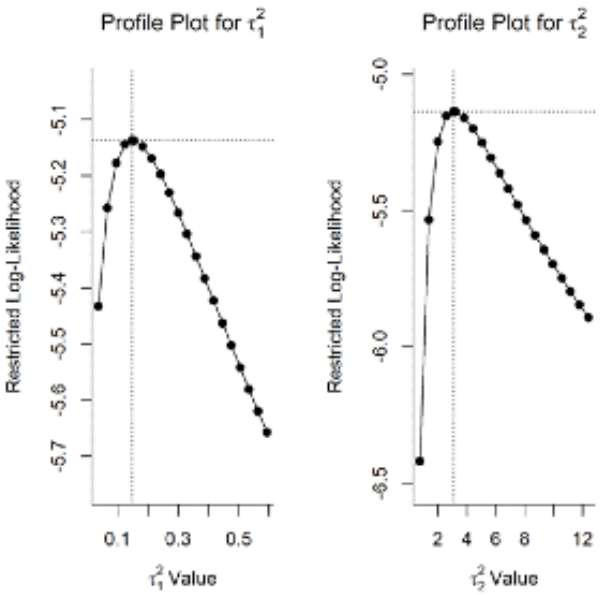
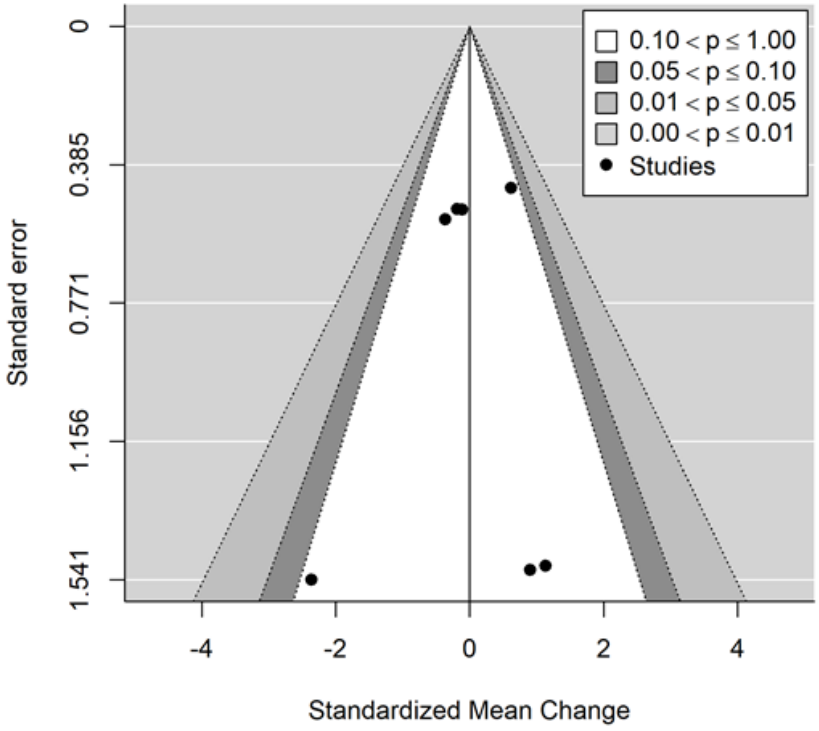


Figure 4. Variance components.



XSL•FO
RenderX

Figure 5. Funnel plot of the standardized mean change versus the standard error.



Discussion

Principal Findings

In this systematic review and meta-analysis of clinical trials, we wanted to determine the use of AR in conjunction with conventional therapy in the different fields of physiotherapy. In our study, favorable results were obtained in balance and gait [54,59], upper limb functionality [51], muscle mass, physical performance, and exercise self-efficacy [57] and in reducing the risk of falls [54-56] and pain in phantom pain syndrome [53]. In addition, significant differences were found with respect to conventional therapy. This intervention was implemented for stroke, amputations, older adults, and Parkinson disease [49,51-56,59]. These findings are consistent with those of other studies in healthy participants, such as the study by Bennour et al [28], which showed promising results for retraining the lower limb in gait through footprint modifications using AR, or the upper limb in the trial by Cavalcanti et al [30] using the AR device *ARkanoidAR*, which improved and corrected movement with the use of auditory, textual, or imaging feedback.

The aspects related to AR interventions and their positive results are as follows. Regarding their use in patients with amputations, the 2 articles found conflicting results on AR in phantom limb pain. In the trial by Ortiz-Catalán et al [53], pain was significantly reduced using AR; however, Rothgangel et al [49] found no additional effect compared with the other groups. As a possible cause, Rothgangel et al [49] argued that an inconsistency during teleprocessing with the representation of the amputated limb could have led to a lack of integration.

In patients who had a stroke, we found improvement in the functionality of the upper limb [51], with high motivation among participants and improvements in the strength of the lower limb, balance, and gait. Protocol studies on these last 2 variables have also been found in stroke, with the AR therapy C-Mill [60] and the Gait Adaptation for Stroke Patients with AR system [61], which have not yet yielded results. AR also appears promising for the rehabilitation of hand-eye coordination and finger dexterity [62].

Regarding geriatrics, favorable results were found in lower limb strength, balance, muscle mass, physical performance, exercise self-efficacy, and fall prevention. It is in this area that we have seen greater consistency in the findings. In this sense, for older adults who normally depend on visual information to achieve balance, AR training could effectively improve proprioception of the lower limbs, favoring static balance. It would be even better if the used system provides visual feedback [54].

There are other areas within physiotherapy where AR could be used to improve these parameters, such as Parkinson disease, where VR has been used to improve balance [63]. Experiments are also being conducted with a platform based on AR and the Microsoft Kinect v1 sensor, where various exercises are implemented with linear or circular movement patterns that

XSL•FO
RenderX

allow the physiotherapist to adjust them to the patient's abilities, although there are still no results [64]. However, AR visual cues did not improve freezing of gait, impaired axial kinematics, or turn scaling and timing [58].

It was possible to conduct the meta-analysis by taking into account the BBS and TUG. The BBS comprises 14 items where the patient is asked to perform several specific tasks to check their balance. Total scores range from 0 (severely affected balance) to 56 (excellent balance) [65]. Individuals with values ≤45 are at greater risk of falling [66]. With respect to the TUG, it is a scale that serves to check a patient's balance and risk of falling [67]. A duration of ≥13.5 seconds on the TUG is associated with a greater risk of falling in older adults and in people with vestibular dysfunction [68]. With the results obtained in both subgroups—BBS [52,54,56,59] and TUG [52,54,59]—the global result of the meta-analysis was favorable so that the intervention using AR is effective for the improvement of balance. However, given the small size of the samples, the heterogeneity of the populations studied, measuring instruments, methods used, times of application, and frequency and duration of the treatments, the results were not conclusive.

Advantageous aspects of AR use have also been described [51-56,59]. However, the procedures used were different in each study [49,51-56,59]. This may lead to uncertainty in the choice of a system for AR and physiotherapy development. Regarding the systems used, although in the past decade they were much more complex [36], they should be simpler in the future. With the present advances in AR systems, such as the HoloLens, its application in clinical settings could be expected to increase [69].

Displays used in AR can be classified into the following categories: head-worn, handheld, and projective [36]. In our research, most of them were projective, except in 2 of the studies [50,58], where head-mounted devices were used. Regarding the classification of the AR system by levels [70], all the systems used in our review were level 3, in which AR is displayed on screens and transformed into augmented vision through projectors that allow the real environment to become an immersive virtual world. The exception was the study by Ortiz-Catalán et al [53], in which a level 1 was used through markers, from which the 3D information contained was extracted, showing it through a device screen. In relation to the type of feedback used, our findings were the same as those of Hussain Al-Issa [36], where it was of a visual type, although in one of the studies in our paper, there was also auditory feedback [59].

There are some obstacles that limit the generalized use of AR, such as technological and user interface limitations [71]. Other negative aspects such as eye fatigue or human factors related to the effects of long-term use, such as latency and the user's adaptation to the equipment, could also reduce task performance. In addition, depth perception can make objects appear farther away than they really are [72]. It also seems that AR has not been used because, for the same objective, other technologies with easier approaches could be used, such as VR [10,69].

However, it should be considered that AR has certain advantages that VR does not have. For example, VR cannot recognize the real dangers that can cause injury, whereas, in AR, the patient is aware of the possible risks [73]. In addition, the participant can interact with the application, the environment, and tangible objects [36], as AR has greater proprioceptive feedback [74]. Game-based rehabilitation would also be of interest to create an interface (means) suitable for AR, encouraging the use of personalized games, which could improve motivation by taking into account whether the game is meaningful and motivating, the type of feedback obtained, the usability, and the interaction technique used with the environment [69].

Furthermore, the benefits of AR in the use of telerehabilitation demonstrate its effectiveness in the remote monitoring of the patient and can even modify according to their progress, providing high-quality attention with reduced costs [75,76]. Thus, the development of an AR system on mobile devices could be a good alternative for patients [77]. It seems that physiotherapy has not yet discovered all the potential promised by AR. What does appear to have been a common approach to the use of AR is lower limb recovery for fall prevention and improved balance.

This study may serve as an aid in clinical practice through the use of AR systems. It may also serve as a preliminary step toward further research with a more homogeneous methodology and the ability to experiment with these technological systems in other areas of physiotherapy where pain, functionality, balance, and fall prevention may be an objective to be pursued.

Limitations

In terms of the limitations found, we must mention the limited number of studies with low quality and the wide variety of AR interventions with respect to the system used, number of sessions, and frequency and duration of the treatment sessions. There was no homogeneity with respect to the instruments used to measure the variables studied or the variables themselves. Similarly, the need for authors to use the same measuring instruments stands out as, in some cases, it was not possible to compare studies statistically because different versions of the same scale or different units of measurement were used.

Comparison With Prior Work

After 10 years of the review by Hussain Al-Issa [36] and with results in promising pilot studies where a great future is always foreseen, our search shows the opposite. We found few studies with considerable heterogeneity and few physiotherapy plots. A difference found with respect to this previous review is that AR is now being used in telerehabilitation [8,34,76,78], although more research is needed.

Conclusions

According to the results obtained, we can say that AR, in combination with conventional therapy, has been used for physical performance, treatment of balance and prevention of falls in geriatrics, functionality of the lower limb and upper limb in stroke, and pain in phantom pain syndrome. However, no positive results were obtained with turning and timing in the freezing of gait in Parkinson disease. Owing to the diversity of the interventions and the variables measured, no consensus can

XSL•FO
RenderX

be reached on the best AR system in each area studied, although the most commonly used were the level 3 projectives.

Future clinical trials are needed using larger sample sizes and with greater homogeneity in terms of the devices used and the frequency and intensity of the interventions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Descriptive data extracted from the selected studies.
[XLS File (Microsoft Excel File), 28 KB - games_v9i4e30985_app1.xls]

References

1. Kipper G, Rampolla J. Augmented Reality An Emerging Technologies Guide to AR. Amsterdam: Elsevier Science; 2012.
2. Badilla Quesada M, Sandoval Poveda A. Realidad aumentada como tecnología aplicada a la educación superior: una experiencia en desarrollo. Innov Educ 2016 May 02;17(23):41-50. [doi: 10.22458/ie.v17i23.1369]
3. Portalés C, Lerma JL, Navarro S. Augmented reality and photogrammetry: a synergy to visualize physical and virtual city environments. ISPRS J Photogrammetry Remote Sensing 2010 Jan;65(1):134-142. [doi: 10.1016/j.isprsjprs.2009.10.001]
4. Davis S, Serrano D. La realidad aumentada como nuevo concepto de la publicidad online a través de los Smartphones | Augmented reality as new concept online advertising through Smartphones. Comunicación como valor de desarrollo social. 2012. URL: http://www.revistarazonypalabra.org/index.php/ryp/article/view/492 [accessed 2021-11-24]
5. Botella C, María B, Azucena G, Soledad Q, Verónica G, José M. La utilización de las nuevas tecnologías de la información y la comunicación en psicología clínica. UOC Papers Revista Sobre La Sociedad Del Conocimiento. 2007. URL: https://www.redalyc.org/articulo.oa?id=79000409 [accessed 2021-11-24]
6. Ortiz Rangel Ing CE. Realidad aumentada en medicina. Revista Colombiana de Cardiología 2011;18(1):4-7. [doi: 10.1016/s0120-5633(11)70160-7]
7. Ni MY, Hui RW, Li TK, Tam AH, Choy LL, Ma KK, et al. Augmented reality games as a new class of physical activity interventions? The impact of Pokémon Go use and gaming intensity on physical activity. Games Health J 2019 Feb;8(1):1-6. [doi: 10.1089/g4h.2017.0181] [Medline: 30153041]
8. Chinthammit W, Merritt T, Pedersen S, Williams A, Visentin D, Rowe R, et al. Ghostman: augmented reality application for telerehabilitation and remote instruction of a novel motor skill. Biomed Res Int 2014;2014:646347 [FREE Full text] [doi: 10.1155/2014/646347] [Medline: 24829910]
9. Manuel N, Navarrete M. La realidad virtual como arma terapéutica en rehabilitación. Rehabil Integral 2010 Jun;5(1):40-45.
10. Sveistrup H. Motor rehabilitation using virtual reality. J Neuroeng Rehabil 2004 Dec 10;1(1):10 [FREE Full text] [doi: 10.1186/1743-0003-1-10] [Medline: 15679945]
11. Postolache O, Monge J, Alexandre R, Geman O, Jin Y, Postolache G. Virtual reality and augmented reality technologies for smart physical rehabilitation. In: Advanced Systems for Biomedical Applications. Cham: Springer; 2021.
12. Cary F, Postolache O, Girao PS. Kinect based system and serious game motivating approach for physiotherapy assessment and remote session monitoring. Int J Smart Sensing Intell Syst 2020 Jan;7(5):1-6 [FREE Full text] [doi: 10.21307/ijssis-2019-131]
13. Lee H, Chang W, Lee J, Hwang J. Therapeutic potential of the home-based exercise program with the augmented reality system on balance in stroke patients: a preliminary report. Annals Physical Rehab Med 2018 Jul;61(Supplement):e36. [doi: 10.1016/j.rehab.2018.05.079]
14. Borghese N, Mainetti R, Pirovano M, Lanzi PL. An intelligent game engine for the at-home rehabilitation of stroke patients. In: Proceedings of the IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH). 2013 Presented at: IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH); May 2-3, 2013; Vilamoura, Portugal. [doi: 10.1109/segah.2013.6665318]
15. Alexandre R, Postolache O, Silva Girão P. Physical rehabilitation based on smart wearable and virtual reality serious game. In: Proceedings of the 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC). 2019 Presented at: 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC); May, 20-23, 2019; Auckland, New Zealand. [doi: 10.1109/i2mtc.2019.8826947]
16. Subramanian S, Knaut LA, Beaudoin C, McFadyen BJ, Feldman AG, Levin MF. Virtual reality environments for post-stroke arm rehabilitation. J Neuroeng Rehabil 2007 Jun 22;4:20 [FREE Full text] [doi: 10.1186/1743-0003-4-20] [Medline: 17587441]
17. Veerbeek JM, van Wegen E, van Peppen R, van der Wees PJ, Hendriks E, Rietberg M, et al. What is the evidence for physical therapy poststroke? A systematic review and meta-analysis. PLoS One 2014 Feb 4;9(2):e87987 [FREE Full text] [doi: 10.1371/journal.pone.0087987] [Medline: 24505342]

XSL•FO
RenderX

18. Laver KE, Lange B, George S, Deutsch JE, Saposnik G, Crotty M. Virtual reality for stroke rehabilitation. Cochrane Database Syst Rev 2017 Nov 20;11:CD008349 [FREE Full text] [doi: 10.1002/14651858.CD008349.pub4] [Medline: 29156493]
19. Ikbali Afsar S, Mirzayev I, Umit Yemisci O, Cosar Saracgil SN. Virtual reality in upper extremity rehabilitation of stroke patients: a randomized controlled trial. J Stroke Cerebrovasc Dis 2018 Dec;27(12):3473-3478. [doi: 10.1016/j.jstrokecerebrovasdis.2018.08.007] [Medline: 30193810]
20. Ghai S, Ghai I. Virtual reality enhances gait in cerebral palsy: a training dose-response meta-analysis. Front Neurol 2019 Mar 26;10:236 [FREE Full text] [doi: 10.3389/fneur.2019.00236] [Medline: 30984095]
21. Casuso-Holgado MJ, Martín-Valero R, Carazo AF, Medrano-Sánchez EM, Cortés-Vega MD, Montero-Bancalero FJ. Effectiveness of virtual reality training for balance and gait rehabilitation in people with multiple sclerosis: a systematic review and meta-analysis. Clin Rehabil 2018 Sep;32(9):1220-1234. [doi: 10.1177/0269215518768084] [Medline: 29651873]
22. Moreno-Verdu M, Ferreira-Sanchez MR, Cano-de-la-Cuerda R, Jimenez-Antona C. [Efficacy of virtual reality on balance and gait in multiple sclerosis. Systematic review of randomized controlled trials]. Rev Neurol 2019 May 01;68(9):357-368 [FREE Full text] [doi: 10.33588/rn.6809.2018350] [Medline: 31017288]
23. Dockx K, Bekkers EM, Van den Bergh V, Ginis P, Rochester L, Hausdorff JM, et al. Virtual reality for rehabilitation in Parkinson's disease. Cochrane Database Syst Rev 2016 Dec 21;12:CD010760 [FREE Full text] [doi: 10.1002/14651858.CD010760.pub2] [Medline: 28000926]
24. Wang B, Shen M, Wang Y, He Z, Chi S, Yang Z. Effect of virtual reality on balance and gait ability in patients with Parkinson's disease: a systematic review and meta-analysis. Clin Rehabil 2019 Jul;33(7):1130-1138. [doi: 10.1177/0269215519843174] [Medline: 31016994]
25. De Miguel-Rubio A, Rubio MD, Alba-Rueda A, Salazar A, Moral-Munoz J, Lucena-Anton D. Virtual reality systems for upper limb motor function recovery in patients with spinal cord injury: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Dec 03;8(12):e22537 [FREE Full text] [doi: 10.2196/22537] [Medline: 33270040]
26. De Miguel-Rubio A, Rubio M, Salazar A, Camacho R, Lucena-Anton D. Effectiveness of virtual reality on functional performance after spinal cord injury: a systematic review and meta-analysis of randomized controlled trials. J Clin Med 2020 Jul 01;9(7):2065 [FREE Full text] [doi: 10.3390/jcm9072065] [Medline: 32630234]
27. Mallari B, Spaeth EK, Goh H, Boyd BS. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis. J Pain Res 2019;12:2053-2085 [FREE Full text] [doi: 10.2147/JPR.S200498] [Medline: 31308733]
28. Bennour S, Ulrich B, Legrand T, Jolles BM, Favre J. A gait retraining system using augmented-reality to modify footprint parameters: effects on lower-limb sagittal-plane kinematics. J Biomech 2018 Jan 03;66:26-35. [doi: 10.1016/j.jbiomech.2017.10.030] [Medline: 29137725]
29. Heeren A, van Ooijen M, Geurts AC, Day BL, Janssen TW, Beek PJ, et al. Step by step: a proof of concept study of C-Mill gait adaptability training in the chronic phase after stroke. J Rehabil Med 2013 Jul;45(7):616-622 [FREE Full text] [doi: 10.2340/16501977-1180] [Medline: 23811818]
30. Cavalcanti VC, Ferreira MI, Teichrieb V, Barioni RR, Correia WF, Da Gama AE. Usability and effects of text, image and audio feedback on exercise correction during augmented reality based motor rehabilitation. Computers & Graphics 2019 Dec;85:100-110. [doi: 10.1016/j.cag.2019.10.001]
31. Klein A, Assis GA. A markeless augmented reality tracking for enhancing the user interaction during virtual rehabilitation. In: Proceedings of the 2013 XV Symposium on Virtual and Augmented Reality. 2013 Presented at: 2013 XV Symposium on Virtual and Augmented Reality; May 28-31, 2013; Cuiaba - Mato Grosso, Brazil. [doi: 10.1109/svr.2013.43]
32. Invernizzi M, Runza L, De Sire A, Lippi L, Blundo C, Gambini D, et al. Integrating augmented reality tools in breast cancer related lymphedema prognostication and diagnosis. J Vis Exp 2020 Feb 06(156):e60093. [doi: 10.3791/60093] [Medline: 32090996]
33. Tunur T, DeBlois A, Yates-Horton E, Rickford K, Columna LA. Augmented reality-based dance intervention for individuals with Parkinson's disease: a pilot study. Disabil Health J 2020 Apr;13(2):100848. [doi: 10.1016/j.dhjo.2019.100848] [Medline: 31679951]
34. Yeo S, Lim J, Do J, Lim J, In Lee J, Hwang J. Effectiveness of interactive augmented reality-based telerehabilitation in patients with adhesive capsulitis: protocol for a multi-center randomized controlled trial. BMC Musculoskelet Disord 2021 Apr 26;22(1):386 [FREE Full text] [doi: 10.1186/s12891-021-04261-1] [Medline: 33902546]
35. van de Venis L, van de Warrenburg BP, Weerdesteyn V, van Lith BJ, Geurts A, Nonnekes J. Improving gait adaptability in patients with hereditary spastic paraplegia (Move-HSP): study protocol for a randomized controlled trial. Trials 2021 Jan 07;22(1):32 [FREE Full text] [doi: 10.1186/s13063-020-04932-9] [Medline: 33413555]
36. Al-Issa H, Regenbrecht H, Hale L. Augmented reality applications in rehabilitation to improve physical outcomes. Physical Ther Rev 2013 Nov 12;17(1):16-28. [doi: 10.1179/1743288x11y.0000000051]
37. Page M, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow C, et al. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. J Clin Epidemiol 2021 Jun;134:103-112. [doi: 10.1016/j.jclinepi.2021.02.003] [Medline: 33577987]

XSL•FO
RenderX

38. Bhogal SK, Teasell RW, Foley NC, Speechley MR. The PEDro scale provides a more comprehensive measure of methodological quality than the Jadad scale in stroke rehabilitation literature. J Clin Epidemiol 2005 Jul;58(7):668-673. [doi: 10.1016/j.jclinepi.2005.01.002] [Medline: 15939217]
39. Moseley AM, Herbert RD, Sherrington C, Maher CG. Evidence for physiotherapy practice: a survey of the Physiotherapy Evidence Database (PEDro). Aust J Physiother 2002;48(1):43-49. [doi: 10.1016/s0004-9514(14)60281-6] [Medline: 11869164]
40. Alderson, D. Scottish Intercollegiate Guidelines Network (SIGN) 87 — the Management of Oesophageal and Gastric Cancer. Clinical Oncology 2008;20(7):530-531. [doi: https://doi.org/10.1016/j.clon.2008.04.015]
41. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions - Version 5. New York: John Wiley & Sons; 2011.
42. Becker B. Synthesizing standardized mean-change measures. Br J Math Stat Psychol 1988 Nov;41(2):257-278. [doi: 10.1111/j.2044-8317.1988.tb00901.x]
43. Morris S. Estimating effect sizes from pretest-posttest-control group designs. Organizational Res Methods 2007 Jul 23;11(2):364-386. [doi: 10.1177/1094428106291059]
44. Gleser L, Olkin I. Stochastically dependent effect sizes. In: The Handbook of Research Synthesis and Meta-analysis. New York: Russell Sage Foundation; 2009.
45. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. Res Synth Methods 2010 Apr;1(2):112-125. [doi: 10.1002/jrsm.11] [Medline: 26061377]
46. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. J Clin Epidemiol 2008 Oct;61(10):991-996. [doi: 10.1016/j.jclinepi.2007.11.010] [Medline: 18538991]
47. Viechtbauer W. Conducting meta-analyses in with the package. J Stat Soft 2010;36(3):1-48. [doi: 10.18637/jss.v036.i03]
48. R Core Team (2020). European Environment Agency. URL: https://www.eea.europa.eu/data-and-maps/indicators/oxygen-consuming-substances-in-rivers/r-development-core-team-2006 [accessed 2021-11-25]
49. Rothgangel A, Braun S, Winkens B, Beurskens A, Smeets R. Traditional and augmented reality mirror therapy for patients with chronic phantom limb pain (PACT study): results of a three-group, multicentre single-blind randomized controlled trial. Clin Rehabil 2018 Dec;32(12):1591-1608. [doi: 10.1177/0269215518785948] [Medline: 30012007]
50. Jung G, Moon T, Park G, Lee J, Lee B. Use of augmented reality-based training with EMG-triggered functional electric stimulation in stroke rehabilitation. J Phys Ther Sci 2013;25(2):147-151 [FREE Full text] [doi: 10.1589/jpts.25.147]
51. Colomer C, Llorens R, Noé E, Alcañiz M. Effect of a mixed reality-based intervention on arm, hand, and finger function on chronic stroke. J Neuroeng Rehabil 2016 May 11;13(1):45 [FREE Full text] [doi: 10.1186/s12984-016-0153-6] [Medline: 27169462]
52. Lee C, Kim Y, Lee B. Augmented reality-based postural control training improves gait function in patients with stroke: randomized controlled trial. Hong Kong Physiother J 2014 Dec;32(2):51-57. [doi: 10.1016/j.hkpj.2014.04.002]
53. Ortiz-Catalan M, Guðmundsdóttir RA, Kristoffersen MB, Zepeda-Echavarria A, Caine-Winterberger K, Kulbacka-Ortiz K, et al. Phantom motor execution facilitated by machine learning and augmented reality as treatment for phantom limb pain: a single group, clinical trial in patients with chronic intractable phantom limb pain. Lancet 2016 Dec 10;388(10062):2885-2894. [doi: 10.1016/S0140-6736(16)31598-7] [Medline: 27916234]
54. Ku J, Kim Y, Cho S, Lim T, Lee H, Kang Y. Three-dimensional augmented reality system for balance and mobility rehabilitation in the elderly: a randomized controlled trial. Cyberpsychol Behav Soc Netw 2019 Feb;22(2):132-141. [doi: 10.1089/cyber.2018.0261] [Medline: 30596530]
55. Lee J, Yoo H, Lee B. Effects of augmented reality-based Otago exercise on balance, gait, and physical factors in elderly women to prevent falls: a randomized controlled trial. J Phys Ther Sci 2017 Sep;29(9):1586-1589 [FREE Full text] [doi: 10.1589/jpts.29.1586] [Medline: 28931993]
56. Yoo H, Chung E, Lee B. The effects of augmented reality-based otago exercise on balance, gait, and falls efficacy of elderly women. J Phys Ther Sci 2013 Jul;25(7):797-801 [FREE Full text] [doi: 10.1589/jpts.25.797] [Medline: 24259856]
57. Jeon S, Kim J. Effects of augmented-reality-based exercise on muscle parameters, physical performance, and exercise self-efficacy for older adults. Int J Environ Res Public Health 2020 May 07;17(9):3260 [FREE Full text] [doi: 10.3390/ijerph17093260] [Medline: 32392833]
58. Janssen S, de Ruyter van Steveninck J, Salim HS, Cockx HM, Bloem BR, Heida T, et al. The effects of augmented reality visual cues on turning in place in Parkinson's disease patients with freezing of gait. Front Neurol 2020 Mar 24;11:185 [FREE Full text] [doi: 10.3389/fneur.2020.00185] [Medline: 32265826]
59. Kim I, Lee B. Effects of augmented reality with functional electric stimulation on muscle strength, balance and gait of stroke patients. J Phys Ther Sci 2012;24(8):755-762. [doi: 10.1589/jpts.24.755]
60. Timmermans C, Roerdink M, van Ooijen MW, Meskers CG, Janssen TW, Beek PJ. Walking adaptability therapy after stroke: study protocol for a randomized controlled trial. Trials 2016 Aug 26;17(1):425 [FREE Full text] [doi: 10.1186/s13063-016-1527-6] [Medline: 27565425]

XSL•FO
RenderX

61. Rossano C, Terrier P. Visually-guided gait training in paretic patients during the first rehabilitation phase: study protocol for a randomized controlled trial. Trials 2016 Oct 27;17(1):523 [FREE Full text] [doi: 10.1186/s13063-016-1630-8] [Medline: 27788679]
62. Shen Y, Gu PW, Ong SK, Nee AY. A novel approach in rehabilitation of hand-eye coordination and finger dexterity. Virtual Real 2011 Aug 5;16(2):161-171. [doi: 10.1007/s10055-011-0194-x]
63. van den Heuvel MR, Kwakkel G, Beek PJ, Berendse HW, Daffertshofer A, van Wegen EE. Effects of augmented visual feedback during balance training in Parkinson's disease: a pilot randomized clinical trial. Parkinsonism Related Disorders 2014 Dec;20(12):1352-1358. [doi: 10.1016/j.parkreldis.2014.09.022]
64. Pachoulakis I, Xilourgos N, Papadopoulos N, Analyti A. A kinect-based physiotherapy and assessment platform for Parkinson's Disease patients. J Med Eng 2016;2016:9413642 [FREE Full text] [doi: 10.1155/2016/9413642] [Medline: 27822467]
65. Berg OL, 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]
66. Guzman ME, Fabiola Y, Cisternas C. Correlation between Berg Balance Scaleand center of pressure variables in older adults. Revista de estudios en movimiento 2017:25-29.
67. Shumway-Cook A, Brauer S, Woollacott M. Phys Ther 2000 Sep;80(9):896-903. [Medline: 10960937]
68. Whitney SL, Marchetti GF, Schade A, Wrisley DM. J Vestib Res 2004;14(5):397-409. [Medline: 15598995]
69. Mubin O, Alnajjar F, Jishtu N, Alsinglawi B, Al Mahmud A. Exoskeletons with virtual reality, augmented reality, and gamification for stroke patients' rehabilitation: systematic review. JMIR Rehabil Assist Technol 2019 Sep 08;6(2):e12010 [FREE Full text] [doi: 10.2196/12010] [Medline: 31586360]
70. Augmented reality and applications. Technology Research and Academy. 2018. URL: https://revistas.udistrital.edu.co/index.php/tia/article/view/11281 [accessed 2021-11-25]
71. Ellis S. Factors influencing operator interaction with virtual objects viewed via head-mounted see-through displays: viewing conditions and rendering latency. In: Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality. 1997 Presented at: Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality; Mar 1-5, 1997; Albuquerque, NM, USA. [doi: 10.1109/vrais.1997.583063]
72. Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B. Recent advances in augmented reality. IEEE Comput Grap Appl 2001;21(6):34-47. [doi: 10.1109/38.963459]
73. Chan ZY, MacPhail AJ, Au IP, Zhang JH, Lam BM, Ferber R, et al. Walking with head-mounted virtual and augmented reality devices: effects on position control and gait biomechanics. PLoS One 2019 Dec 4;14(12):e0225972 [FREE Full text] [doi: 10.1371/journal.pone.0225972] [Medline: 31800637]
74. Cameirão MS, Badia SB, Duarte E, Frisoli A, Verschure PF. The combined impact of virtual reality neurorehabilitation and its interfaces on upper extremity functional recovery in patients with chronic stroke. Stroke 2012 Oct;43(10):2720-2728. [doi: 10.1161/STROKEAHA.112.653196] [Medline: 22871683]
75. Nardi AD, Real A, Santos TD, Rocha RD, Lenzi T. Effect of inspiratory muscle training on patients undergoing bariatric surgery: a systematic review. Fisioter Pesqui. 2016. URL: https://www.scielo.br/j/fp/a/BVfqz7YXFpYLFtYZTFwh48d/?lang=pt [accessed 2021-11-25]
76. Cerdán de Las Heras J, Tulppo M, Kiviniemi A, Hilberg O, Løkke A, Ekholm S, et al. Augmented reality glasses as a new tele-rehabilitation tool for home use: patients' perception and expectations. Disabil Rehabil Assist Technol 2020 Aug 04:1-7. [doi: 10.1080/17483107.2020.1800111] [Medline: 32750254]
77. Ruiz Torres D. Augmented reality: a new resource in Information and Communication Technology (ICT) for museums in the 21st century. Intervención (Méx. DF) 2012;3(5):39-44.
78. Borresen A, Wolfe C, Lin C, Tian Y, Raghuraman S, Nahrstedt K, et al. Usability of an immersive Augmented Reality Based Telerehabilitation System with Haptics (ARTESH) for synchronous remote musculoskeletal examination. Int J Telerehabil 2019 Jun 12;11(1):23-32 [FREE Full text] [doi: 10.5195/ijt.2019.6275] [Medline: 31341544]

Abbreviations

AR: augmented reality
BBS: Berg Balance Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta‐Analyses
PROSPERO: International Prospective Register of Systematic Reviews
SMCR: standardized mean change with raw standardization
TUG: Timed Up and Go
VR: virtual reality

XSL•FO
RenderX

Edited by N Zary; submitted 04.06.21; peer-reviewed by JJ Jimenez-Rejano, R Ciorap, E Sadeghi-Demneh, S Braun; comments to author 22.07.21; revised version received 11.09.21; accepted 24.09.21; published 15.12.21.

Please cite as:
Vinolo Gil MJ, Gonzalez-Medina G, Lucena-Anton D, Perez-Cabezas V, Ruiz-Molinero MDC, Martín-Valero R
Augmented Reality in Physical Therapy: Systematic Review and Meta-analysis
JMIR Serious Games 2021;9(4):e30985
URL: https://games.jmir.org/2021/4/e30985
doi:10.2196/30985
PMID:34914611

©Maria Jesus Vinolo Gil, Gloria Gonzalez-Medina, David Lucena-Anton, Veronica Perez-Cabezas, María Del Carmen Ruiz-Molinero, Rocío Martín-Valero. Originally published in JMIR Serious Games (https://games.jmir.org), 15.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Dissecting Digital Card Games to Yield Digital Biomarkers for the Assessment of Mild Cognitive Impairment: Methodological Approach and Exploratory Study

Karsten Gielis[1], MSc, PhD; Marie-Elena Vanden Abeele[2], MD; Robin De Croon[3], MSc, PhD; Paul Dierick[4], MA, PhD; Filipa Ferreira-Brito[5], MA; Lies Van Assche[6,7], MA, PhD; Katrien Verbert[3], MA, PhD; Jos Tournoy[8,9], MD, PhD; Vero Vanden Abeele[1], MA, PhD

[1]e-Media Research Lab, Katholieke Universiteit Leuven, Leuven, Belgium
[2]Memory Clinic, Jessa Hospital, Hasselt, Belgium
[3]Department of Computer Science, Katholieke Universiteit Leuven, Leuven, Belgium
[4]Department of Gerontopsychiatry, University Psychiatric Center, Duffel, Belgium
[5]Instituto de Saúde Ambiental, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal
[6]Section of Geriatric Psychiatry, University Hospital Leuven, Katholieke Universiteit Leuven, Leuven, Belgium
[7]Department of Psychiatry, University Hospital Leuven, Katholieke Universiteit Leuven, Leuven, Belgium
[8]Department of Geriatric Medicine, University Hospital Leuven, Leuven, Belgium
[9]Department of Public Health and Primary Care, Gerontology and Geriatrics, Katholieke Universiteit Leuven, Leuven, Belgium

Corresponding Author:
Karsten Gielis, MSc, PhD
e-Media Research Lab
Katholieke Universiteit Leuven
Andreas Vesaliusstraat 13
Leuven, 3000
Belgium
Phone: 32 16376866
Email: karsten.gielis@kuleuven.be

Abstract

Background: Mild cognitive impairment (MCI), the intermediate cognitive status between normal cognitive decline and pathological decline, is an important clinical construct for signaling possible prodromes of dementia. However, this condition is underdiagnosed. To assist monitoring and screening, digital biomarkers derived from commercial off-the-shelf video games may be of interest. These games maintain player engagement over a longer period of time and support longitudinal measurements of cognitive performance.

Objective: This paper aims to explore how the player actions of Klondike Solitaire relate to cognitive functions and to what extent the digital biomarkers derived from these player actions are indicative of MCI.

Methods: First, 11 experts in the domain of cognitive impairments were asked to correlate 21 player actions to 11 cognitive functions. Expert agreement was verified through intraclass correlation, based on a 2-way, fully crossed design with type consistency. On the basis of these player actions, 23 potential digital biomarkers of performance for Klondike Solitaire were defined. Next, 23 healthy participants and 23 participants living with MCI were asked to play 3 rounds of Klondike Solitaire, which took 17 minutes on average to complete. A generalized linear mixed model analysis was conducted to explore the differences in digital biomarkers between the healthy participants and those living with MCI, while controlling for age, tablet experience, and Klondike Solitaire experience.

Results: All intraclass correlations for player actions and cognitive functions scored higher than 0.75, indicating good to excellent reliability. Furthermore, all player actions had, according to the experts, at least one cognitive function that was on average moderately to strongly correlated to a cognitive function. Of the 23 potential digital biomarkers, 12 (52%) were revealed by the generalized linear mixed model analysis to have sizeable effects and significance levels. The analysis indicates sensitivity of the derived digital biomarkers to MCI.

XSL•FO
RenderX

Conclusions: Commercial off-the-shelf games such as digital card games show potential as a complementary tool for screening and monitoring cognition.

Trial Registration: ClinicalTrials.gov NCT02971124; https://clinicaltrials.gov/ct2/show/NCT02971124

(JMIR Serious Games 2021;9(4):e18359) doi:10.2196/18359

KEYWORDS

mild cognitive impairment; Klondike Solitaire; card games; generalized linear mixed effects analysis; expert study; monitoring; screening; cognition; dementia; older adults; mobile phone

Introduction

Assessing Cognitive Performance

Mild cognitive impairment (MCI) is a clinical entity defined as a transitional state between normal and pathological aging, where one or more cognitive domains are significantly impaired, yet activities of daily living are still preserved [1]. Early detection of MCI is important for signaling possible prodromes of dementia, monitoring the progression of possible decline, taking supportive measures, and detecting any possible underlying causes. However, cognitive impairment is still underdiagnosed [2-4]. In response, governmental bodies have called for novel, scalable, and longitudinal tools to assist in the early screening and monitoring of dementia [5-7]. To answer this call, researchers have explored the use of digital games as a suitable medium for assessing cognitive impairment [8-11]. Games are autotelic in nature, tapping into the intrinsic motivation to play [12,13], hence captivating a player's attention. Furthermore, digital games are a natural source of information on player behavior, cognitive performance, motor skills, social conduct, and affective experiences [14].

As such, digital games may help by providing *digital* biomarkers of cognitive performance. Biomarkers, defined as "objectively measured and evaluated indicators of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention [15]" have a longstanding tradition in dementia research [16,17]. Complementary to their biological counterparts, *digital* biomarkers are "user-generated physiological and behavioral measures collected through connected *digital* devices to explain, influence and/or predict health-related outcomes [18]." User interaction with digital games produces dense and detailed behavioral traces that may inform on the users' social health, praxis, and *cognition*.

Today, the focal point of research assessing cognitive performance is *serious games*, that is, games intentionally designed and built for a serious purpose and not solely to entertain [19]. Although serious gaming interventions show potential, they are typified by less funding, shorter development cycles, and missing know-how compared with traditional video games, which affect the in-game hallmarks of quality, such as graphics, music, and storytelling [20,21]. This may lead to frustrating player experiences, a lack of engagement, and less attention being paid during gameplay, which may lower the reliability and validity of any findings and possibly cause attrition in longitudinal studies [22-24]. Therefore, most recently, the study by Mandryk and Birk [14] argued in favor of turning to commercial off-the-shelf (COTS) video games instead [14]. Instead of spending limited resources on building a serious game, researchers can devote themselves to investigating existing games already enjoyed by the target population. Although not designed to measure cognition, COTS games are woven into the fabric of everyday life and may be able to provide digital biomarkers of cognitive performance that are reflective of cognitive status [14,25].

This study aims to explore the possibility of using COTS card games to screen cognition among patients living with MCI. This study involved 11 experts in the domain of MCI coming together to craft 23 candidate digital biomarkers from the digital card game Klondike Solitaire. Subsequently, a data acquisition campaign was set up involving 46 participants: 23 (50%) healthy older adults and 23 (50%) older adults with MCI. The participants were asked to play 3 games on a tablet. We examined the game data for differences at a group level for the candidate digital biomarkers using a generalized linear mixed model (GLMM) analysis. The results show that of the 23 candidate digital biomarkers, 12 (52%) differed significantly between both groups, taking age, tablet experience, and Klondike Solitaire experience into account. By providing a methodological approach and an exploratory study for crafting digital biomarkers, as well as articulating the rationale and the different steps taken, we hope to inform future researchers who aim to leverage the use of COTS video games to yield digital biomarkers.

MCI Classification

Persons diagnosed with MCI show a deficit in cognition in at least one cognitive domain that cannot be attributed to age or any other disease; yet, they do not fulfill the diagnosis of dementia [26]. Persons with MCI, however, have a higher risk of progressing to a form of dementia such as Lewy body dementia [27], vascular dementia [28], or the most common form of dementia, Alzheimer disease [29]. Depending on the early symptoms, persons with MCI can be classified into 2 groups: amnestic MCI (aMCI) and nonamnestic MCI (naMCI). Persons in the aMCI group show a significant memory deficit, whereas in persons with naMCI, mainly a nonmemory impairment (eg, language) is present [30]. For both aMCI and naMCI, a further distinction can be made between persons with 1 cognitive domain impaired (single domain MCI) and those with multiple cognitive domains impaired (multiple-domain MCI). Although no treatments exist with the current state of modern medicine to cure the neuronal damage of these progressive forms of dementia [31,32], early diagnosis matters [33] because there are several measures that can be taken to slow down disease progression [34], including starting nonpharmacological treatment for delaying symptoms

XSL•FO
RenderX

[31,32,34], or support patients and families with the appropriate counseling [35].

Detecting MCI

Typically, the process leading to a diagnosis of MCI is set into motion by a cognitive complaint from the older adult, relative, or (informal) caregiver, followed by a presumptive identification through a screening test. The cognitive screening tests used most often to detect MCI are the Montreal Cognitive Assessment (MoCA) [36] and the Mini-Mental State Examination (MMSE) [37]. These cognitive screening tests primarily focus on evaluating language, visual skills, memory, orientation, attention, and executive functions [38]. Despite their widespread use, the psychometric properties of the screening tests by themselves are insufficient to draw firm conclusions regarding an MCI diagnosis [39].

Therefore, this presumptive identification is in turn followed by an elaborate neuropsychological assessment (ie, a cognitive test battery) and possibly a biomarker scan or a neuroimaging scan [26,30]. This neuropsychological assessment assesses cognitive skills and level of impairment more thoroughly. In addition, the assessment may include a semiguided interview with a relative or caregiver to evaluate the change in symptoms over time, such as in the Clinical Dementia Rating (CDR) Scale [40]. However, this neuropsychological assessment is laborious and time-intensive, requiring skilled test administrators, who, despite their training, are still subject to interassessor variability [41]. In addition, from a patient perspective, the process has been described as bewildering, highly stressful, and even humiliating [42,43], contributing to malperformance. This in turn can make patients self-aware of impairment, leading to feelings of distress or helplessness, possibly spiraling into even worse performance [44,45]. Although biological and imaging biomarkers are becoming more common to support diagnosis, they remain expensive and invasive, making them equally unfit for high-frequency measurements [41]. As a result, health professionals and policy makers welcome additional tools that support the monitoring of cognition [1,46-49] while reducing patient-level barriers and are more considerate of patients' experiences [44].

Serious Games for the Assessment of Cognitive Functions

Serious (digital) games are "games that do not have entertainment, enjoyment, or fun as their primary purpose [19]." An early and longstanding tradition [50] is the use of serious games for cognitive evaluation [51]. Space Fortress [19,50] is perhaps the first research game to measure and train cognitive control and related cognitive functions. Ever since, the popularity of creating serious games and game-based interventions to measure, detect, and train cognition has only increased, as indicated by systematic reviews on this topic by Ferreira-Brito et al [52], Lumsden et al [53], and Valladares-Rodriguez et al [51].

Serious games may provide certain advantages for the assessment of cognitive performance *compared with* standard cognitive assessment tests. First, by offering an interactive and immersive audiovisual experience, serious games can be considered to be more engaging than classical tests [23,51,53]. As ensuring the full attention of the participant is paramount in neuropsychological testing, such increased engagement may also result in more reliable research results; previous research has linked effort to cognitive test performance in healthy undergraduate students [54]. Second, games allow for embedding of cognitive tasks in a virtual, audiovisual world, which more closely mimics the actual lived-in world, allowing for better transfer of task results and providing *higher ecological validity* [55]. However, it has to be noted that the skills learned through serious games might still be difficult to generalize to the skills needed in a real-life context [55]. Third, serious games can be designed in such a manner that they minimize the need for the presence of a trained administrator. Setting a pace, reading out loud, or cueing can be integrated into the game itself. In this manner, test administration bias is reduced, and white-coat effects can be minimized [56,57]. If assessments are possible with less supervision and manual effort, they also become more *scalable* because testing becomes less resource intensive [55]. However, this lack of supervision while performing measurements has an important caveat. When measurements are performed in a personal environment, it becomes more difficult to prevent distractions that influence gameplay behavior.

Although serious games show promising results and have merit for both patient and physician, serious games are at risk of being dismissed as "chocolate-covered broccoli": neuropsychological tests embellished with a thin layer of gameplay [58-60]. This can lead to games that are suboptimal in terms of esthetic quality and game mechanics [20], negatively affecting gameplay [58]. A meta-analysis of *serious games* [22] shows that although they can be more effective and improve retention compared with conventional methods, they are not found to be more motivating. Similar signs of lack of motivation have been noted in game-based interventions designed to train cognitive functions [23,24].

This lack of sustained engagement contrasts with surveys on gameplay among older adults. A large-scale (N=3737) survey of older adults' attitude toward video games, conducted in 2019 by the American Association of Retired Persons [61], highlights that older adults enjoy playing digital games. Of the nine reasons to play, *to have fun* was indicated to be the top reason (78%) to play video games, with *to help stay mentally sharp* coming in second (69%). In the 70-years-or-older age category, this difference became marginal, with 73% indicating *to have fun* and 72% indicating *to stay mentally sharp* as the reasons to play. Therefore, to increase engagement and to tap into intrinsic motivation, popular COTS video games may present an interesting alternative. These games are already woven into the daily life of the older adult, providing *meaningful* play [62,63].

COTS Video Games for Mental Health

COTS games may have the ability to retain players over a longer period and to support *continuous measurements of cognitive performance*. As frequent measurements are more sensitive to detecting small deviations in the cognitive performance of older adults [64], this could lead to a better interpretation of the patient's cognitive trajectory. Furthermore, fluctuations in

XSL•FO
RenderX

cognitive performance [65], a common feature of dementia, may be more easily detected. In addition, this continuous monitoring enables establishing an intraindividual cognitive baseline [66]. This cognitive baseline can be used to compare patients with themselves, as opposed to comparisons with normed references. In turn, this can lead to improved management and care [1]. Nevertheless, a prominent disadvantage of COTS games is that researchers have less control over which cognitive functions are measured in the game [60].

Recent studies on using COTS games to measure cognitive impairment have generated promising results. The study by Jimison et al [8] used FreeCell, a Solitaire variant, to compare cognitive performance between a group of people living with MCI and a healthy control group by means of an *optimized solver*. The results indicated that based on gameplay, the group with MCI could be discerned from the healthy control group [8]. In the case of sudoku, another popular game among older adults, the study by Grabbe [67] showed that performance in the game was significantly related to measures of working memory. Using a set of smartphone-based puzzle games, which also contained sudoku, the study by Thompson et al [68] explored smartphone-based games as a means of portable cognitive assessment and monitoring. The participants' performance in these games correlated to several measures of cognition, among which were visual memory, verbal learning, and reasoning. Finally, the study by Wallace et al [11] developed a word search game and sudoku that incorporated hints to reduce frustration among patients living with MCI. Their first study with 2 patients indicated that cognitive performance could be measured with COTS gameplay, comparing game performance with the MoCA and the Repeatable Battery for the Assessment of Neuropsychological Status [69]. Synthesizing these results, these studies suggest that COTS games yield promise for the assessment of cognitive impairment but that further research is warranted.

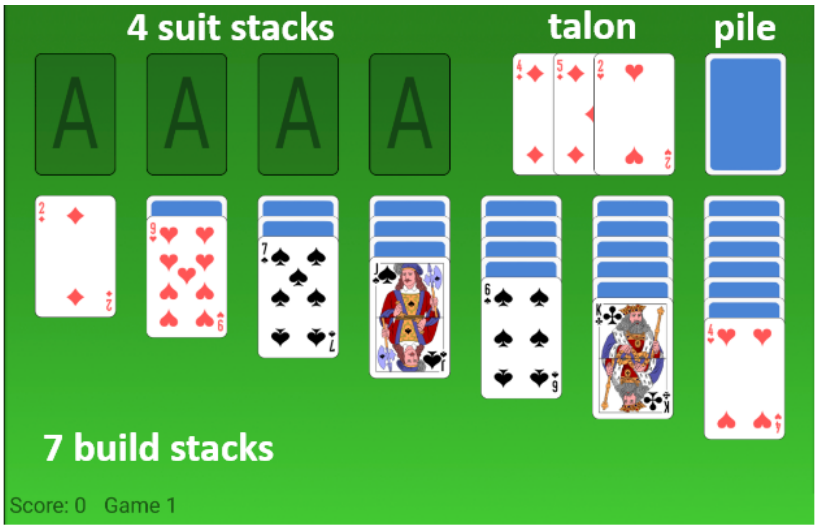
Across the aforementioned studies, different lines of reasoning have been given to justify the game of choice as suitable for neuropsychological evaluation. The study by Grabbe [67] analyzed the components of sudoku and linked them to working memory based on a subjective analysis. The study by Jimison et al [8] chose FreeCell because it was the most popular game in their focus group. The study by Wallace et al [11] chose a word search game and sudoku above other games because of certain properties such as the percentage of successful deals. Finally, the study by Thompson et al [68] chose games based on face validity with regard to targeting cognitive functions. Although these reasons are valid arguments for choosing a game, it should be noted that these studies have no arguments rooted in empirical evidence for their game of choice.

Klondike Solitaire

One of the most popular card games among older adults is Klondike Solitaire, also known as Patience, Fascination, or even just Solitaire [70]. The popularity of Klondike Solitaire among older adult gamers was recently noted in the study by Boot et al [71]. For 1 year, participants had access to a computer where 11 games were installed, among which were sudoku, Klondike Solitaire, and crossword puzzles. The study noted that "There was a strong, clear preference for Solitaire […]. After Solitaire, there was no clear second choice, and on average participants infrequently played the other games." In addition, the results showed that of all 11 games, 1 (9%)—Solitaire—was being played most consistently.

This popular card game is played with a standard 52-card deck, with 28 (54%) cards dealt in 7 build stacks and the other 24 (46%) cards placed in a pile, as can be seen in Figure 1. The goal of the game is to order all cards from the ace to the king on the 4 corresponding suit stacks. Cards can be moved on top of other build stacks if their rank is one lower than the current top card and of the opposite color. Cards can be requested from the pile to be placed on the talon (Figure 1).

Figure 1. Klondike Solitaire. The seven build stacks can be seen at the bottom, the suit stacks are at the top left. The pile of undealt cards can be seen in the top right.



XSL•FO
RenderX

Study Objective

Given the popularity of Klondike Solitaire among the older population, and given the need for engaging, ecologically valid, scalable tools to assist in the screening and monitoring of MCI, this paper aims to investigate the feasibility of Klondike Solitaire game sessions to yield digital biomarkers of MCI. To this end, the study comprised the following investigations:

1. an exploration of the digital biomarkers of cognitive performance, based on the player actions (PAs) of Klondike Solitaire and
2. an evaluation of the candidate digital biomarkers captured in Klondike Solitaire to measure the differences between healthy older adults and older adults living with MCI.

Methods

Crafting Candidate Digital Biomarkers in Klondike Solitaire

To explore the potential digital biomarkers of cognitive performance in Klondike Solitaire, we first conducted an expert consensus study, involving 11 experts. In this first part of the paper, we discuss the 3 steps taken to compile a final list of 23 candidate digital biomarkers.

Step 1: Defining PAs

To transform gameplay into PAs, a methodical approach was applied. In all, 4 researchers in the field of human-computer interaction carried out the following tasks. First, the literature on Klondike Solitaire was studied, ranging from scientific work [72-78] to more general sources [79-81]. Afterward, the game was played for a minimum of 10 sessions of 30 minutes by each of the researchers. Combining the theoretical background with this practical experience, the experts independently drafted a list of game events, which was then reviewed as a team. This list was iterated 3 times until no more game events were found. The game events included, but were not limited to, game outcomes (eg, the game was won or lost), correct player moves (eg, the player moves a card among the build stacks), and erroneous player moves (eg, the player places cards on other cards such that they are not in descending order on the build stack).

These game events were then converted to PAs; they were described as an action of the player rather than as an event of the game. Next, all these PAs were transformed into their negative equivalents, for example, *The player takes little time to think of a move* was reworded as *The player takes a lot of time to think of a move*. The reason for this was 2-fold. It enabled inverse PAs (the positive and negative equivalents, eg, moving cards fast or moving them slowly) to be combined, reducing the rating complexity for the professionals. Furthermore, the negative equivalent aimed to facilitate the rating process because impaired cognition leads to reduced performance in gameplay. After this step was completed, 21 PAs (Table 1) were defined for evaluation.

XSL•FO
RenderX

Table 1. Average of the experts' ratings for each player action and cognitive function.

Player actions	Mental flexibility, mean (SD)	Inhibitory control, mean (SD)	Working memory, mean (SD)	Selective attention, mean (SD)	Visuospatial ability, mean (SD)	Object recognition, mean (SD)	Apraxia, mean (SD)	Cognitive planning, mean (SD)	Processing speed, mean (SD)
PA[a] 1. Player takes a lot of time to think of a move.	1.64 (1.12)	0.73 (1.01)	1.82 (0.4)	1.55 (0.82)	1.18 (0.98)	1.27 (0.79)	0.27 (0.47)	*2.45*[b] (0.52)	*2.55* (0.82)
PA 2. Player takes a lot of time to move the card.	0.73 (1.01)	0.73 (1.1)	0.64 (0.5)	0.64 (0.67)	1.45 (1.04)	0.64 (0.67)	1.64 (0.92)	0.91 (0.83)	*2.09* (1.04)
PA 3. Player does not move a suitable card from the talon to the suit stack.	*2.27* (0.79)	0.73 (0.9)	*2* (0.77)	*2.55* (0.69)	1.45 (0.93)	1.64 (0.67)	0.73 (1.01)	*2.18* (0.75)	0.91 (1.04)
PA 4. Player does not move a suitable card from the build stack to the suit stack.	1.82 (0.98)	0.91 (0.83)	1.82 (0.75)	*2.36* (0.67)	1.36 (0.92)	1.36 (0.92)	0.27 (0.47)	1.73 (0.79)	1 (1)
PA 5. Player does not move a suitable card from the talon to the build stack.	*2.18* (0.75)	1.27 (0.9)	1.91 (0.7)	*2.64* (0.67)	1.55 (1.04)	1.73 (0.79)	0.18 (0.4)	*2.09* (0.94)	0.82 (1.08)
PA 6. Player does not move a suitable card from 1 build stack to another build stack.	*2.36* (0.81)	1.09 (0.94)	*2* (0.77)	*2.45* (0.69)	1.45 (0.93)	1.64 (0.92)	0.18 (0.4)	*2.18* (0.98)	0.73 (1.01)
PA 7. Player does not place an ace immediately on an empty suit stack.	1.27 (1.01)	0.73 (1.01)	*2.18* (0.4)	*2.36* (0.92)	1 (0.77)	1.18 (1.17)	0.45 (0.69)	*2.09* (0.7)	1.09 (1.04)
PA 8. Player does not place a king on an empty build stack.	1.45 (1.13)	0.73 (0.9)	*2* (0.77)	*2.27* (0.9)	1 (0.77)	1.55 (1.13)	0.36 (0.67)	*2.09* (0.7)	1 (0.89)
PA 9. Player moves cards without benefit (eg, moving a jack from 1 queen to another).	1.45 (1.04)	1.82 (0.98)	*2.18* (1.17)	1.64 (1.03)	0.82 (0.75)	1.45 (1.21)	0.18 (0.4)	*2.27* (1.01)	0.45 (0.52)
PA 10. Player flips a lot through the pile.	*2* (0.89)	*2.55* (0.69)	1.73 (1.01)	1.82 (1.08)	1 (0.89)	1.45 (1.13)	1 (1)	*2.09* (1.04)	0.91 (0.7)
PA 11. Player moves a card onto a card with the same color.	1.73 (1.1)	*2.55* (0.52)	*2.18* (0.98)	*2.18* (0.98)	1 (1)	1.82 (1.17)	0.27 (0.65)	1.36 (0.81)	0.45 (0.69)
PA 12. Player moves a card to another card with the wrong number (eg, placing a 5 on a 10).	1.18 (1.08)	*2* (1)	*2.27* (0.9)	1.91 (0.94)	1.09 (1.04)	2.09 (0.94)	0.45 (0.69)	1.45 (0.93)	0.36 (0.5)
PA 13. Player selects the cards with a very bad precision (taps on edge of card or next to the card).	0.45 (0.69)	0.73 (0.79)	0.27 (0.47)	0.64 (0.81)	*2.27* (0.9)	0.82 (0.75)	*2.27* (0.79)	0.45 (0.82)	0.45 (0.69)
PA 14. Player starts tapping on the playfield with no apparent target (with short intervals, fidget tapping).	0.73 (0.79)	*2.27* (1.01)	0.27 (0.47)	0.82 (0.87)	0.73 (0.9)	0.45 (0.52)	1.55 (1.29)	0.91 (1.04)	0.73 (0.9)
PA 15. Player presses the undo button a lot.	1.82 (0.6)	*2.45* (0.69)	1.73 (1.1)	1.36 (1.12)	0.64 (0.67)	0.64 (0.67)	0.73 (1.01)	*2.27* (1.01)	1.27 (1.1)
PA 16. Player requests a lot of hints.	1.91 (1.04)	1.73 (1.01)	*2* (1)	1.45 (0.93)	0.64 (0.81)	1 (0.77)	0.45 (0.69)	*2.27* (1.01)	1.18 (0.75)
PA 17. Player takes a very long time to finish games.	*2.18* (1.25)	1 (1.34)	*2.18* (0.75)	1.64 (1.21)	1.09 (0.83)	1.18 (0.98)	0.91 (0.83)	*2.64* (0.5)	*2.91* (0.3)
PA 18. Player does not have a high score in the game.	*2.18* (0.98)	*2* (1)	*2.36* (1.03)	1.91 (1.04)	1.45 (0.93)	1.36 (0.92)	0.91 (0.94)	*2.27* (0.9)	1.55 (1.04)
PA 19. Player does not win a lot of games (low win ratio).	*2.36* (0.67)	1.82 (1.08)	*2.64* (0.5)	2 (1)	1.36 (0.92)	1.18 (0.87)	1 (0.89)	*2.82* (0.4)	1.64 (0.81)
PA 20. Player's scores in different games vary greatly.	*2.27* (1.1)	1.64 (1.12)	*2.27* (0.79)	*2.36* (1.12)	0.73 (0.9)	0.73 (0.9)	0.64 (0.92)	*2.18* (1.08)	1.82 (1.08)
PA 21. Player's win ratio decreases rapidly as the game's level of difficulty increases.	*2.36* (0.67)	1.91 (0.94)	*2.64* (0.67)	*2.18* (0.87)	1.18 (0.87)	1.09 (0.94)	0.82 (0.87)	*2.64* (0.81)	1.64 (1.03)

XSL•FO
RenderX

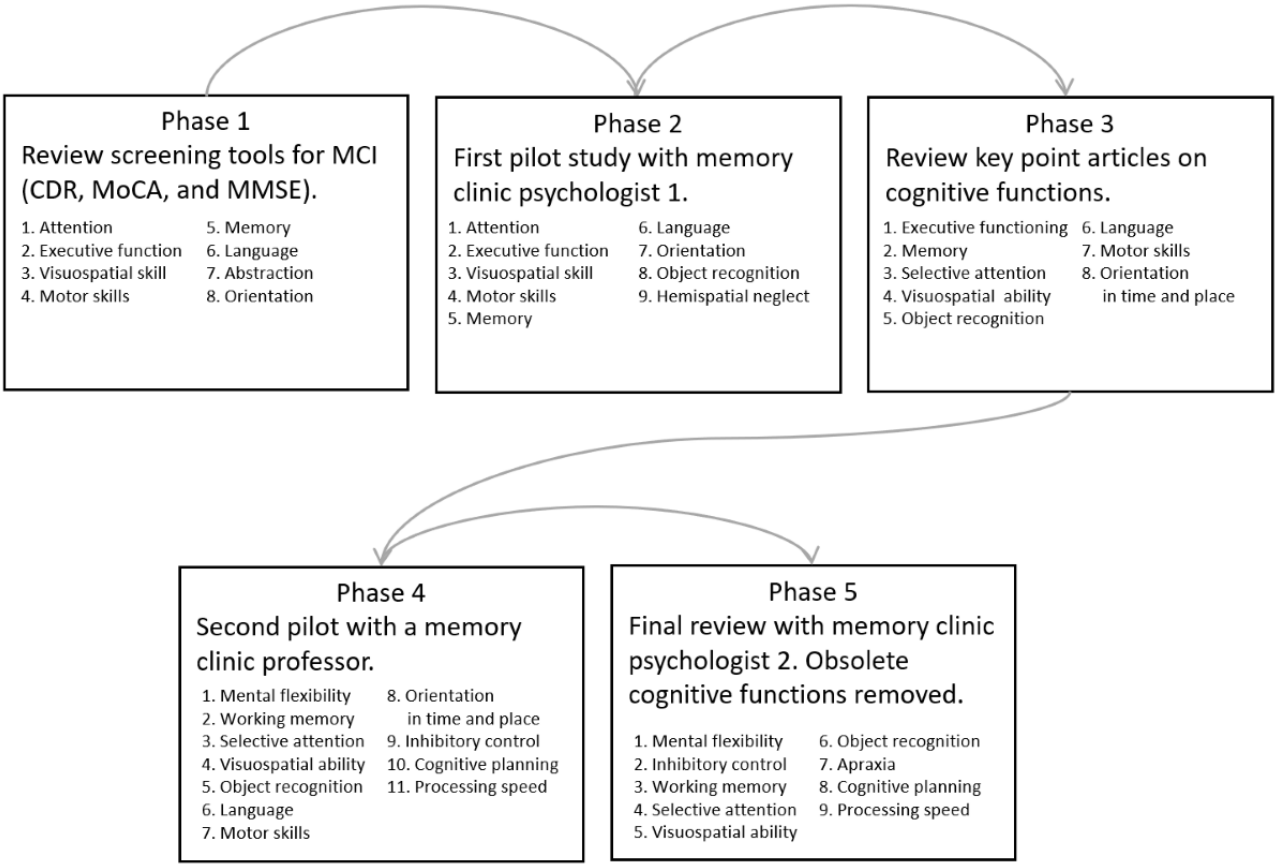
[a]PA: player action.

[b]Values scoring moderately strong are in italics.

Step 2: Defining Cognitive Functions

A set of cognitive functions was drafted in 5 phases (Figure 2). A first draft was prepared (phase 1), beginning with the cognitive functions measured using the screening tests used most often for MCI [36,37,82]. In phase 2, during a trial with a psychologist, we replaced abstraction with object recognition to more clearly indicate the problems with finding suitable cards, based on key articles on cognitive aging and cognition [83-87]. Next (phase 3), to better delineate attention, it was specified as selective attention. In phase 4, a pilot study was conducted with an expert on memory and age-related disorders (with 23 years of clinical and research experience). On the basis of this pilot testing, it was decided to split executive functioning into inhibitory control, cognitive planning, and mental flexibility. Memory was further specified as working memory and lack of motor skills as apraxia. In the final iteration, cognitive functions ostensibly not present in Klondike Solitaire, that is, orientation in time and space, as well as language, were removed to reduce the rating complexity. This resulted in a set of 9 cognitive functions (Figure 2, phase 5).

Figure 2. The 5 phases through which the cognitive functions were defined. CDR: Clinical Dementia Rating, MoCA: Montreal Cognitive Assessment, MMSE: Mini-Mental State Examination.



Protocol for Rating Functions and Actions

Overview

In the next step, the experts were asked to rate the extent to which each PA was related to a specific cognitive function.

These experts were recruited from 2 leading memory clinics in Belgium using a snowball sampling method. Of the 11 experts, 3 (27%) held Doctor of Medicine degrees and were experienced in treating cognitive decline, whereas the other 8 (73%) were neuropsychologists; 7 (64%) were women; and 4 (36%) were men. The average age of the experts was 45 (SD 13.3) years, and their average working experience was 20 (SD 14) years. In all, 3 coauthors of this paper (LVA, PD, and FFB) also participated as experts. None of the experts were compensated for participating in the study.

Before they began the rating process, the experts received a standardized introduction comprising a video that explained all concepts of the game [88], a video that visualized all 21 PAs [89], and a document that provided explanatory notes for all 9 cognitive functions. The aim of providing this introduction was to prevent confusion regarding game terminology, interpretation of PAs, and cognitive functions. The introduction also included a delineation of the target group to persons living with multiple-domain aMCI. The experts could revisit the videos and document at any time.

XSL•FO
RenderX

Next, each expert received a coding sheet in which they could map the 21 PAs to the 9 cognitive functions. Each cell had to be filled in according to the following 4-point scale:

0: This cognitive function has no significant correlation to the PA. 1: This cognitive function correlates weakly to the PA. 2: This cognitive function correlates moderately to the PA. 3: This cognitive function correlates strongly to the PA.

Finally, they were also given the choice to explain their train of thought in an optional clarification column.

Expert Agreement on PAs and Cognitive Functions

The intraclass correlation (ICC) for each PA as variables of interest with cognitive functions was computed. In addition, we computed ICCs for each of the cognitive functions as variables of interest, with all PAs considered as observations [90]. All calculations were executed using SPSS software (version 23.0; IBM Corp) [91]. The ICC was calculated to verify the rater agreement [92] on PAs and cognitive functions based on a 2-way, random, fully crossed design with type consistency [93]. According to the criteria described in the study by Koo et al [94], ICCs below 0.5 are indicative of low reliability, ICCs between 0.5 and 0.75 are indicative of moderate reliability, ICCs between 0.75 and 0.9 are indicative of good reliability, and ICCs above 0.9 are indicative of excellent reliability.

We found that the ICCs for all PAs scored higher than 0.75, suggesting good to excellent reliability according to the study by Koo et al [94]. With the exception of 4 cognitive functions (ie, mental flexibility, visuospatial ability, object recognition, and apraxia with scores of 0.68, 0.42, 0.66, and 0.71, respectively), the ICCs of the cognitive functions scored higher than 0.75, suggesting good to excellent reliability.

Cognitive Functions Present in Klondike Solitaire

An overview of the associations between individual PAs and cognitive functions, according to the expert mapping, is presented in Table 1. In addition, an overview of all ICCs with 95% CIs is presented in Multimedia Appendix 1. Each PA was related by the experts to one or more cognitive functions with an average association above two, which indicated a moderate to strong relationship with the cognitive functions. Similarly, each cognitive function was associated with at least one PA, with an average association of more than two.

Step 3: Defining Candidate Digital Biomarkers

These PAs were captured through the game as potential digital biomarkers, that is, measurable factors of the game, such as score, duration of the game, and detailed moves. These candidate digital biomarkers were enriched with additional information about the game. This contextualization is important to ensure an unambiguous interpretation of the cognitive information derived from the gameplay. For example, whereas a game played with many moves on the pile may indicate that a player progressed in the game, it may indicate equally that the player did not realize that they were stuck. By calculating the percentage of pile moves by dividing the number of pile moves by the total number of moves, a more informative metric can be obtained. In this manner, 23 potential digital biomarkers were defined that we further classified into 1 of 6 categories: time-based, performance-based, error-based, execution-based, auxiliary-based, and result-based. Time-based digital biomarkers are biomarkers related to the speed of PAs. Performance-based digital biomarkers are biomarkers related to optimal gameplay (ie, if the game was played according to strategies that ensure optimal performance). Error-based digital biomarkers relate to making incorrect moves according to the rules of Klondike Solitaire. Auxiliary-based digital biomarkers are interactions that are not part of the core gameplay (ie, requesting undo moves and hints). Execution-based digital biomarkers relate to the accuracy in moving cards and the presence of accidental taps. Finally, result-based digital biomarkers are biomarkers that evaluate the final outcome of the game (eg, how far did the participant get in the game). An overview of all digital biomarkers and their contextualizations is presented in Table 2.

XSL•FO
RenderX

Table 2. Digital biomarkers related to the player actions in Klondike Solitaire[a].

Related PA[b]	Digital biomarker	Description	Contextualization	Value
Time-based				
PA 1	Think time	Time spent thinking of a move. Defined as the time necessary to find and touch a suitable card	Average (SD)	Number (ms)
PA 2	Move time	Time spent moving cards. Defined as the time necessary to move a suitable card to the destination	Average (SD)	Number (ms)
PA 1, PA 2	Total time	Total time to make a move. Defined as the combination of think time and move time	Average (SD)	Number (ms)
Performance-based				
PA 3	Final β error	Whether there were still moves possible when quitting a game	None	Boolean
PA 3, PA 4, PA 5, PA 6	β error	Number of pile moves made with moves remaining on the board divided by the total number of pile moves	Percentage	0%-100%
PA 7	Ace β error	Number of missed opportunities to place an ace on the suit stacks divided by the total number of game moves	Percentage	0%-100%
PA 8	King β error	Number of missed opportunities to place a king on an empty spot divided by the total number of game moves	Percentage	0%-100%
PA 10	Pile move	Number of pile moves divided by the total number of board moves	Percentage	0%-100%
Error-based				
PA 11, PA 12	Successful move	Number of successful moves divided by the total number of board moves	Percentage	0%-100%
PA 11, PA 12	Erroneous move	Number of erroneous moves divided by the total number of board moves	Percentage	0%-100%
Execution-based				
PA 13	Accuracy	Accurateness of selecting a card, defined by how close to the center a card was touched	Average (SD)	0%-100%
PA 14	Taps	Number of actuations on nonuser interface elements	None	Number
Auxiliary-based				
PA 15	Undo move	Number of undo moves requested.	Percentage	0%-100%
PA 16	Hint move	Number of hints requested	Percentage	0%-100%
Result-based				
PA 17	Game time	Total time spent playing a game	None	Number (ms)
PA 18	Score	Final score of a game	None	Number
PA 19	Solved	Whether the game was completed. Indicator of how successfully the game was played	None	Boolean

XSL•FO
RenderX

Related PA[b]	Digital biomarker	Description	Contextualization	Value
PA 19	Cards moved	Number of cards selected for each move. An additional indicator of how successfully the game was played; as games of Klondike Solitaire progress, longer stacks of cards are moved	Average (SD)	Number

[a]Player actions (PAs) 20 and 21 were not captured because the single-point-in-time setup would not allow a comparison of scores and win ratios with ranging difficulty over time. In addition, PA 9 was not tested because the current software would not allow for detection of these moves.

[b]PA: player action.

Evaluating Digital Biomarkers

The aim of this second study is to explore the potential of these candidate digital biomarkers of cognitive performance. Relying on 46 participants, we captured data and performed a GLMM analysis to examine the differences between healthy participants and those diagnosed with MCI.

Participants

In total, 46 participants—23 (50%) healthy participants and 23 (50%) with MCI—were enrolled. The older adults with MCI were recruited by 2 leading memory clinics in Belgium. Healthy participants were recruited from multiple senior citizen organizations, using a snowball sampling method. All healthy participants were aged 65 years or older, were fluent in written and verbal Dutch, had 20/20 (corrected) vision, no motor impairments, and lived independently or semi-independently at home, in a service flat, or at a care home. The exclusion criteria for healthy participants were subjective-memory concerns expressed by the participant, caretaker, or clinician. In addition, they were screened using the MMSE, MoCA, and CDR Scale. To minimize the risk of including potential individuals with MCI among the healthy participants, cutoff scores of 27 on the MMSE, 26 on the MoCA, and 0 on the CDR Scale were enforced. The participants living with MCI had been formally diagnosed with multiple-domain aMCI by 1 of the 2 collaborating memory clinics, based on the diagnostic criteria described in the study by Petersen [95]. Participants with MCI were excluded if they scored lower than 23 on the MMSE to avoid including participants who could be on the borderline between the diagnosis of MCI and dementia. All recruited participants had prior experience with Klondike Solitaire. This familiarity with the rules was imperative because participants with MCI may have problems with memorizing and recalling new game rules because of short-term memory issues. Moreover, the rationale underlying this study is that it is imperative to draw from games already played and enjoyed by participants and where the rules are crystallized in their memory. Demographic and basic neuropsychological data of both groups are presented in Table 3.

XSL•FO
RenderX

Table 3. Demographic and neuropsychological data (N=46).

	Healthy participants (n=23)	Participants diagnosed with MCI[a] (n=23)
Age (years), mean (SD)	70 (5.4)	80 (5.2)
Education (ISCED[b] level) [96], n (%)		
Levels 1-2	5 (22)	4 (17)
Levels 3-4	7 (30)	13 (57)
Levels 5-6	11 (48)	6 (26)
Sex, n (%)		
Female	11 (47)	13 (57)
Male	12 (53)	10 (43)
Tablet proficiency, n (%)		
Daily	12 (52)	3 (13)
Weekly	2 (9)	2 (9)
Monthly	0 (0)	2 (9)
Yearly or less	2 (9)	1 (4)
Never	7 (30)	15 (65)
Klondike Solitaire proficiency, n (%)		
Daily	3 (13)	7 (30)
Weekly	6 (26)	8 (35)
Monthly	3 (13)	2 (9)
Yearly or less	11 (47)	6 (26)
Never	0 (0)	0 (0)
MMSE[c] score, mean (SD)	29.61 (0.65)	26.17 (1.75)
MoCA[d] score, mean (SD)	28.09 (1.28)	N/A[e]
CDR[f] Scale score, mean (SD)	0 (0)	N/A

[a]MCI: mild cognitive impairment.
[b]ISCED: International Standard Classification of Education.
[c]MMSE: Mini-Mental State Examination.
[d]MoCA: Montreal Cognitive Assessment.
[e]N/A: not applicable.
[f]CDR: Clinical Dementia Rating.

Data Collection Tools

All game sessions were completed on a Lenovo Tab3 10 Business tablet (Lenovo Group Ltd) running Android 6.0 (Google LLC). A Solitaire app created by Bielefeld [97] under the Lesser General Public License 3 was modified to capture and store game metrics that served as building blocks for the digital biomarkers of cognitive performance.

Data Collection Procedure

Each observation was made between 9 AM and 5 PM in the home environment of the participant to ensure a familiar and comfortable environment. Each observation took between 2 and 3 hours and consisted of 2 main parts:

1. a game session where game-based digital biomarkers of Klondike Solitaire were collected on a tablet and
2. a neuropsychological examination where cognitive information was collected.

Each game session started with a standardized 5-minute introduction during which the tablet, the game mechanics, and possible touch interactions were explained. This was followed by a practice game, identical for all participants, where questions to the researcher were allowed and the participant could get used to the touch controls. Data from this practice game were not used for analysis. After this practice game, each participant played 3 games in succession. The order and games were identical across all participants. All games were purposefully chosen through prior playtesting, in that they were solvable and varied in difficulty level. During these 3 games, no questions were allowed, and gameplay continued until the participants either finished the game or indicated that they deemed further progress impossible. All game sessions and cognitive

XSL•FO
RenderX

evaluations were conducted by the same researcher to avoid differences arising from researcher bias.

Ethical Statement

This study is in accordance with the declaration of Helsinki and General Data Protection Regulation compliant. Ethical approval was provided by the ethics committee of UZ/KU Leuven, Belgium (CTC S59650). Because of the fragile nature of our participants' health, utmost care was taken when providing information to them about the game sessions. The tests were conducted only after we received written informed consent.

Statistical Analysis

To assess the difference between the healthy participants and those diagnosed with MCI, a GLMM analysis was performed using R software (The R Foundation for Statistical Computing) [98] with the lme4 library [99]. Concerning the design of our GLMM, the fixed effects consisted of MCI, age, tablet proficiency, and Klondike Solitaire proficiency. The random effects were modeled as random intercepts for game seed and participant. In addition, by-participant random slopes for the effect of MCI were modeled.

Continuous digital biomarkers (eg, think time average) were modeled using a GLMM with the identity link function. Binary outcomes (eg, solved or not solved) were modeled using a GLMM with the logit link function. The significance of the effect of MCI was determined using the likelihood ratio test, which compares the model with a model without the effect of MCI, both estimated without restricted maximum likelihood [100,101]. Assumptions of homoscedasticity and normality were visually inspected using residual plots. To provide supplemental information on the fit of the models, the marginal R^2 and the conditional R^2 were given as specified in the study by Nakagawa and Schielzeth [102]. Given the exploratory nature of this study, we did not correct for family-wise inflation error [103].

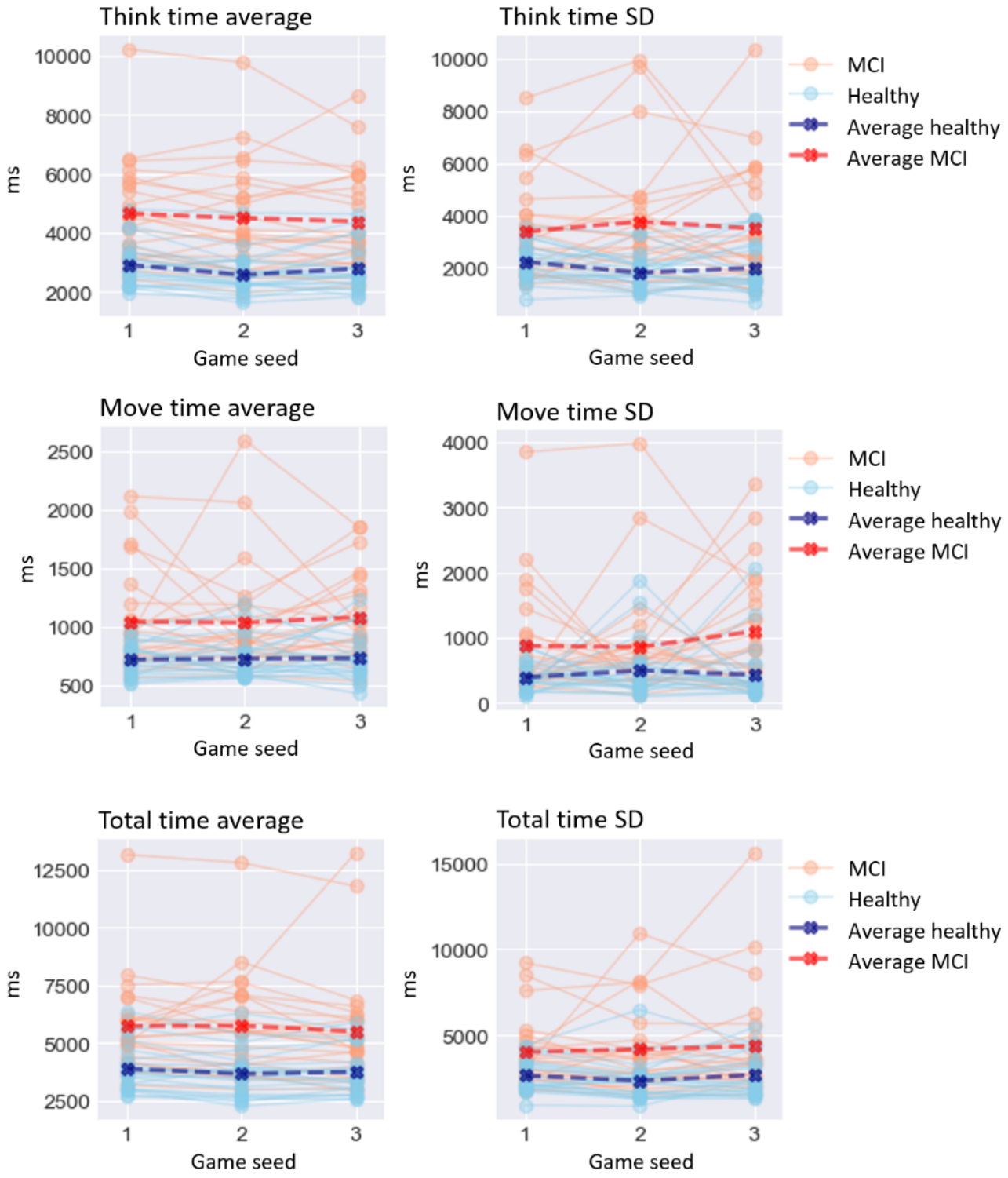
Results

Overview

The results of the GLMM analysis on the effect of MCI are presented below. A visualization of digital biomarker performance for all groups across all games is presented in Figures 3-8. A summary is presented in Table 4.

XSL•FO
RenderX

Figure 3. Performance on time-based digital biomarkers for both groups. MCI: mild cognitive impairment.



XSL•FO
RenderX

Table 4. Generalized linear mixed model analysis results for each digital biomarker.

Digital biomarker	Value, constant (SD)	Value, β (SD)	*P* value (χ^2)	Value, R^2m (R^2c)
Time-based				
Think time average	–1371.778 (1415.444)	1119.947 (405.815)	.006	0.416 (0.904)
Think time SD	–814.527 (1720.073)	1112.533 (490.53)	.02	0.211 (0.655)
Move time average	–508.575 (373.89)	156 (95.547)	.1	0.257 (0.579)
Move time SD	–856.605 (847.852)	323.599 (202.032)	.1	0.137 (0.419)
Total time average	–912.419 (2149.177)	1278.263 (573.839)	.02	0.318 (0.870)
Total time SD	206.569 (2676.062)	1315.598 (673.665)	.04	0.176 (0.715)
Performance-based				
Final β error	–7.233 (4.131)	0.435 (0.922)	.65	0.096 (0.068)
β error percentage	–7.203 (33.849)	6.108 (6.879)	.36	0.089 (0.371)
Ace β error percentage	–0.132 (0.629)	0.051 (0.137)	.73	0.023 (0.209)
King β error percentage	–3.682 (5.918)	0.907 (1.323)	.48	0.028 (0.230)
Pile move percentage	71.759 (24.052)	13.333 (4.88)	.006	0.097 (0.513)
Error-based				
Successful move percentage	87.486 (9.443)	–8.913 (3.595)	.02	0.104 (0.795)
Erroneous move percentage	9.486 (6.529)	3.624 (1.651)	.03	0.081 (0.466)
Execution-based				
Accuracy average	92.134 (9.167)	–3.817 (1.903)	.04	0.246 (0.805)
Accuracy SD	4.519 (3.746)	0.137 (0.772)	.85	0.056 (0.196)
Taps	–5.113 (10.704)	5.334 (2.762)	.05	0.098 (0.500)
Auxiliary-based				
Undo move percentage	0.228 (0.955)	0.135 (0.205)	.49	0.008 (0.151)
Hint move percentage	–0.58 (0.87)	–0.311 (0.204)	.12	0.046 (0.491)
Result-based				
Gametime	–167427.7 (187325.5)	93211.27 (53699.25)	.08	0.198 (0.690)
Score	29.03 (1389.752)	–744.433 (286.576)	.009	0.105 (0.612)
Solved	–2.954 (4.578)	–2.63 (1.007)	.008	0.186 (0.152)
Cards moved average	1.111 (0.262)	–0.119 (0.054)	.03	0.061 (0.093)
Cards moved SD	0.135 (0.705)	–0.38 (0.147)	.009	0.072 (0.152)

Time-Based Digital Biomarkers

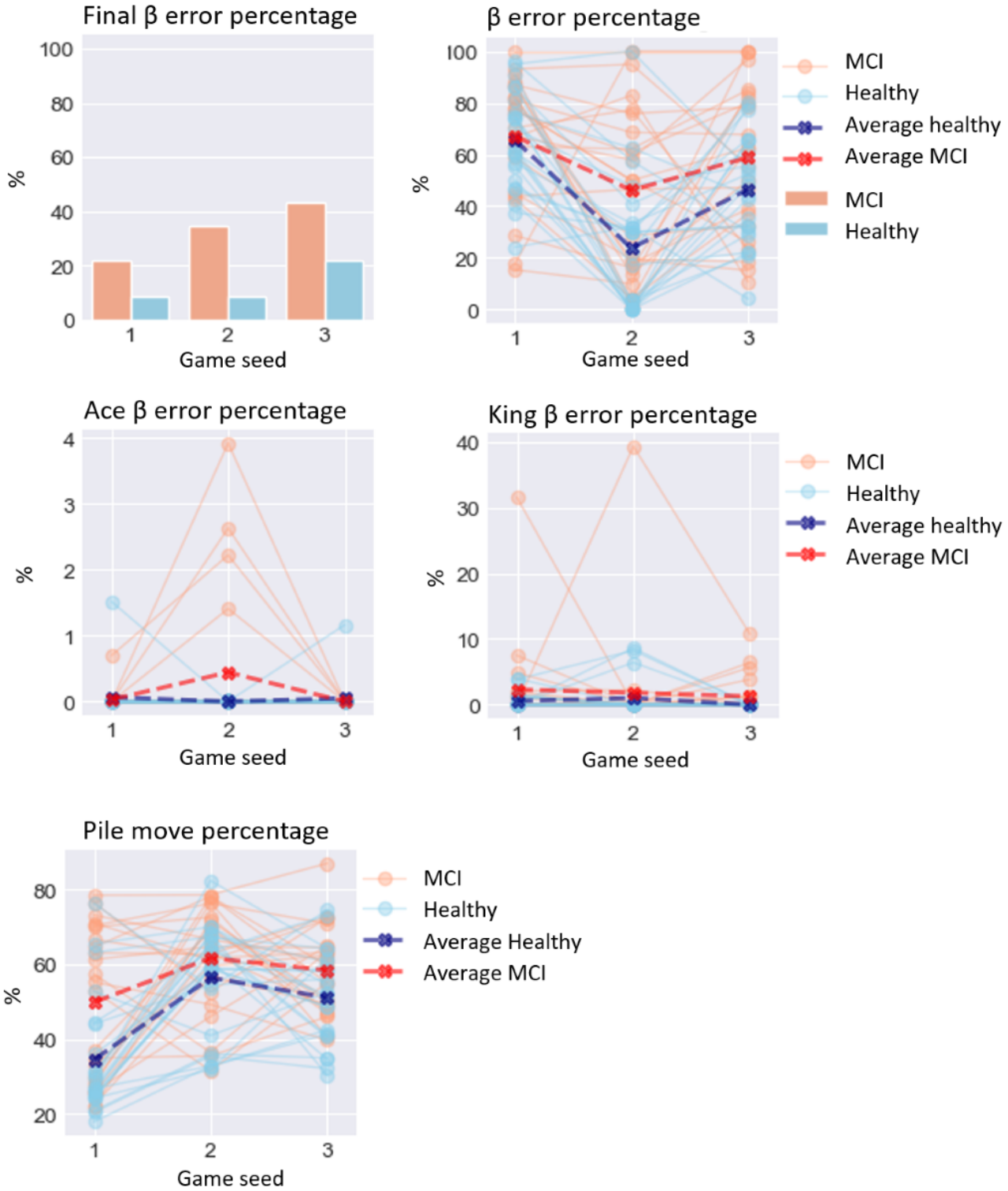
For time-based digital biomarkers (Figure 3), MCI significantly affected think time average (χ^2_1=7.7; *P*=.006), increasing it by 1119.947 ms (SD 405.81). Equally significantly, MCI affected think time SD (χ^2_1=5.1; *P*=.02), increasing it by 1112.533 ms (SD 490.53). However, MCI did not significantly affect move time average (χ^2_1=2.7; *P*=.10) or move time SD (χ^2_1=2.6; *P*=.10). MCI significantly affected total time average (χ^2_1=5.2; *P*=.02), increasing it by 1278.263 ms (SD 573.84), and total time SD (χ^2_1=4.1; *P*=.04), increasing it by 1315.598 ms (SD 673.67).

Performance-Based Digital Biomarkers

For performance-based digital biomarkers (Figure 4), MCI did not significantly affect final β error percentage (χ^2_1=0.2; *P*=.65). Equally, MCI did not significantly affect β error percentage (χ^2_1=0.8; *P*=.36), ace β error percentage (χ^2_1=0.1; *P*=.73), or king β error percentage (χ^2_1=0.5; *P*=.48). MCI significantly affected pile move percentage (χ^2_1=7.5; *P*=.006), increasing it by 13.333% (4.88).

XSL•FO
RenderX

Figure 4. Performance on performance-based digital biomarkers for both groups. MCI: mild cognitive impairment.

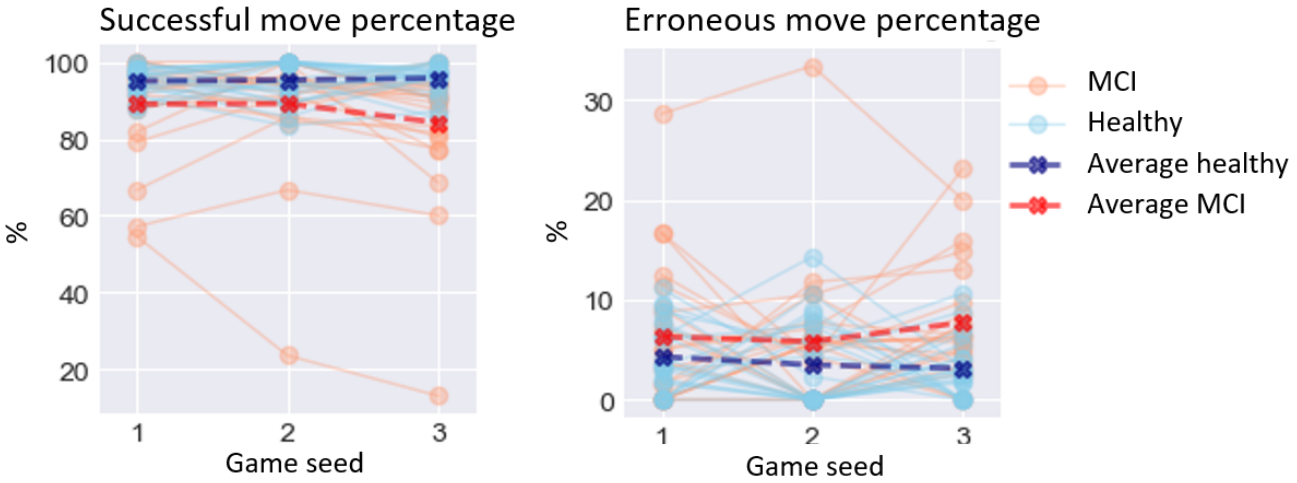


Error-Based Digital Biomarkers

For error-based digital biomarkers (Figure 5), MCI significantly affected successful move percentage, (χ^2_1=5.9; *P*=.02), lowering it by 8.913% (SD 3.6). MCI also significantly affected erroneous move percentage, (χ^2_1=4.8; *P*=.03), increasing it by 3.624% (SD 1.65).

XSL•FO
RenderX

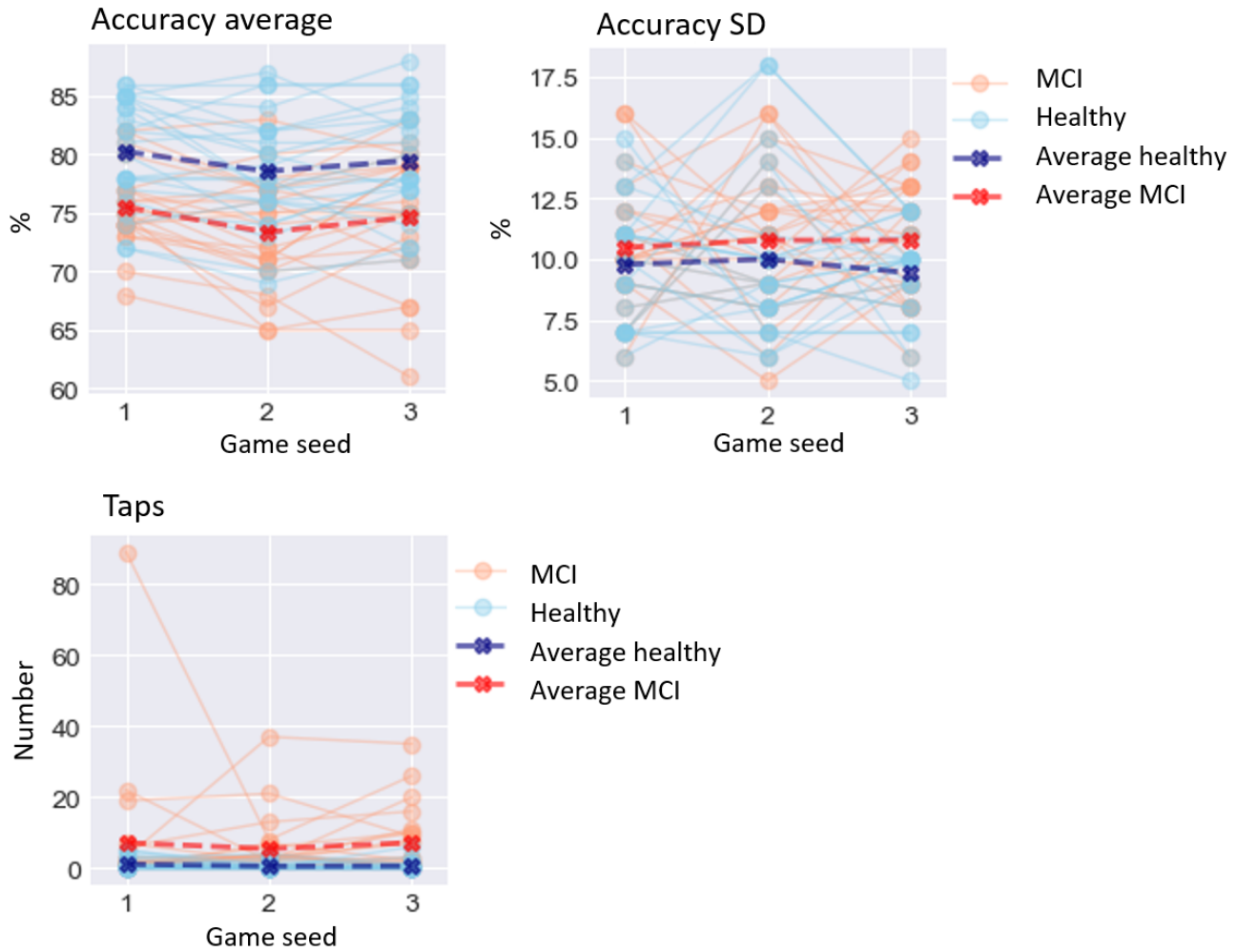
Figure 5. Performance on error-based digital biomarkers for both groups. MCI: mild cognitive impairment.



Execution-Based Digital Biomarkers

For execution-based digital biomarkers (Figure 6), MCI significantly affected accuracy average (χ^2_1=4.1; *P*=.04), lowering it by 3.817% (SD 1.9). MCI did not significantly affect accuracy SD (χ^2_1=0.04; *P*=.85) or taps (χ^2_1=3.8; *P*=.05).

Figure 6. Performance on execution-based digital biomarkers for both groups. MCI: mild cognitive impairment.



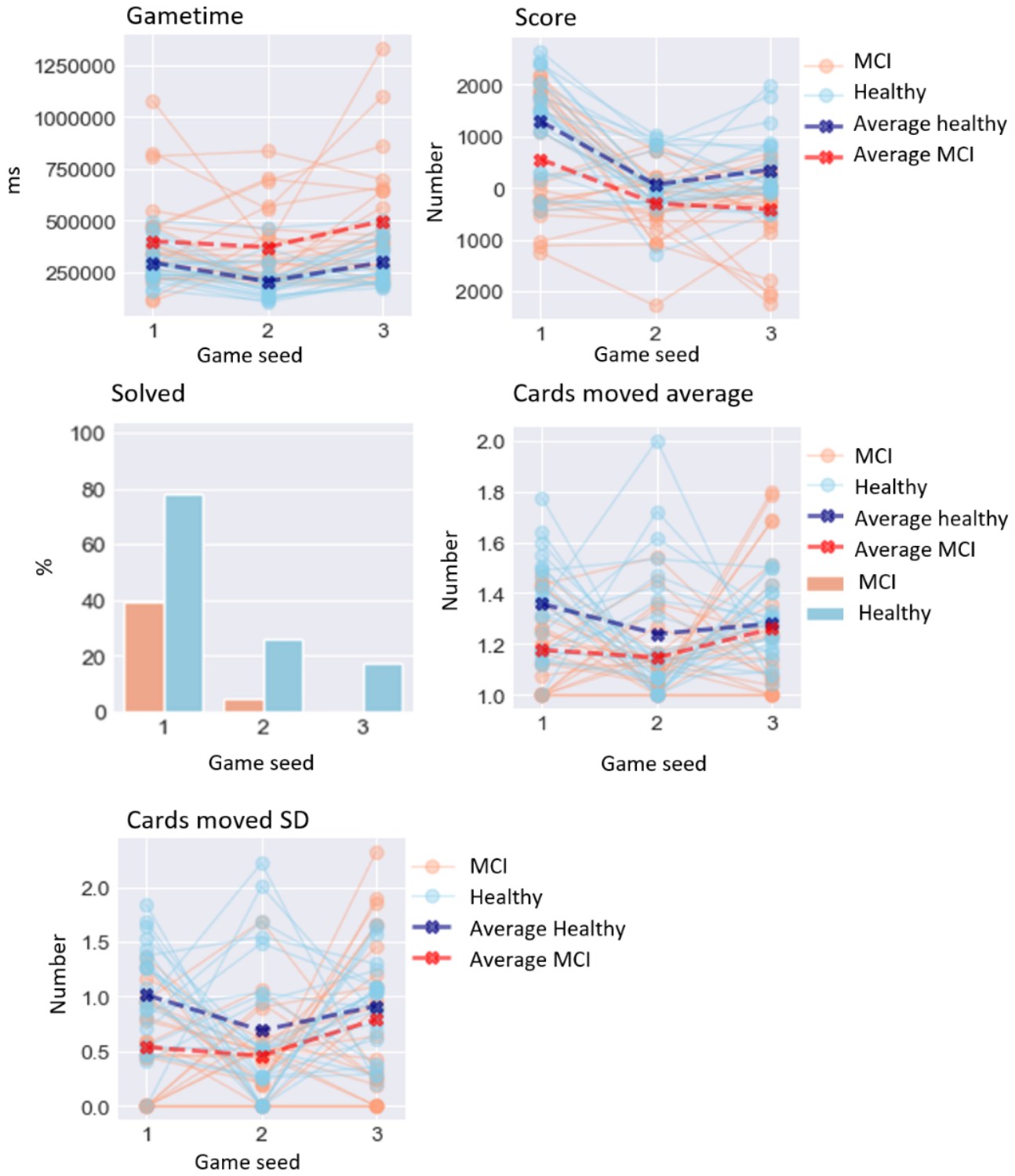
Result-Based Digital Biomarkers

For result-based digital biomarkers (Figure 7), MCI did not significantly affect gametime (χ^2_1=3.1; *P*=.08). MCI significantly affected solved (χ^2_1=6.9; *P*=.008), lowering it by 2.63 (SD 1.01). MCI also significantly affected cards moved average (χ^2_1=4.9; *P*=.03), lowering it by 0.119 cards (SD 0.05),

XSL•FO
RenderX

and cards moved SD (χ^2_1=6.7; *P*=.009), lowering it by 0.38 cards (SD 0.15).

Figure 7. Performance on result-based digital biomarkers for both groups. MCI: mild cognitive impairment.

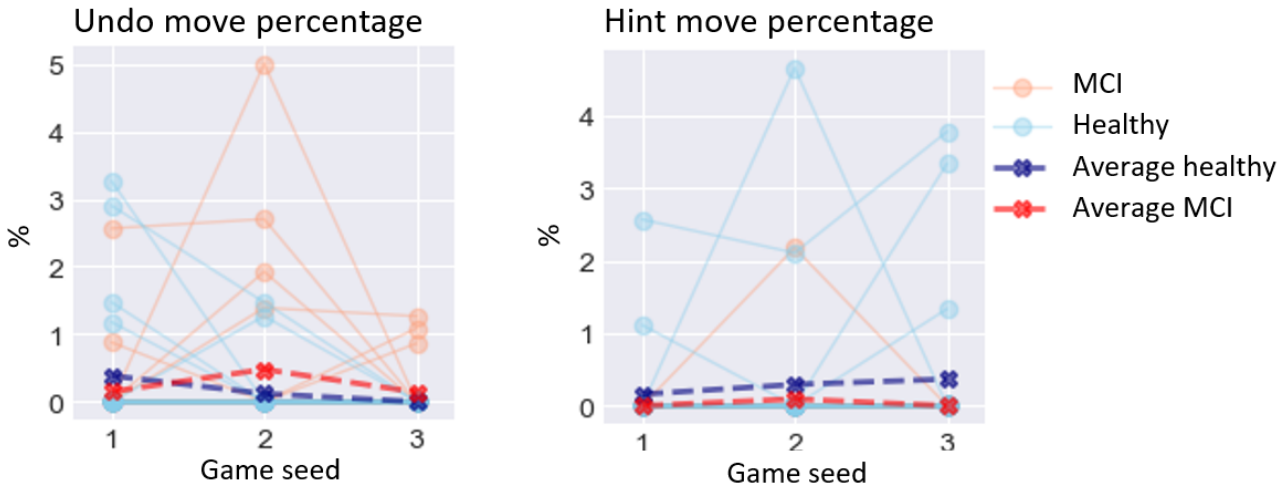


Auxiliary-Based Digital Biomarkers

With regard to auxiliary-based digital biomarkers (Figure 8), none of these candidate biomarkers reached significance: undo move percentage (χ^2_1=0.4; *P*=.49) and hint move percentage (χ^2_1=2.4; *P*=.12).

XSL•FO
RenderX

Figure 8. Performance on auxiliary-based digital biomarkers for both groups. MCI: mild cognitive impairment.



Discussion

Overview

MCI is a neurological disorder that is linked to an increased risk of developing dementia. As such, early detection of cognitive deterioration is essential for timely diagnosis and for allowing tailored care and treatment. Collecting digital biomarkers through COTS games may help by providing cognitive information through behavior traces of activities already integrated into the daily life of older adults. In this study, we investigated in particular whether Klondike Solitaire game sessions could yield digital biomarkers. In the paragraphs below, we discuss our findings and reflect on the different potential digital biomarkers, their relationship with cognitive functions, and the ethical implications of their use for cognitive assessment purposes.

Dissecting Digital Biomarkers

Of the 23 candidate digital biomarkers, 12 (52%) differed significantly between older adults with MCI and a healthy control group. This supports the use of digital card games for monitoring cognitive performance and possibly detecting differences in cognitive performance caused by MCI.

Although the overall findings are promising, not all candidate biomarkers performed equally. In the case of time-based digital biomarkers, the biomarkers related to coming up with a move—think time average and think time SD—were significantly affected by MCI. In contrast, the biomarkers related to the actual physical movement of cards—move time average and move time SD—were not significantly affected. Total time average (P=.02), which consists of move time as well as think time, was significantly affected; yet, it was less significant than think time average (P=.006). These results indicate that segmenting in-game actions can be beneficial because they can more accurately isolate cognitive functions such as praxis and cognitive planning.

In the case of performance-based digital biomarkers, in contrast with expectations, none of the biomarkers related to β errors were proven to differ significantly. Upon rewatching gameplay, it became clear that there were two different types of β errors: strategic and unintentional. However, because of the current configuration of the app, it was impossible to discriminate between the two types. This is discussed further in the *Limitations* section. In contrast, pile move percentage was proven to differ significantly. This may indicate that older adults with MCI may not recognize the same cards being returned as quickly as their healthy counterparts.

Equally, the results indicated that participants with MCI made more mistakes because both error-based digital biomarkers (ie, successful move percentage and erroneous move percentage) were significant. In contrast, none of the auxiliary-based digital biomarkers differed significantly. Upon inspecting the data, it was noted that neither group consistently used these functionalities, which may have contributed to the lack of significance.

Finally, of the 5 digital biomarkers in the result-based category, 4 (80%) were significant, 3 (60%) with P<.01 (ie, score, solved, and cards moved SD). The outcome of these measures is the result of a series of consequent moves, each of them being potentially crucial to complete the game. For example, a lapse in attention or executive functioning can cause important moves to be overlooked, in turn making the game unsolvable. Although overall gametime was not significant, this can be explained by the fact that time spent in the game by itself does not indicate a lesser performance. Time-based digital biomarkers, which are equally measures of time but contextualized with the number of moves made, show more significant results (ie, think time average, think time SD, total time average, and total time SD), stressing the importance of contextualization.

In sum, our findings are in accordance with those of the study by Jimison et al [8], which used FreeCell, a Solitaire variant, to compare cognitive performance between a group of people living with MCI and a healthy control group. Using card gameplay, we can discern older adults with MCI from a healthy control group. Moreover, the results gathered from this study are in line with those of previous studies by Bankiqued et al [60] and Ángeles Quiroga et al [104]. The study by Bankiqued et al [60] found that casual games that tap working memory and reasoning can be robustly related to performance on working memory and fluid intelligence. Similar research on commercial

XSL•FO
RenderX

video games by Ángeles Quiroga et al [104] found a strong relationship between performance in video games and general intelligence test performance. Our results confirm these findings at a finer granularity and show that when scrutinizing PAs, time-based, error-based, and result-based biomarkers yield promise in particular.

Future Work

In this study, the participants with MCI were diagnosed with multiple-domain aMCI based on the diagnostic criteria described in the study by Petersen [95]. As MCI is a multidimensional clinical entity, it would be interesting to explore whether Klondike Solitaire game sessions are suitable for monitoring the cognitive status of participants with non-aMCI as well. The focus on executive functioning can be useful for identifying both MCI subtypes because it has been shown that both have a similar decrease in executive functioning [105]. Although we acknowledge that the evaluation of other cognitive functions such as anterograde memory, retrograde memory, orientation, and language is paramount to obtain a complete overview of the patient's cognitive profile, these cognitive functions were not identified by the experts and were thus not included in our analysis.

Reflections on the Use of COTS Games to Assess Cognitive Performance

COTS games also have their limitations. First, neuropsychological assessments are typically designed to assess a broad yet targeted spectrum of cognitive functions. Moreover, different tests are devised to measure 1 cognitive function in particular. COTS games, and more particularly digital card games, were found to be more limited in terms of the cognitive functions that they can specifically assess. When using COTS games, it may be hard to separate the evaluations of specific cognitive functions. In this study, experts judged every single PA to be moderately to strongly related to at least one cognitive function.

Second, using COTS games as an instrument to measure cognitive performance and possibly flag MCI necessitates ethical reflection. We envisioned that COTS games would be used only in accordance with the informed consent of the patient, with the positive aspiration that this could aid in the longitudinal monitoring of cognitive deterioration, more accurately measuring cognitive performance and variance. This project grew out of an ambition to escape the limitations of serious games and provide meaningful play to older adults. Nonetheless, we have to acknowledge that we may have transformed an activity previously considered enjoyable and innocent into an instrumental activity that may even trigger a sense of being under health surveillance [106]. Observational notes taken during this study did not reveal any verbal remarks of stress from the participants diagnosed with MCI. However, such remarks were made by some of the healthy participants because they felt pressure to outperform the participants living with MCI. Further research is needed to understand how the instrumentalization of COTS games affects the playing experience of patients.

Third, it has to be noted that deriving digital biomarkers from digital games may not be relevant for all older adults. Not everyone is an avid gamer, and even avid gamers may have preferences for different game genres. In addition, these preferences might change over time [14]. Although digital card games such as Klondike Solitaire are in general a popular pastime for the population susceptible to MCI [61,71,107-109], they might not be so for the coming generations. Therefore, it is important to identify other accessible games suitable for cognitive monitoring with a broad appeal.

Finally, the interaction between health care professional and patient, often stimulating and motivating in and of itself, is crucial for full assessment. Hence, we argue that COTS games for screening and monitoring of MCI should not be used as a replacement for current neuropsychological examination but rather as a source of additional information.

Limitations

Fine-Tuning β Errors

In contrast with expectations, β error–related digital biomarkers proved to be insignificant. Upon inspecting the games of both groups, it became clear that there are two types of β errors: build stack β errors and suit stack β errors. The former represents missed moves among the build stacks. These errors were rarely made on purpose and occurred fewer times in the healthy participants' group, based on observation. In contrast, the latter represent missed moves between build stacks and suit stacks. We observed that this latter category was used strategically to prevent the inability to place future cards. Our observations suggest that these occurred more often in the healthy participants' group. However, because of the current configuration of the app, it was impossible to discriminate between these two types of β errors. Hence, this points to the importance of further contextualization and refinement of the measurement of β errors, and biomarkers in general, which should be addressed in future work.

Limited Sample Size

An a priori power analysis [110] estimated an adequate sample size to be between 32 and 88 participants (assuming comparable effect sizes as cognitive screening instruments to detect MCI [111]). Because of the strict inclusion criteria, only 46 participants were eligible. Although this strict protocol was designed with data quality in mind, the sample size may have affected the effects estimated in this study. It could be that our study was underpowered, leading to some digital biomarkers to be wrongly found insignificant. Future studies should therefore critically inspect the different digital biomarkers and the results obtained.

In addition, because of the average age difference between the 2 groups, we chose a GLMM for our statistical analysis because it can factor in confounding effects. A side exploration included using trained machine learning models on the same data set to predict age instead of MCI. These models were found to be less performant than the ones modeling MCI, underscoring that the effect of MCI was greater than the effect of age in our data set. Nevertheless, it is a limitation that we have to acknowledge and take into account while interpreting the results.

XSL•FO
RenderX

Conclusions

This study provides insight into the cognitive functions addressed while playing digital card games and assesses the potential of digital card game sessions for screening for MCI. To this end, 11 experts in neuropsychology or geriatrics mapped the associations of PAs in Klondike Solitaire with cognitive functions. On the basis of this exercise, which showed that the experts agreed that PAs were related to cognitive functions, 23 potential digital biomarkers of cognitive performance were crafted. A GLMM analysis, taking the effects of age, tablet experience, and Klondike Solitaire experience into account, compared digital biomarker performance between a group consisting of people living with MCI and a healthy control group. We found that of the 23 digital biomarkers, 12 (52%) had a significant and sizeable effect, despite the strict inclusion criteria and natural variations in human cognition. These exploratory results support the notion of detecting MCI through Klondike Solitaire game sessions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Intraclass correlations of the cognitive functions and player actions with 95% CIs.
[DOCX File , 15 KB - games_v9i4e18359_app1.docx]

References

1. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Practice guideline update summary: mild cognitive impairment: report of the guideline development, dissemination, and implementation subcommittee of the American Academy of Neurology. Neurology 2018 Jan 16;90(3):126-135 [FREE Full text] [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
2. Bartfay E, Bartfay WJ, Gorey KM. Prevalence and correlates of potentially undetected dementia among residents of institutional care facilities in Ontario, Canada, 2009-2011. Int J Geriatr Psychiatry 2013 Oct;28(10):1086-1094. [doi: 10.1002/gps.3934] [Medline: 23382109]
3. Cherubini A, Ruggiero C, Dell'Aquila G, Eusebi P, Gasperini B, Zengarini E, et al. Underrecognition and undertreatment of dementia in Italian nursing homes. J Am Med Dir Assoc 2012 Oct;13(8):7-13. [doi: 10.1016/j.jamda.2012.05.015] [Medline: 22727993]
4. Valcour VG, Masaki KH, Curb JD, Blanchette PL. The detection of dementia in the primary care setting. Arch Intern Med 2000 Oct 23;160(19):2964-2968. [doi: 10.1001/archinte.160.19.2964] [Medline: 11041904]
5. A dementia strategy for Canada internet. Public Health Agency of Canada. 2019. URL: https://www.canada.ca/content/dam/phac-aspc/images/services/publications/diseases-conditions/dementia-strategy/National%20Dementia%20Strategy_ENG.pdf [accessed 2021-09-15]
6. National Dementia Strategy 2014-2019. Switzerland: Federal Office of Public Health. URL: https://www.admin.ch/gov/de/start/dokumentation/medienmitteilungen.msg-id-76802.html [accessed 2021-09-15]
7. National action plan on dementia 2025. Danish Health Authorit. 2018. URL: https://www.sst.dk/-/media/Udgivelser/2018/National-Research-Strategy-on-Dementia-2025_Final.ashx [accessed 2021-09-15]
8. 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]
9. Valladares-Rodriguez S, Fernández-Iglesias MJ, Anido-Rifón L, Facal D, Pérez-Rodríguez R. Episodix: a serious game to detect cognitive impairment in senior adults. A psychometric study. PeerJ 2018 Sep 05;6:e5478 [FREE Full text] [doi: 10.7717/peerj.5478] [Medline: 30202646]
10. Leduc-McNiven K, White B, Zheng H, D McLeod R, R Friesen M. Serious games to assess mild cognitive impairment: 'The game is the assessment'. Res Rev Insights 2018;2(1):1000128. [doi: 10.15761/rri.1000128]
11. Wallace B, Goubran R, Knoefel F, Petriu M, McAvoy A. Design of games for measurement of cognitive impairment. In: Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI). 2014 Presented at: IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI); June 1-4, 2014; Valencia, Spain. [doi: 10.1109/bhi.2014.6864318]
12. Suits B, Hurka T, Newfeld F. The Grasshopper: Games, Life and Utopia, 3rd Edition. Peterborought, Ontario: Broadview Press; May 2014.
13. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. Motiv Emot 2006 Nov 29;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]
14. Mandryk RL, Birk MV. The potential of game-based digital biomarkers for modeling mental health. JMIR Ment Health 2019 Apr 23;6(4):e13485 [FREE Full text] [doi: 10.2196/13485] [Medline: 31012857]
15. Puntmann VO. How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. Postgrad Med J 2009 Oct;85(1008):538-545. [doi: 10.1136/pgmj.2008.073759] [Medline: 19789193]

XSL•FO
RenderX

16. Swarbrick S, Wragg N, Ghosh S, Stolzing A. Systematic review of miRNA as biomarkers in Alzheimer's disease. Mol Neurobiol 2019 Oct;56(9):6156-6167 [FREE Full text] [doi: 10.1007/s12035-019-1500-y] [Medline: 30734227]
17. Ahmed RM, Paterson RW, Warren JD, Zetterberg H, O'Brien JT, Fox NC, et al. Biomarkers in dementia: clinical utility and new directions. J Neurol Neurosurg Psychiatry 2014 Dec;85(12):1426-1434 [FREE Full text] [doi: 10.1136/jnnp-2014-307662] [Medline: 25261571]
18. Torous J, Rodriguez J, Powell A. The new digital divide for digital biomarkers. Digit Biomark 2017 Sep;1(1):87-91 [FREE Full text] [doi: 10.1159/000477382] [Medline: 29104959]
19. Michael DR. In: Chen S, editor. Serious Games: Games That Educate, Train and Inform. Boston, Massachusetts: Thomson Course Technology; 2006.
20. Carras MC, Van Rooij AJ, Spruijt-Metz D, Kvedar J, Griffiths MD, Carabas Y, et al. Commercial video games as therapy: a new research agenda to unlock the potential of a global pastime. Front Psychiatry 2018 Jan 22;8:300 [FREE Full text] [doi: 10.3389/fpsyt.2017.00300] [Medline: 29403398]
21. Cannon-Bowers J, Bowers C. Serious Game Design and Development: Technologies for Training and Learning Internet. Hershey, Pennsylvania: IGI Global; 2010.
22. Wouters P, van Nimwegen C, van Oostendorp H, van der Spek ED. A meta-analysis of the cognitive and motivational effects of serious games. J Educ Psychol 2013 May;105(2):249-265. [doi: 10.1037/a0031311]
23. Toril P, Reales JM, Ballesteros S. Video game training enhances cognition of older adults: a meta-analytic study. Psychol Aging 2014 Sep;29(3):706-716. [doi: 10.1037/a0037507] [Medline: 25244488]
24. Boot W, Souders D, Charness N, Blocker K, Roque N, Vitale T. The gamification of cognitive training: older adults' perceptions ofattitudes toward digital game-based interventions. In: Human Aspects of IT for the Aged Population. Design for Aging. Cham: Springer; 2016:290-300.
25. Gielis K, Verbert K, Tournoy J, Vanden AV. Age? It's in the game: an exploratory study on detection of cognitive aging through card games. In: Proceedings of the Annual Symposium on Computer-Human Interaction in Play. 2019 Presented at: CHI PLAY '19: The Annual Symposium on Computer-Human Interaction in Play; October 22 - 25, 2019; Barcelona, Spain p. 651-664. [doi: 10.1145/3311350.3347193]
26. Tangalos EG, Petersen RC. Mild cognitive impairment in geriatrics. Clin Geriatr Med 2018 Nov;34(4):563-589. [doi: 10.1016/j.cger.2018.06.005] [Medline: 30336988]
27. Kane JP, Surendranathan A, Bentley A, Barker SA, Taylor J, Thomas AJ, et al. Clinical prevalence of lewy body dementia. Alzheimers Res Ther 2018 Feb 15;10(1):19 [FREE Full text] [doi: 10.1186/s13195-018-0350-6] [Medline: 29448953]
28. Munoz D, Weishaupt N. Vascular dementia. In: The Cerebral Cortex in Neurodegenerative and Neuropsychiatric Disorders: Experimental Approaches to Clinical Issues. Cambridge, MA: Academic Press; 2017:119-139.
29. Reitz C, Brayne C, Mayeux R. Epidemiology of Alzheimer disease. Nat Rev Neurol 2011 Feb 8;7(3):137-152. [doi: 10.1038/nrneurol.2011.2]
30. Tampi RR, Tampi DJ, Chandran S, Ghori A, Durning M. Mild cognitive impairment: a comprehensive review. Healthy Aging Res 2015:A. [doi: 10.12715/har.2015.4.39]
31. Knight R, Khondoker M, Magill N, Stewart R, Landau S. A systematic review and meta-analysis of the effectiveness of acetylcholinesterase inhibitors and memantine in treating the cognitive symptoms of dementia. Dement Geriatr Cogn Disord 2018;45(3-4):131-151 [FREE Full text] [doi: 10.1159/000486546] [Medline: 29734182]
32. Webster L, Groskreutz D, Grinbergs-Saull A, Howard R, O'Brien JT, Mountain G, et al. Core outcome measures for interventions to prevent or slow the progress of dementia for people living with mild to moderate dementia: systematic review and consensus recommendations. PLoS One 2017;12(6):e0179521 [FREE Full text] [doi: 10.1371/journal.pone.0179521] [Medline: 28662127]
33. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Kokmen E, Tangelos EG. Aging, memory, and mild cognitive impairment. Int Psychogeriatr 2005 Jan 10;9(S1):65-69. [doi: 10.1017/s1041610297004717]
34. Rodakowski J, Saghafi E, Butters MA, Skidmore ER. Non-pharmacological interventions for adults with mild cognitive impairment and early stage dementia: an updated scoping review. Mol Aspects Med 2015;43-44:38-53. [doi: 10.1016/j.mam.2015.06.003] [Medline: 26070444]
35. Livingston G, Barber J, Rapaport P, Knapp M, Griffin M, Romeo R, et al. START (STrAtegies for RelaTives) study: a pragmatic randomised controlled trial to determine the clinical effectiveness and cost-effectiveness of a manual-based coping strategy programme in promoting the mental health of carers of people with dementia. Health Technol Assess 2014 Oct;18(61):1-242 [FREE Full text] [doi: 10.3310/hta18610] [Medline: 25300037]
36. Nasreddine ZS, Phillips NA, 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]
37. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. J Am Geriatr Soc 1992 Sep;40(9):922-935. [doi: 10.1111/j.1532-5415.1992.tb01992.x] [Medline: 1512391]
38. Pinto TC, Machado L, Bulgacov TM, Rodrigues-Júnior AL, Costa ML, Ximenes RC, et al. Is the montreal cognitive assessment (MoCA) screening superior to the mini-mental state examination (MMSE) in the detection of mild cognitive

XSL•FO
RenderX

impairment (MCI) and alzheimer's disease (AD) in the elderly? Int Psychogeriatr 2019 Apr;31(4):491-504. [doi: 10.1017/S1041610218001370] [Medline: 30426911]
39. Petersen RC. Mild cognitive impairment. Continuum (Minneap Minn) 2016 Apr;22(2 Dementia):404-418 [FREE Full text] [doi: 10.1212/CON.0000000000000313] [Medline: 27042901]
40. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. Int Psychogeriatr 1997;9 Suppl 1:173-176. [doi: 10.1017/s1041610297004870] [Medline: 9447441]
41. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild alzheimer disease and implications for clinical care: systematic review. J Med Internet Res 2019 Aug 30;21(8):e12785 [FREE Full text] [doi: 10.2196/12785] [Medline: 31471958]
42. Martin S, Kelly S, Khan A, Cullum S, Dening T, Rait G, et al. Attitudes and preferences towards screening for dementia: a systematic review of the literature. BMC Geriatr 2015 Jun 16;15:66 [FREE Full text] [doi: 10.1186/s12877-015-0064-6] [Medline: 26076729]
43. Krohne K, Slettebø A, Bergland A. Cognitive screening tests as experienced by older hospitalised patients: a qualitative study. Scand J Caring Sci 2011 Dec;25(4):679-687. [doi: 10.1111/j.1471-6712.2011.00878.x] [Medline: 21362007]
44. Wong S, Jacova C. Older adults' attitudes towards cognitive testing: moving towards person-centeredness. Dement Geriatr Cogn Dis Extra 2018 Oct 09;8(3):348-359 [FREE Full text] [doi: 10.1159/000493464] [Medline: 30483302]
45. Lai JM, Hawkins KA, Gross CP, Karlawish JH. Self-reported distress after cognitive testing in patients with Alzheimer's disease. J Gerontol A Biol Sci Med Sci 2008 Aug;63(8):855-859 [FREE Full text] [doi: 10.1093/gerona/63.8.855] [Medline: 18772474]
46. Summers MJ, Saunders NL. Neuropsychological measures predict decline to Alzheimer's dementia from mild cognitive impairment. Neuropsychology 2012 Jul;26(4):498-508. [doi: 10.1037/a0028576] [Medline: 22612573]
47. Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol 2007 Aug;6(8):734-746. [doi: 10.1016/S1474-4422(07)70178-3] [Medline: 17616482]
48. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
49. Weintraub S, Carrillo MC, Farias ST, Goldberg TE, Hendrix JA, Jaeger J, et al. Measuring cognition and function in the preclinical stage of Alzheimer's disease. Alzheimers Dement 2018 Feb 13;4(1):64-75. [doi: 10.1016/j.trci.2018.01.003]
50. Mané A, Donchin E. The space fortress game. Acta Psychologica 1989 Aug;71(1-3):17-22. [doi: 10.1016/0001-6918(89)90003-6]
51. 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]
52. Ferreira-Brito F, Fialho M, Virgolino A, Neves I, Miranda AC, Sousa-Santos N, et al. Game-based interventions for neuropsychological assessment, training and rehabilitation: which game-elements to use? A systematic review. J Biomed Inform 2019 Oct;98:103287. [doi: 10.1016/j.jbi.2019.103287] [Medline: 31518700]
53. 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]
54. DeRight J, Jorgensen RS. I just want my research credit: frequency of suboptimal effort in a non-clinical healthy undergraduate sample. Clin Neuropsychol 2015;29(1):101-117. [doi: 10.1080/13854046.2014.989267] [Medline: 25494327]
55. 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]
56. Schlemmer M, Desrichard O. Is medical environment detrimental to memory? A test of a white coat effect on older people's memory performance. Clin Gerontol 2018;41(1):77-81. [doi: 10.1080/07317115.2017.1307891] [Medline: 28406393]
57. Shehab A, Abdulle A. Cognitive and autonomic dysfunction measures in normal controls, white coat and borderline hypertension. BMC Cardiovasc Disord 2011 Jan 11;11:3 [FREE Full text] [doi: 10.1186/1471-2261-11-3] [Medline: 21223543]
58. Boot WR. Video games as tools to achieve insight into cognitive processes. Front Psychol 2015 Jan 21;6:3 [FREE Full text] [doi: 10.3389/fpsyg.2015.00003] [Medline: 25653634]
59. Valladares-Rodriguez S, Pérez-Rodríguez R, Anido-Rifón L, Fernández-Iglesias M. Cognitive assessment through casual video games and machine learning - doctoral consortium contributions. Doctoral Consortium - DCICT4AW, (ICT4AgeingWell 2015). 2015. URL: https://www.scitepress.org/Papers/2015/55322/55322.pdf [accessed 2021-09-15]
60. 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]

XSL•FO
RenderX

61. Tanaka Y, Shiosaka T. Gaming attitudes and habits of adults ages 50-plus. AARP Research. 2019 Dec. URL: https://www.aarp.org/content/dam/aarp/research/surveys_statistics/technology/2019/2020-gaming-trends-older-adults.doi.10.26419-2Fres.00328.001.pdf [accessed 2021-09-15]
62. Salen K, Zimmerman E. Game design and meaningful play. In: Handbook of Computer Game Studies. Cambridge, MA: MIT Press; 2005.
63. De Schutter B. Never too old to play: the appeal of digital games to an older audience. Games Cult 2010 May 07;6(2):155-170. [doi: 10.1177/1555412010364978]
64. Pavel M, Jimison H, Hagler S, McKanna J. Using behavior measurement to estimate cognitive function based on computational models. In: Cognitive Informatics in Health and Biomedicine. Cham: Springer; 2017:137-163.
65. Escandon A, Al-Hammadi N, Galvin JE. Effect of cognitive fluctuation on neuropsychological performance in aging and dementia. Neurology 2010 Jan 19;74(3):210-217 [FREE Full text] [doi: 10.1212/WNL.0b013e3181ca017d] [Medline: 20083796]
66. Dorsey E, Papapetropoulos S, Xiong M, Kieburtz K. The first frontier: digital biomarkers for neurodegenerative disorders. Digit Biomark 2017 Jul 04;1(1):6-13 [FREE Full text] [doi: 10.1159/000477383] [Medline: 32095743]
67. Grabbe JW. Sudoku and working memory performance for older adults. Act Adapt Aging 2011 Jul;35(3):241-254. [doi: 10.1080/01924788.2011.596748]
68. Thompson O, Barrett S, Patterson C, Craig D. Examining the neurocognitive validity of commercially available, smartphone-based puzzle games. Psychology 2012 Jul;03(07):525-526. [doi: 10.4236/psych.2012.37076]
69. Joshi V, Wallace B, Shaddy A, Knoefel F, Goubran R, Lord C. Metrics to monitor performance of patients with mild cognitive impairment using computer based games. In: Proceedings of the 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI). 2016 Presented at: 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI); Feb. 24-27, 2016; Las Vegas, NV, USA. [doi: 10.1109/bhi.2016.7455949]
70. Morehead A, Mott-Smith G. The Complete Book of Solitaire and Patience Games: The Most Comprehensive Book of Its Kind: Over 225 Games. New York, USA: Bantam; 1983.
71. Boot W, Moxley J, Roque N, Andringa R, Charness N, Czaja S, et al. Exploring older adults' video game use in the PRISM computer system. Innov Aging 2018 Jan 24;2(1):igy009 [FREE Full text] [doi: 10.1093/geroni/igy009] [Medline: 30480133]
72. Bjarnason R, Fern A, Tadepalli P. Lower bounding klondike solitaire with monte-carlo planning. In: Proceedings of the Nineteenth International Conference on International Conference on Automated Planning and Scheduling. 2009 Presented at: Nineteenth International Conference on International Conference on Automated Planning and Scheduling; September 19 - 23, 2009; Thessaloniki Greece p. 26-33 URL: https://dl.acm.org/doi/10.5555/3037223.3037228
73. Blake C, Gent I. The winnability of klondike and many other single-player card games. ArXiv. 2019. URL: https://arxiv.org/abs/1906.12314 [accessed 2021-10-04]
74. Goadrich MH, Droscha J. Improving solvability for procedurally generated challenges in physical solitaire games through entangled components. IEEE Trans Games 2020;12(3):260-269. [doi: 10.1109/tg.2019.2918223]
75. Jilani R, Crampton A, Kitchin D, Vallati M. Have a little patience: let planners play cards. In: Proceedings of the 34th Workshop of the UK Planning and Scheduling Special Interest Group. 2017 Presented at: 34th Workshop of the UK Planning and Scheduling Special Interest Group, PlanSIG 2016; December 15-16, 2016; Huddersfield, UK URL: http://ceur-ws.org/Vol-1782/paper_3.pdf
76. Longpré L, McKenzie P. The complexity of solitaire. Theor Comput Sci 2009 Nov 17;410(50):5252-5260. [doi: 10.1016/j.tcs.2009.08.027]
77. Roscoe AW. Card games as pointer structures: case studies in mobile CSP modelling. ArXiv. 2016. URL: https://arxiv.org/abs/1611.08418 [accessed 2021-10-04]
78. Yan X, Diaconis P, Rusmevichientong P, Roy B. Solitaire: man versus machine. In: Saul LK, Weiss Y, Bottou L, editors. Advances in Neural Information Processing Systems 17. Cambridge, MA: MIT Press; 2005.
79. Basic klondike (solitaire) tips. CardGameHeaven. 2013. URL: https://cardgameheaven.com/basic-klondike-solitaire-tips/ [accessed 2019-10-03]
80. Winning at klondike. Semicolon Software. URL: https://semicolon.com/Solitaire/Articles/Klondike.html [accessed 2019-10-03]
81. Sandberg B. Klondike solitaire - winning strategy. BVS Development Corporation. URL: https://www.bvssolitaire.com/rules/klondike-solitaire-strategy.htm [accessed 2019-10-03]
82. Tan J, Strauss E, Sherman E. Clinical dementia rating. In: Encyclopedia of Clinical Neuropsychology. New York, NY: Springer; 2011.
83. Glisky E. Memory. In: Encyclopedia of Clinical Neuropsychology. Cham: Springer International Publishing; 2017.
84. Harada CN, Natelson Love MC, Triebel KL. Normal cognitive aging. Clin Geriatr Med 2013 Nov;29(4):737-752 [FREE Full text] [doi: 10.1016/j.cger.2013.07.002] [Medline: 24094294]
85. Diamond A. Executive functions. Annu Rev Psychol 2013;64:135-168 [FREE Full text] [doi: 10.1146/annurev-psych-113011-143750] [Medline: 23020641]
86. Verhaeghen P, Cerella J. Aging, executive control, and attention: a review of meta-analyses. Neurosci Biobehav Rev 2002 Nov;26(7):849-857. [doi: 10.1016/s0149-7634(02)00071-4] [Medline: 12470697]

XSL•FO
RenderX

87. Owen AM. Cognitive planning in humans: neuropsychological, neuroanatomical and neuropharmacological perspectives. Prog Neurobiol 1997 Nov;53(4):431-450. [doi: 10.1016/s0301-0082(97)00042-7] [Medline: 9421831]
88. Gielis K. ExplanationSolitaire_Short. Figshare. 2019. URL: https://figshare.com/articles/media/ExplanationSolitaire_Short/10304780 [accessed 2021-09-15]
89. Gielis K. PlayerActionsSolitaire_Short. Figshare. 2019. URL: https://figshare.com/articles/media/PlayerActionsSolitaire_Short/10304801 [accessed 2021-09-15]
90. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [doi: 10.1037//0033-2909.86.2.420] [Medline: 18839484]
91. IBM SPSS Statistics 23 Documentation. IBM Corp. 2015. URL: https://www.ibm.com/support/pages/ibm-spss-statistics-23-documentation [accessed 2019-11-14]
92. von Eye A, Young Mun E. Analyzing Rater Agreement Manifest Variable Methods. Milton, UK: Taylor & Francis Group; 2014.
93. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8(1):23-34 [FREE Full text] [doi: 10.20982/tqmp.08.1.p023] [Medline: 22833776]
94. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
95. 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]
96. International Standard Classification of Education, ISCED 1997. United Nations Educational, Scientific and Cultural Organization (UNESCO) and UNESCO Institute for Statistics (UIS). 2003. URL: http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf [accessed 2021-09-15]
97. Bielefeld T. Android Simple Solitaire game collection. Contribute to TobiasBielefeld/Simple-Solitaire development by creating an account on GitHub. URL: https://github.com/TobiasBielefeld/Simple-Solitaire [accessed 2019-08-16]
98. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/ [accessed 2021-09-15]
99. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft 2015;67(1):A. [doi: 10.18637/jss.v067.i01]
100. Harrison X, Donaldson L, Correa-Cano M, Evans J, Fisher D, Goodwin C, et al. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ 2018 May 23;6:e4794 [FREE Full text] [doi: 10.7717/peerj.4794] [Medline: 29844961]
101. Winter B. Linear models and linear mixed effects models in R with linguistic applications. arXivcs. 2013 Aug. URL: http://arxiv.org/abs/1308.5499 [accessed 2020-06-09]
102. Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol Evol 2012 Dec 03;4(2):133-142 [FREE Full text] [doi: 10.1111/j.2041-210x.2012.00261.x]
103. Robertson J, Kaptein M, editors. Modern Statistical Methods for HCI. New York, NY: Springer International Publishing; 2016.
104. Quiroga MA, Escorial S, Román FJ, Morillo D, Jarabo A, Privado J, et al. Can we reliably measure the general factor of intelligence (G) through commercial video games? Yes, we can!. Intelligence 2015;53:1-7. [doi: 10.1016/j.intell.2015.08.004]
105. Csukly G, Sirály E, Fodor Z, Horváth A, Salacz P, Hidasi Z, et al. The differentiation of amnestic type MCI from the non-amnestic types by structural MRI. Front Aging Neurosci 2016 Mar 30;8:52 [FREE Full text] [doi: 10.3389/fnagi.2016.00052] [Medline: 27065855]
106. MacLean S, Hatcher S. Constructing the (Healthy) neoliberal citizen: using the walkthrough method “do” critical health communication research. Front Commun 2019 Oct 4;4:52. [doi: 10.3389/fcomm.2019.00052]
107. De SB, Vanden AV. Designing meaningful play within the psycho-social context of older adults. In: Proceedings of the 3rd International Conference on Fun and Games. 2010 Presented at: Fun and Games '10: 3rd International Conference on Fun and Games; September 15 - 17, 2010; Fun and Games '10: Fun and Games 2010 p. 84-93. [doi: 10.1145/1823818.1823827]
108. 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. Comput Human Behav 2013 Jul;29(4):1302-1306. [doi: 10.1016/j.chb.2013.01.014]
109. Essential facts about the video game industry. Entertainment Software Association. 2020. URL: https://www.theesa.com/resource/2020-essential-facts/ [accessed 2021-09-15]
110. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
111. Larner AJ. Effect size (Cohen's d) of cognitive screening instruments examined in pragmatic diagnostic accuracy studies. Dement Geriatr Cogn Dis Extra 2014 Jul 03;4(2):236-241 [FREE Full text] [doi: 10.1159/000363735] [Medline: 25177332]

XSL•FO
RenderX

Abbreviations

aMCI: amnestic mild cognitive impairment
CDR: Clinical Dementia Rating
COTS: commercial off-the-shelf
GLMM: generalized linear mixed model
ICC: intraclass correlation
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
naMCI: nonamnestic mild cognitive impairment
PA: player action

Edited by N Zary; submitted 21.02.20; peer-reviewed by P Haring, B Rooney; comments to author 12.08.20; revised version received 28.01.21; accepted 11.08.21; published 04.11.21.

Please cite as:
Gielis K, Vanden Abeele ME, De Croon R, Dierick P, Ferreira-Brito F, Van Assche L, Verbert K, Tournoy J, Vanden Abeele V Dissecting Digital Card Games to Yield Digital Biomarkers for the Assessment of Mild Cognitive Impairment: Methodological Approach and Exploratory Study
JMIR Serious Games 2021;9(4):e18359
URL: https://games.jmir.org/2021/4/e18359
doi:10.2196/18359
PMID:34734825

©Karsten Gielis, Marie-Elena Vanden Abeele, Robin De Croon, Paul Dierick, Filipa Ferreira-Brito, Lies Van Assche, Katrien Verbert, Jos Tournoy, Vero Vanden Abeele. Originally published in JMIR Serious Games (https://games.jmir.org), 04.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Examining the Academic Trends in Neuropsychological Tests for Executive Functions Using Virtual Reality: Systematic Literature Review

Euisung Kim[1], BSc, MSc; Jieun Han[1], BSc, MSc, PhD; Hojin Choi[2], MD, PhD; Yannick Prié[3], BSc, MSc, PhD; Toinon Vigier[3], BSc, MSc, PhD; Samuel Bulteau[4,5], MD, PhD; Gyu Hyun Kwon[1], BSc, MSc, PhD

[1]Graduate School of Technology and Innovation Management, Hanyang University, Seoul, Republic of Korea

[2]Department of Neurology, College of Medicine, Hanyang University, Seoul, Republic of Korea

[3]Laboratory of Digital Science of Nantes (LS2N), CNRS UMR6004, Nantes Université, Nantes, France

[4]CHU Nantes, Psychiatry Department, Nantes, France

[5]INSERM U1246, SPHERE, University of Nantes, University of Tours, Nantes, France

Corresponding Author:
Gyu Hyun Kwon, BSc, MSc, PhD
Graduate School of Technology and Innovation Management
Hanyang University
Wangsimni-ro 222
Multidisciplinary Lecture Hall 703
Seoul, 04763
Republic of Korea
Phone: 82 222202414
Email: ghkwon@hanyang.ac.kr

Abstract

Background: In neuropsychology, fully immersive virtual reality (VR) has been spotlighted as a promising tool. It is considered that VR not only overcomes the existing limitation of neuropsychological tests but is also appropriate for treating executive functions (EFs) within activities of daily living (ADL) due to its high ecological validity. While fully immersive VR offers new possibilities of neuropsychological tests, there are few studies that overview the intellectual landscape and academic trends in the research related to mainly targeted EFs with fully immersive VR.

Objective: The objective of this study is to get an overview of the research trends that use VR in neuropsychological tests and to analyze the research trends using fully immersive VR neuropsychological tests with experimental articles.

Methods: This review was carried out according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Articles were searched in three web databases using keywords related to VR, EFs, and cognitive abilities. The study was conducted in two steps, keyword analysis and in-depth systematic review. In the web database search from 2000 to 2019, 1167 articles were initially collected, of which 234 articles in the eligibility phase were used to conduct keyword analysis and a total of 47 articles were included for systematic review.

Results: In keyword analysis, the number of articles focused on dementia including the keywords “MCI,” “SCD,” and “dementia” were highlighted over the period, rather than other symptoms. In addition, we identified that the use of behavioral and physiological data in virtual environments (VEs) has dramatically increased in recent studies. In the systematic review, we focused on the purpose of study, assessment, treatment, and validation of usability and structure. We found that treatment studies and uncategorized studies including presence and cybersickness issues have emerged in the recent period. In addition, the target symptoms and range of participants were diversified.

Conclusions: There has been a continuously increasing interest in dealing with neuropsychology by using fully immersive VR. Target cognitive abilities have been diversified, as well as target symptoms. Moreover, the concept of embodied cognition was transplanted in this research area.

(JMIR Serious Games 2021;9(4):e30249) doi:10.2196/30249

XSL•FO
RenderX

KEYWORDS

virtual reality; neuropsychological test; executive function; cognitive ability; brain disorder; immersive; digital health; cognition; academic trends; neurology

Introduction

Virtual reality (VR) is a state-of-the-art technology at present. With this technology, we can create realistic worlds from the real world and generate artificial experiences in real time [1]. For the past decade, technological progress has enabled VR to become more popular through decreasing costs and increasing convenience.

With regard to defining mixed reality, Milgram and Kishino [2] suggested the concept of the reality-virtuality continuum. It is powerful to overview the dimension of reality. Considering the scope of our research, we partly adopted their perspective and focused on "immersion" to classify VR. Immersion in VR can be explained as the perception that is created by various stimuli that provide an absorbing environment so that users think they are physically in a virtual world [3]. The VR system can be divided into fully immersive and semi- or non-immersive, depending on the degree of immersion [4]. Fully immersive VR provides 3D displays (eg, head-mounted display [HMD]) that effectively make the user feel they are existing inside the virtual environment (VE) for the highest level of immersion. Non-immersive VR is based on a desktop with a flat-screen monitor with low interaction (eg, personal computer, tablet). With a large, curved monitor or projector, semi-immersive VR offers moderate immersion and interaction (eg, Kinect). Thomas et al [5], who studied the impact of immersion with a virtual avatar, said that the degree of immersion significantly influences the virtual body ownership and the feeling of presence. In addition, an enhanced effect of fully immersive VR on presence was reported by Waltemate et al [6]. In this study, we define fully immersive VR as a cover with up to a 360° screen-like Cave Automatic Virtual Environment (CAVE), dome-shaped screens, and HMDs [7].

Neuropsychological assessment is one way of examining the brain by its behavioral product [8]. Neuropsychological tests are typically used to assess/treat one's cognitive abilities and further diagnose mental illness. By using neuropsychological tests, findable neuropsychological domains are extensive, for example, attention/concentration, language, visuospatial abilities, motor coordination, and executive functions (EFs) [9].

There are some popular traditional neuropsychological tests such as the Stroop test and the trail-making task [10,11]. These tests are widely accepted in professional society. Nevertheless, there are a few questionable limitations. One of the most critical limitations that is commonly pointed out is the lack of ecological validity. It is difficult to exactly measure the cognitive demands of real-world activities of daily living (ADL) with existing traditional neuropsychological tests [12,13].

As mentioned, the progress in VR technology gives rise to the possibility of applying existing neuropsychological tests in VR as an advanced model of the tests [9]. Pen-and-paper tests and computerized tests are widely used in neuropsychology, and they are composed of a set of predefined stimuli that exist in a controlled environment [14]. It is still necessary to increase ecological validity in measuring real-world performance or cognitive impairment [15-18]. However, VR environments can provide a realistic experience by multisensory stimulation and interactive factors as in daily life [19,20], with a strong sense of presence ("being there") [21]; the means to test multitasking ability [22]; the influence of distractors that may not be used in a real-world laboratory [22]; and a good level of motivation of participants [23]. By using VR neuropsychological tests, it is possible to measure and evaluate cognitive abilities and EFs in a daily-living situation, even the instrumental activities of daily living (IADL) [24], and also, collecting behavioral data is enabled. IADL indicate activities using instruments that allow an individual to improve their quality of life, such as cooking, cleaning, and managing finance. There are many previous studies that have reported the specific connection between IADL and EFs [25-28]. These support that VR neuropsychological tests are significant, in that the IADL situation can be experimentally implemented through VR to measure or manage related cognitive functions. Therefore, VR can be considered an alternative or complementary innovative neuropsychological tool [29].

The term "executive function" was first defined by Muriel Deutsch Lezak in 1982 [30]. She said that "executive functions comprise those mental capacities necessary for formulating goals, planning how to achieve them, and carrying out the plans effectively" and proposed four classes as capacities of (1) formulating goals, (2) planning, (3) carrying out the plan, (4) and performing effectively [30]. However, the definition and functional categories of EFs are prescribed slightly differently by other researchers. Anderson [31] described EFs as "a collection of inter-related processes responsible for purposeful, goal-directed behavior" such as "anticipation, goal selection, planning, initiation of activity, self-regulation, mental flexibility, deployment of attention, and utilization of feedback." Similarly, Hughes [32] described an EF as "a complex cognitive construct that encompasses the set of processes that underlie flexible goal-directed behavior (e.g., planning, inhibitory control, attentional flexibility and working memory)."

Therefore, there is a common agreement for categorizing three core EFs in general [33,34]: inhibition, working memory, and shifting (also called cognitive flexibility). Higher-order EFs, such as decision making (also called reasoning), problem solving, and planning, are usually established by extending from the core EFs [35,36]. Moreover, one of the core EFs, working memory, can be divided into more detail, such as general working memory, which holds information in the mind, and updating, which conducts translating instructions into action plans and involving new information into thinking or action plans [37].

In this paper, with advice from experts in neuropsychology based on the theoretical background, we categorized the EFs into seven sub-abilities:

XSL•FO
RenderX

- Inhibition: The ability to control impulsive and automatic responses and generate responses using attention and reasoning
- Working memory: The ability to temporarily store and handle information in order to do complex cognitive tasks
- Shifting: The ability to adapt your thoughts and behaviors to new, changing, and unexpected situations
- Decision making: The ability to efficiently and thoughtfully choose an option among different alternatives
- Problem solving: The ability to come to a logical conclusion when considering an unknown
- Planning: The ability to think about future events and mentally anticipate the correct way to carry out a task or reach a specific goal
- Updating: The ability to supervise behavior and ensure that you are properly carrying out the established plan of action

These days, fully immersive VR has been in the spotlight because it is considered an innovative tool to exceed the existing limitations of neuropsychological tests. Consequently, researchers who are studying in related areas have worked to improve the quality of VR and optimally apply it to neuropsychological tests. Many earlier studies have been conducted in the neuropsychological area with VR technology. However, there are only few studies that use fully immersive VR to mainly target EFs.

This systematic review article aims (1) to get an overview of the research trends using VR (including non-immersive) to conduct neuropsychological tests targeting EFs from 2000 to 2019 by keyword analysis and (2) to analyze the specific research trends using fully immersive VR with experimental articles.

Methods

Registration

This systematic review was conducted using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [38]. In this study, we present a systematic review to summarize the past studies in the neuropsychological area with VR. Because there were few studies before 2000, the range of publishing years was limited from 2000 to 2019. The earlier studies were also significant and important. However, we assumed most of the concepts and topics in those studies could reflect on future studies, since there were a negligible number of studies that were related to VR and neuropsychological tests.

Search Strategy

Initial articles were searched in three web databases: Scopus, Web of Science, and PubMed. These engines are adequate to cover the wide spectrum of topics in the target area. The articles were searched in March 2020, and the search query string was as follows: (“virtual reality” OR “VR” OR “virtual environment”) AND (“executive function(s)” OR “EF(s)” OR “cognitive function(s)” OR “cognition”). The publishing years were limited from 2000 to 2019.

Inclusion Criteria

Journal articles and conference (proceeding) papers written in English were chosen, and other forms of publication, such as short reports, chapters of books, and dissertations, were excluded. To aim more deeply at EFs among the cognitive abilities, we included articles that were only related to “executive function” and “virtual reality” were used to see the effectiveness of neuropsychological tests in VR.

Exclusion Criteria

After duplicates were removed, all studies were checked by using keywords and abstracts in the screening process. Review and protocol articles and chapters of books were all excluded. Moreover, articles not mainly focusing on EFs were excluded. At the end of the screening process, 234 studies were left.

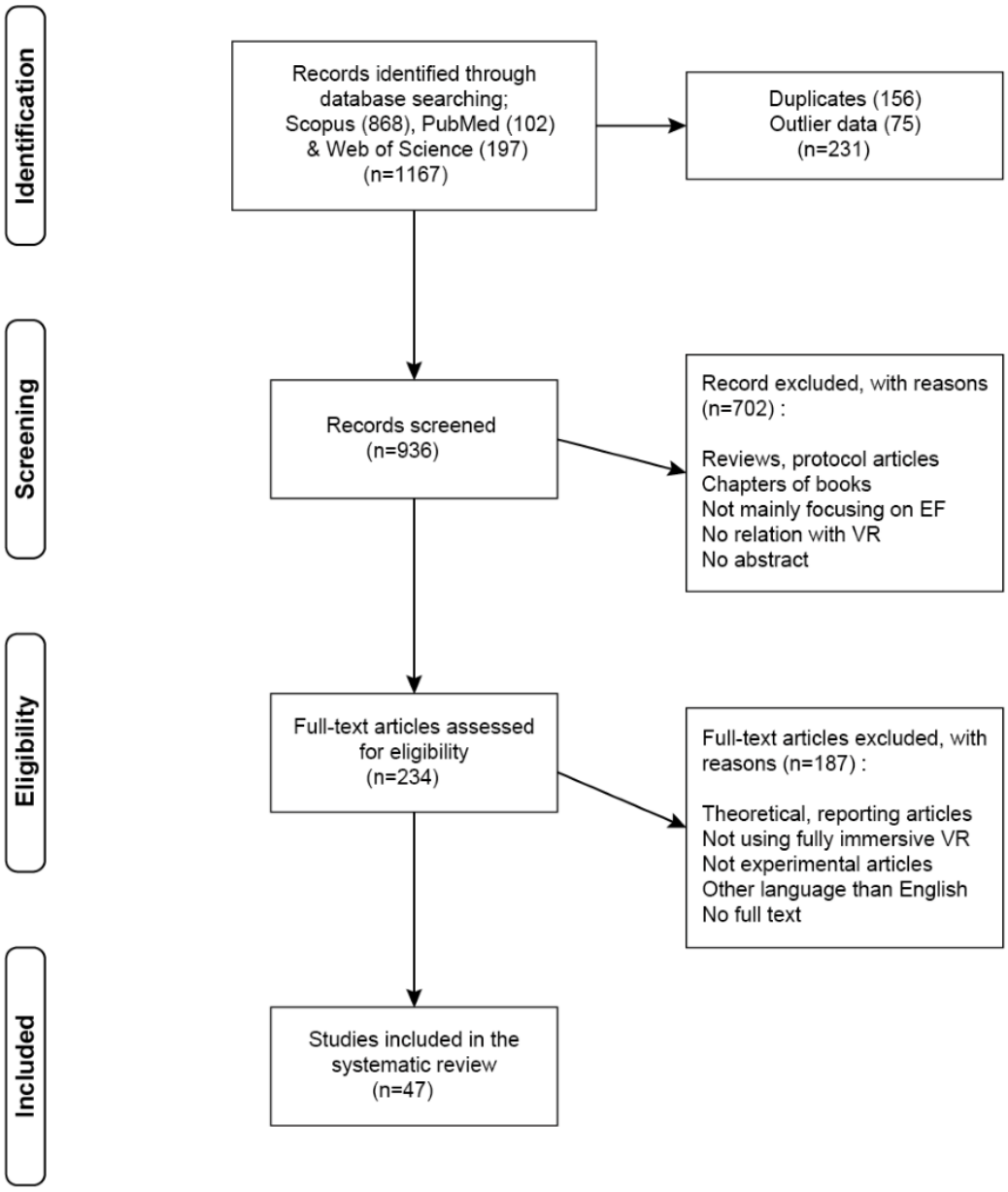
Evaluation of full-text articles was conducted during the eligibility process. In this phase, we excluded theoretical articles, reporting articles, and non-experimental articles. The term “virtual reality” commonly corresponds not only to articles about immersive environments but also to articles about non-immersive environments, such as the use of a flat-screen monitor, so this study excluded them as not fully immersive VR articles. A few records where we could not find the full text were also excluded.

Study Selection

The overall procedure of the study and the number of selected articles are shown in Figure 1. After searching the web database, 1167 articles were collected (Scopus 868, PubMed 102, and Web of Science 197). After removing duplicates and outliers, the remaining articles were used in the next phase. As a result of the screening phase, 702 articles were excluded by titles, keywords, and abstract review. With abstracts of 234 papers (data set 1), keyword analysis was conducted to examine the research trends. Then, we conducted a full-text review in the eligibility phase, and 187 papers were excluded based on the exclusion criteria. Finally, 47 papers (data set 2) were selected for the descriptive analysis.

XSL•FO
RenderX

Figure 1. Flowchart of the process of selecting studies. EF, executive function; VR, virtual reality.



Data Analysis

Data analysis was conducted by two sub-analysis methods.

First, a keyword analysis was conducted with data set 1. It was important to specify keywords for the analysis, so we conducted keyword extraction with Python NLTK. The Natural Language Toolkit (NLTK) is a Python package that has been developed for natural language processing and document analysis. The NLTK extracted top 200 keywords based on their frequency from each article's title, keywords, and abstract. Among the results of the extraction, we selected significant keywords in line with the study's purpose and categorized them according to measuring targets, names of mental disorders, and behavioral and physiological measures. Moreover, some keywords were added from the 47 full-text descriptive analysis.

Second, a descriptive analysis was conducted in detail with data set 2. The analysis was conducted with almost the same categories as the keyword analysis, but we also added some other elements: purpose of article, measuring targets, age of participants, names of mental disorders, real-time walking, and environment of VR.

All the analyses were conducted within a certain timescale. Considering the evolution of yearly publication frequencies, the time periods were divided into 4 stages of 5-year blocks: period 1, 2000-2004; period 2, 2005-2009; period 3, 2010-2014; and period 4, 2015-2019.

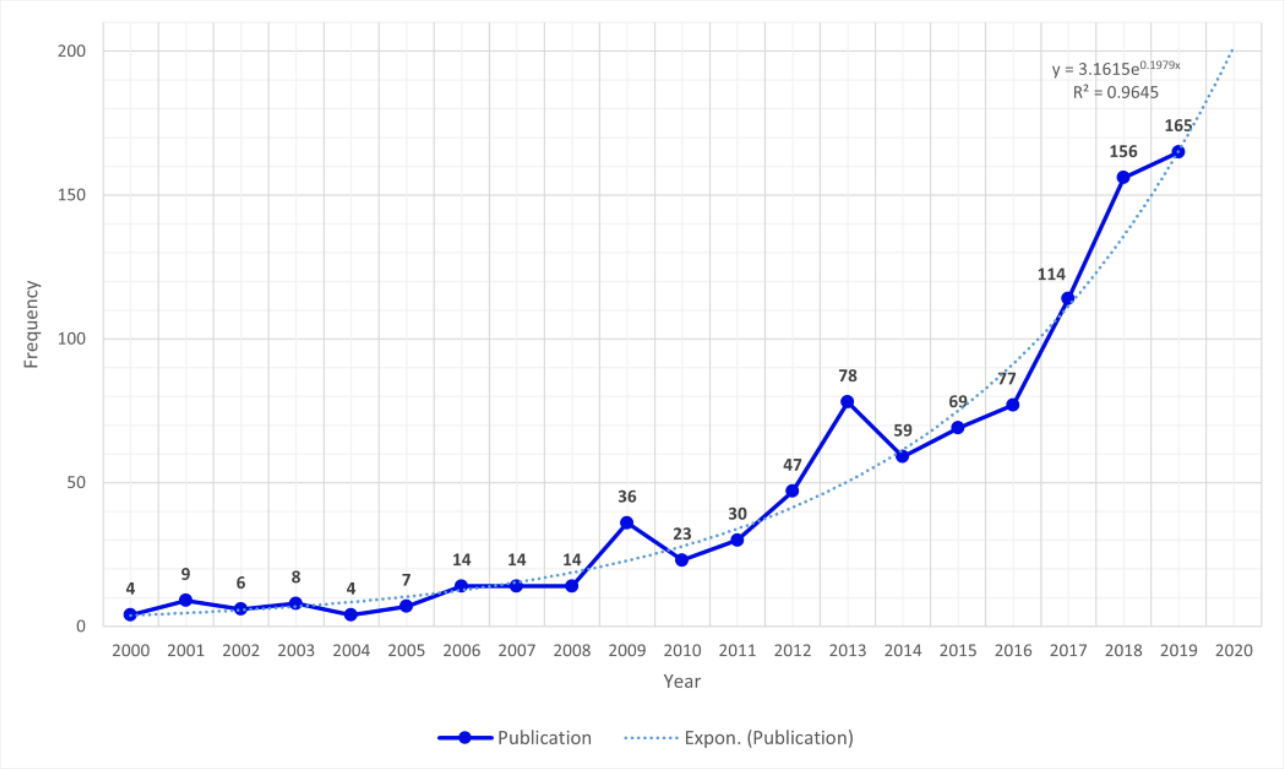
XSL•FO
RenderX

Results

Evolution of the Publication Frequency

In this section, we give an overview of the study of virtual neuropsychological tests in connection with EFs. There was a change of pace of the publication frequency from 2000 to 2019, with an exponential curve, as shown in Figure 2, with 936 articles after removing duplicates and outliers. Studies maintained a certain level all through the 2000s. The publication frequency began to increase conspicuously in 2009 and has rapidly increased in the past 10 years. It expresses the high interest in VR for dealing with mental disorders.

Figure 2. The evolution of publication frequency.



Keyword Analysis

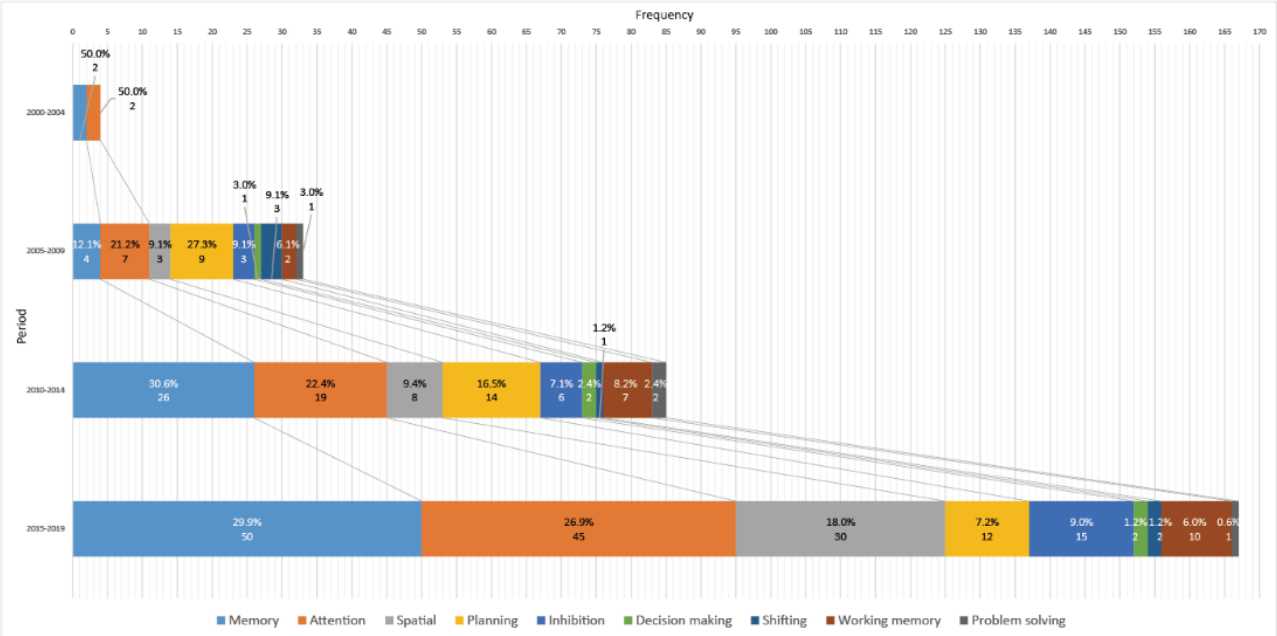
A keyword analysis was applied to data set 1 that mainly focused on EFs. For the keyword analysis, we segmented the total period (2000-2019) into four periods: 2000-2004, 2005-2009, 2010-2014, and 2015-2019. A significant increase in the publications in Figure 2 was the main rationale to classify the periods. In the keyword analysis, we examined the trends in VR applying to neuropsychology.

Target Cognitive Abilities

Cognitive ability includes various information-handling processes that occur in the human brain [39]. Target EFs brought from the theoretical background were complemented with "memory," "attention," and "spatial" since these frequently appeared with keyword extraction. Memory and attention are wide and basic concepts when assessing cognitive abilities [39,40]. They had a high proportion through all periods (Figure 3) because they occurred when measuring specific cognitive abilities together.

XSL•FO
RenderX

Figure 3. The evolution of cognitive abilities targeted in the 234 screened studies.



In the early 2000s, period 1, all studies tried measuring basic abilities, memory, and attention, but researchers began to try to measure other targets, including higher-order EFs. It seems that the diversification accelerated in line with the popularization of VR in the late 2000s.

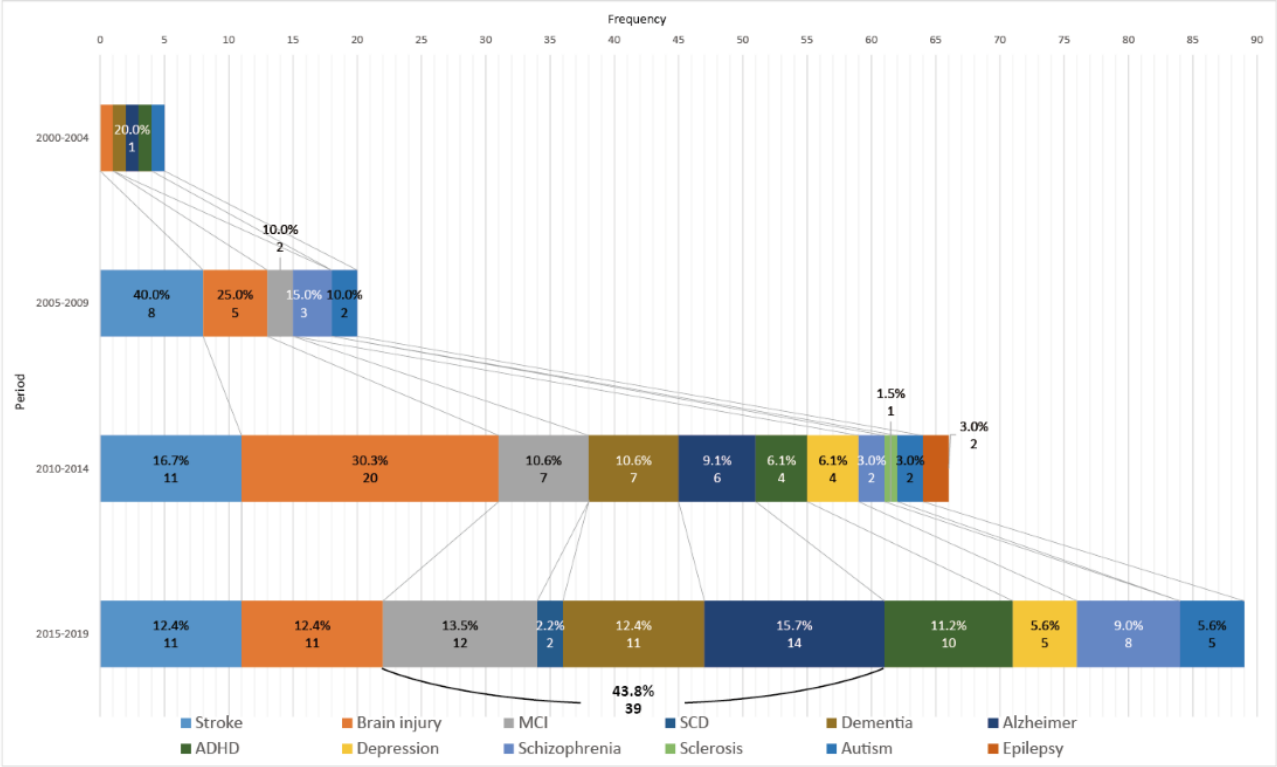
Target Symptoms

Our mental disorder list was formed by the result of keyword extraction, and we manually added several mental disorders from the 47 articles we reviewed for full-text descriptive analysis. There was a transition stage from the late 1990s to the early 2000s, when traditional pen-and-paper tests turned into computerized tests [41]. As shown in Figure 4, only a few studies conducted experimental attempts to apply VR to neuropsychology. Period 1 was the beginning stage of using VR technology for neuropsychological assessment, and this was ongoing through period 2. After that, from the early 2010s, studies expanded to treat various types of mental disorders. Until period 3, brain injury (stroke is included in brain injury in a broad view) was the most common mental disorder in this research area, maybe due to its high incidence and fatal long-term consequences [42]. However, in period 4, studies gradually spread to other brain disorders.

Through period 4, research on the topics mild cognitive impairment (MCI), subjective cognitive decline (SCD), dementia, Alzheimer's disease (AD), attention deficit hyperactivity disorder (ADHD), and schizophrenia increased. Among them, MCI, SCD, AD, and dementia are age-related symptoms that increased from period 3 (36.4%) to period 4 (43.8%). In this article, we view SCD, MCI, AD, and dementia as one continuous process and call it "age-related cognitive decline." In addition, studies on ADHD increased from period 3 (6.1%) to period 4 (11.2%), and studies on schizophrenia increased from 3.0% to 9.0%. Despite a high prevalence compared with schizophrenia, autism, ADHD, and depression remained less investigated.

XSL•FO
RenderX

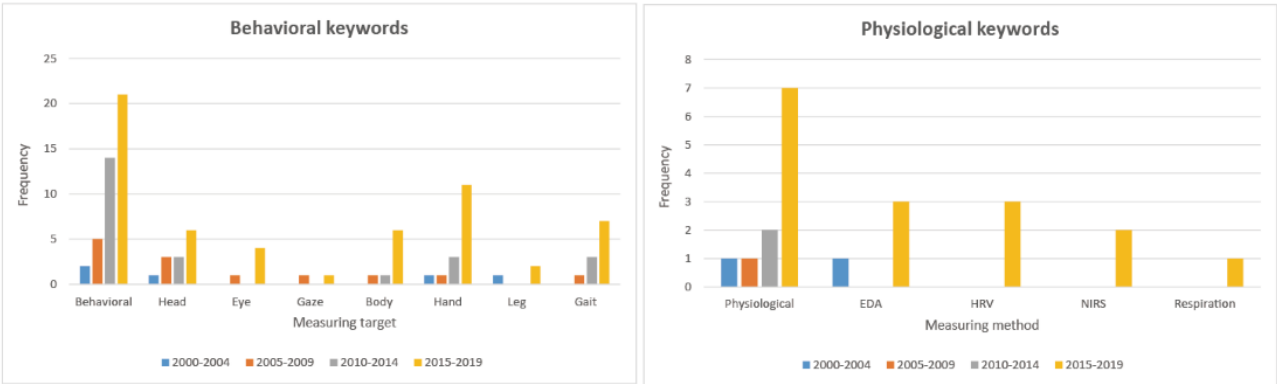
Figure 4. The evolution of symptoms targeted in the 234 screened studies. AD: Alzheimer's disease; ADHD, attention deficit hyperactivity disorder; MCI: mild cognitive impairment; SCD: subjective cognitive decline.



Use of Behavioral and Physiological Data

Attempts to obtain a subject's behavior data continuously increased from the early 2000s, while attempts to obtain physiological data rapidly increased in period 4 (Figure 5). In terms of content, there were attempts to use head tracking, gaze (eye) tracking, body tracking, hand tracking, and gait (leg) tracking, as shown in Figure 5. In the graph for physiological data, there were studies completed in EDA, HRV, NIRS, and respiration measurement, but EDA was the only attempt before period 4, and the others were conducted in period 4.

Figure 5. (Left) The use of behavioral data over each period. (Right) The use of physiological data over each period. EDA: electrodermal activity; HRV: heart rate variability; NIRS: near-infrared spectroscopy.



Descriptive Analysis

The final 47 articles included through PRISMA were experimental studies that conducted neuropsychological tests using fully immersive VR. The article set was divided into the same four periods as the keyword analysis. Studies on fully immersive VR were encouraged in the early 2010s, and most of the research was done within the past 5 years (period 4) with every year. Table 1 represents a summary of the 47 studies.

XSL•FO
RenderX

Table 1. Summary of the characteristics of the 47 studies.

Study	Journal/conference	Purpose of study	Target cognitive functions	Subject status	VE[a]	BE[b]
Areces et al (2018) [43]	*PLoS ONE*	As[c] Val.S[d]	Attention Inhibition	ADHD[e] Age 6-16 years	Classroom	—
Bailey et al (2019) [44]	*Journal of Applied Developmental Psychology*	Un[f]	Inhibition	Normal only Age 4-6 years	Other(s)	—
Blume et al (2017) [45]	*Trials*	Tr[g] Val.U[h]	Attention Inhibition Working memory	ADHD Age 6-10 years	Classroom	O[i]
Chicchi et al (2019) [46]	*Applied Neuropsychology: Adult*	Val.U	Attention Decision making Inhibition Planning Shifting	Normal only Mean age 30.14 years	Other(s)	—
Chicchi et al (2019) [47]	*Frontiers in Psychology*	Un	Attention Planning Shifting	Normal only Age 18-48 years	Kitchen	O
Climent et al (2019) [48]	*Applied Neuropsychology: Adult*	Un	Attention Working memory	Normative study Age 16-90 years	Aquarium	—
Coleman et al (2019) [49]	*Frontiers in Psychology*	Tr	Attention Working memory	ADHD Age 6-13 years	Classroom	—
Dahdah et al (2017) [50]	*Neurorehabilitation*	Tr Val.U	Inhibition	Stroke, TBI[j] Mean age 40.3 years	Apartment Classroom	—
Davison et al (2018) [51]	*Acta Neuropsychiatrica*	Val.S Val.U	Planning Working memory Motor	Normal only Age 19-24, 66-77 years	Parking lot Laboratory	—
De Lillo et al (2014) [52]	*American Journal of Primatology*	Un	Working memory	Normal only Age 18-36 years	Other(s)	—
De Luca et al (2020) [42]	*Applied Neuropsychology: Child*	Tr	Inhibition Planning Problem solving Updating Working memory Motor Spatial	TBI Age 15 years	A number of VEs	O
Díaz-Orueta et al (2014) [53]	*Child Neuropsychology*	As Val.U	Attention Inhibition	ADHD Age 6-16 years	Classroom	—
Eom et al (2019) [54]	*Cyberpsychology, Behavior and Social Networking*	As Un	Attention Inhibition	ADHD Age 6-17 years	Classroom	—
Grewe et al (2013) [55]	*Journal of NeuroEngineering and Rehabilitation*	Val.S	Working memory Spatial	Epilepsy Age 25-47 years	Market	—
Grewe et al (2014) [56]	*Epilepsy and Behavior*	As Val.S	Working memory Spatial	Epilepsy Age>18 years	Market	—
Henry et al (2012) [57]	*Journal of Neuroscience Methods*	Val.U	Attention Inhibition Motor	Normal only Age 19-58 years	Apartment	—

XSL•FO
RenderX

Study	Journal/conference	Purpose of study	Target cognitive functions	Subject status	VE[a]	BE[b]
Huang (2020) [58]	*Cyberpsychology Behavior and Social Networking*	Tr Un	Inhibition Shifting Working memory	Normal only Age>50 years	Other(s)	—
Ijaz et al (2019) [59]	*Journal of Medical Internet Research*	As Val.S	Working memory Spatial	Normal only Mean age 73.22 years	Google Street View	O
Iriarte et al (2012) [60]	*Journal of Attention Disorder*	Un	Attention Inhibition	Normative study Age 6-16 years	Classroom	—
Kang et al (2008) [61]	*Cyberpsychology and Behavior*	As	Attention Planning Working memory	Stroke Mean age 55.4 years	Market	—
Lalonde et al (2013) [29]	*Journal of Neuroscience Methods*	Val.U	Inhibition	Normal only Age 13-17 years	Classroom	—
Liao et al (2019) [62]	*Frontiers in Aging Neuroscience*	Tr Val.U	Attention Inhibition Planning Shifting Working memory Motor	MCI[k] Age>65 years	A number of VEs	O
Lo Priore et al (2003) [63]	*Cyberpsychology and Behavior*	Un	Attention Planning Shifting	Normal only Young age	Market	O
Mondellini et al (2018) [64]	*2018 IEEE 6th International Conference on Serious Games and Appli cations for Health, SeGAH 2018*	Tr	Working memory	Normal only Age<40 years	Market	—
Negut et al (2017) [65]	*Child Neuropsychology*	As Val.S Val.U	Attention Inhibition	ADHD Age 7-13 years	Classroom	—
Nolin et al (2009) [66]	*Annual Review of CyberTherapy and Telemedicine*	Val.U	Attention Inhibition	TBI Age 8-12 years	Classroom	—
Nolin et al (2016) [67]	*Computers in Human Behavior*	Val.S Un	Attention Inhibition	Normal only Age 7-16 years	Classroom	—
Ouellet et al (2018) [68]	*Journal of Neuroscience Methods*	Val.S Un	Working memory	SCD[l] Mean age 67.2 years	Market	—
Parsons et al (2013) [69]	*Journal of Clinical and Experimental Neuropsy chology*	Val.U	Attention Inhibition	Normal only Age 18-28 years	In a car	O
Parsons et al (2016) [70]	*Journal of Autism and Developmental Disorder*	As Val.U	Inhibition	Autism Age 18-34 years	Classroom	—
Parsons et al (2017) [71]	*Journal of Neuroscience Methods*	As Val.S	Working memory	Normal only Age>17 years	Market	—
Parsons et al (2018) [72]	*IEEE Transactions on Affective Computing*	As	Attention Inhibition	Normal only Age 18-28 years	In a car	O
Parsons et al (2018) [73]	*Journal of Neuroscience Methods*	Val.U	Inhibition	Normal only Age 18-30 years	Apartment	—
Parsons et al (2019) [74]	*Applied Neuropsychology: Adult*	As	Attention Inhibition	Normal only Older adults and undergraduate young adults	Apartment	—

XSL•FO
RenderX

Study	Journal/conference	Purpose of study	Target cognitive functions	Subject status	VE[a]	BE[b]
Plechatá et al (2019) [75]	*Frontiers in Psychology*	As Un	Decision making Shifting Working memory	Normal only Age 19-39, 60-91 years	Market	—
Plotnik et al (2017) [76]	*International Conference on Virtual Rehabilitation (ICVR)*	Val.S	Attention Planning	Normal only Mean age 37.1 years	Other(s)	O
Proulx et al (2018) [77]	*Annual Review of CyberTherapy and Teleme dicine*	Un	Attention Inhibition Working memory	Normal only Age 20-60 years	Laboratory Tent	—
Rizzo et al (2000) [78]	*Cyberpsychology and Behavior*	As	Attention Inhibition	ADHD Age 8-12 years	Classroom	—
Robitaille et al (2017) [79]	*Disability and Rehabili tation: Assistive Technology*	As	Attention Motor	TBI Mean age 30.3 years	Middle-east Village	O
Serino et al (2018) [80]	*Sensors (Switzerland)*	As	Working memory Spatial	AD[m] Age>65 years	Other(s)	—
Tarnanas et al (2013) [81]	*International Conference on Virtual Re habilitation, ICVR 2013*	As Un	Inhibition Planning Working memory Motor	MCI Age >60 years	Apartment	O
Tarnanas et al (2013) [82]	*Journal of Medical Internet Research Serious Game*	As Val.S	Inhibition Shifting Updating	MCI, AD Age>65 years	Apartment	O
Tarnanas et al (2014) [83]	*Alzheimer's and Dementia*	As	Inhibition Shifting Updating	MCI Mean age 71.6 years	Apartment	O
Voinescu et al (2019) [84]	*26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 – Proceeding*	Val.S Val.U	Attention Inhibition	Normal only Age 23-51 years	Aquarium	—
Yasuda et al (2017) [85]	*Topics in Stroke Rehabilitation*	Tr Val.S	Spatial	Stroke Age 45-85 years	Other(s)	—
Yeh et al (2012) [86]	*2012 IEEE-EMBS Conference on Biomedical Engineering and Scien ces, IECBES 2012*	As	Attention Decision making Shifting Working memory	ADHD Age 7-13 years	Classroom	O
Yeh et al (2012) [87]	*2012 IEEE-EMBS Conference on Biomedical Engineering and Scien ces, IECBES 2012*	As	Planning Working memory	Dementia Age 60-90 years	Market	O

[a]VE: virtual environment.
[b]BE: behavior or ecological measures.
[c]As: assessment.
[d]Val.S: structure validation.
[e]ADHD: attention deficit hyperactivity disorder.
[f]Un: uncategorized.
[g]Tr: treatment.
[h]Val.U: usability validation.
[i]O: includes BE measures.
[j]TBI: traumatic brain injury.
[k]MCI: mild cognitive impairment.
[l]SCD: subjective cognitive decline.
[m]AD: Alzheimer's disease.

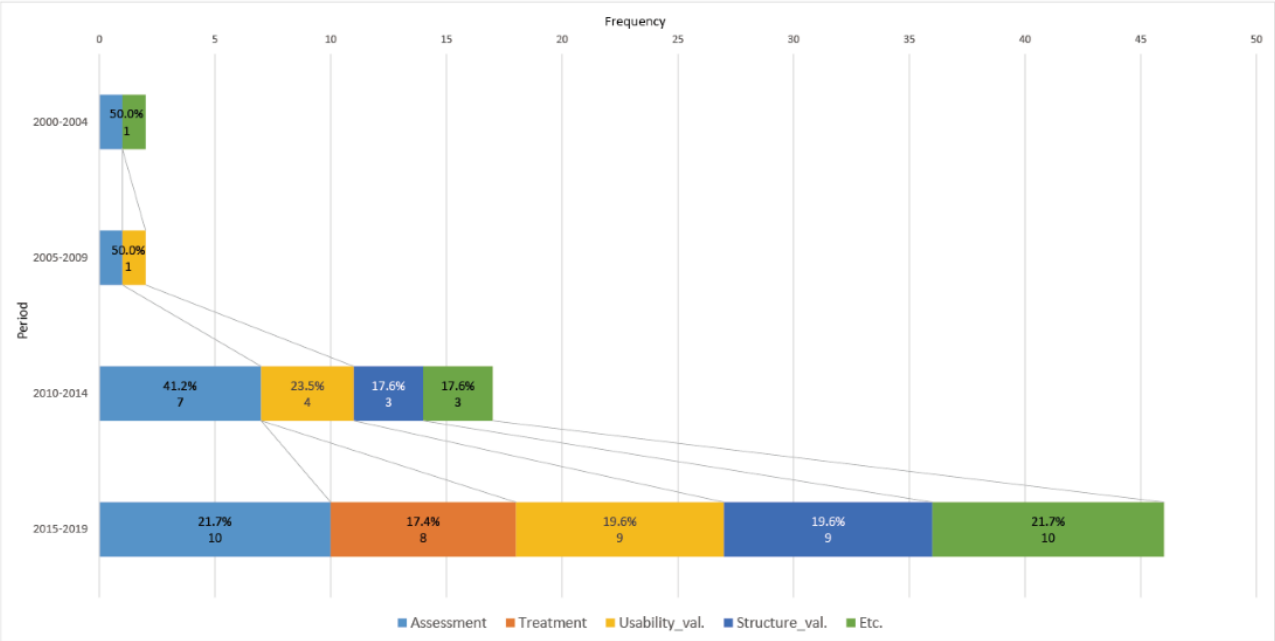
XSL•FO
RenderX

Analysis of Purpose of the Studies

This study identified that the purpose of the 47 articles is mainly clustered into four groups, and some additional purposes also existed. Thus, in this study, we divided the articles into five groups: assessment, treatment, usability validation, structure validation, and, for additional purposes, uncategorized: (1) assessment: to assess, diagnose, or evaluate a neuropsychological condition of the subject; (2) treatment: to treat or rehabilitate the patient; (3) usability validation: to validate the efficiency of a VR application by comparing it with an existing traditional neuropsychological test; (4) structure validation: to attempt to test the construct integrity and ecological and temporal stability of the VR itself; and (5) uncategorized: to evaluate presence, cybersickness, aging, and normative studies. Since there could be multiple purposes in one article, we checked all articles for all purposes.

As shown in Figure 6, the majority of the studies were conducted for assessment and validation. The number of validation studies was nearly steady through all periods. The number of studies dramatically increased as we moved to period 3. In period 3, 7 studies (41.2%) reported the results of assessing EF(s), 7 studies (41.2%) tried to validate their VR neuropsychological tests, but there were no studies for applying a VR test to treatment until period 4. Treatment studies first appeared in 2017, and the tendency began to expand. In period 4, 10 (21.7%), 8 (17.4%), and 18 (39.2%) studies were conducted on assessment, treatment, and validation, respectively.

Figure 6. The evolution of the objective of the final 47 studies.



Meanwhile, the proportion of uncategorized studies increased rapidly in period 4. Before period 4, there were just 4 studies, but 10 studies corresponded to the uncategorized type. Of the four studies, one was a study for the sense of presence that was conducted early in VR history, and the others were conducted in period 3 to obtain information about certain cognitive functions associated with functional impairment in ADL. However, the uncategorized studies in period 4 were slightly different: 7 (70%) of the 10 studies were conducted to evaluate the sense of presence and cybersickness of users in the VE, and the remainder were studies on age-specific differences.

Cognitive Abilities and VEs

The VEs used to target EF(s) are shown in Table 2. In addition to the cognitive abilities selected in the Target Cognitive Abilities section, we added spatial (navigation, spatial perception, etc) and motor (psychomotor, motion, movement, etc) abilities, because these are significant features within fully immersive VR tests. In general, most articles not only target a certain cognitive ability but also treat several abilities, so we allowed for multiple checking.

XSL•FO
RenderX

Table 2. Frequency of VEs[a] by EFs[b].

Cognitive ability	VEs	Reference	Number of papers that studied the specific cognitive ability (n)
Inhibition (n=27)	Apartment	[50,57,73,74,81-83]	7
	Classroom	[29,43,45,50,53,54,60,65-67,70,78]	12
	Aquarium	[84]	1
	Laboratory	[77]	1
	Tent outdoors	[77]	1
	In a car	[69,72]	2
	Others[c]	[42,44,46,58,62]	5
Working memory (n=21)	Laboratory	[51,77]	2
	Parking lot	[51]	1
	Market	[55,56,61,64,68,71,75,87]	8
	Classroom	[45,49,86]	3
	Tent outdoors	[77]	1
	Google Street View	[59]	1
	Aquarium	[48]	1
	Others	[42,52,58,62,80]	5
Shifting (n=9)	Apartment	[82,83]	2
	Market	[63,75]	2
	Classroom	[86]	1
	Kitchen	[47]	1
	Others	[46,58,62]	3
Decision making (n=3)	Classroom	[86]	1
	Market	[75]	1
	Others	[46]	1
Problem solving (n=1)	Others	[42]	1
Planning (n=10)	Parking lot	[51]	1
	Laboratory	[51]	1
	Market	[61,63,87]	3
	Apartment	[81]	1
	Kitchen	[47]	1
	Others	[42,46,62,76]	4
Updating (n=3)	Apartment	[82,83]	2
	Others	[42]	1

XSL•FO
RenderX

Cognitive ability	VEs	Reference	Number of papers that studied the specific cognitive ability (n)
Attention (n=24)	Classroom	[43,45,49,53,54,60,65-67,78,86]	11
	Market	[61,63]	2
	Aquarium	[48,84]	2
	In a car	[69,72]	2
	Kitchen	[47]	1
	Laboratory	[77]	1
	Tent outdoors	[77]	1
	Apartment	[57]	1
	Middle-east village	[79]	1
	Others	[46,62,76]	3
Spatial (n=6)	Market	[55,56]	2
	Google Street View	[59]	1
	Others	[42,80,85]	3
Motor (n=6)	Parking lot	[51]	1
	Laboratory	[51]	1
	Apartment	[57,81]	2
	Middle-east village	[79]	1
	Others	[42,62]	2

[a]VE: virtual environment.

[b]EF: executive function.

[c]Others: does not use specific environments, uses multiple environments as a minigame with multiple scenarios, or uses a certain unrealistic environment that could not be experienced in real life.

Since the study of fully immersive VR began, numerous VEs have been designed. In the initial stage, until 2012, only classrooms and supermarkets were studied, classrooms four times [60,66,78,86] and supermarkets three times [61,63,87], but VEs began to diversify as research progressed.

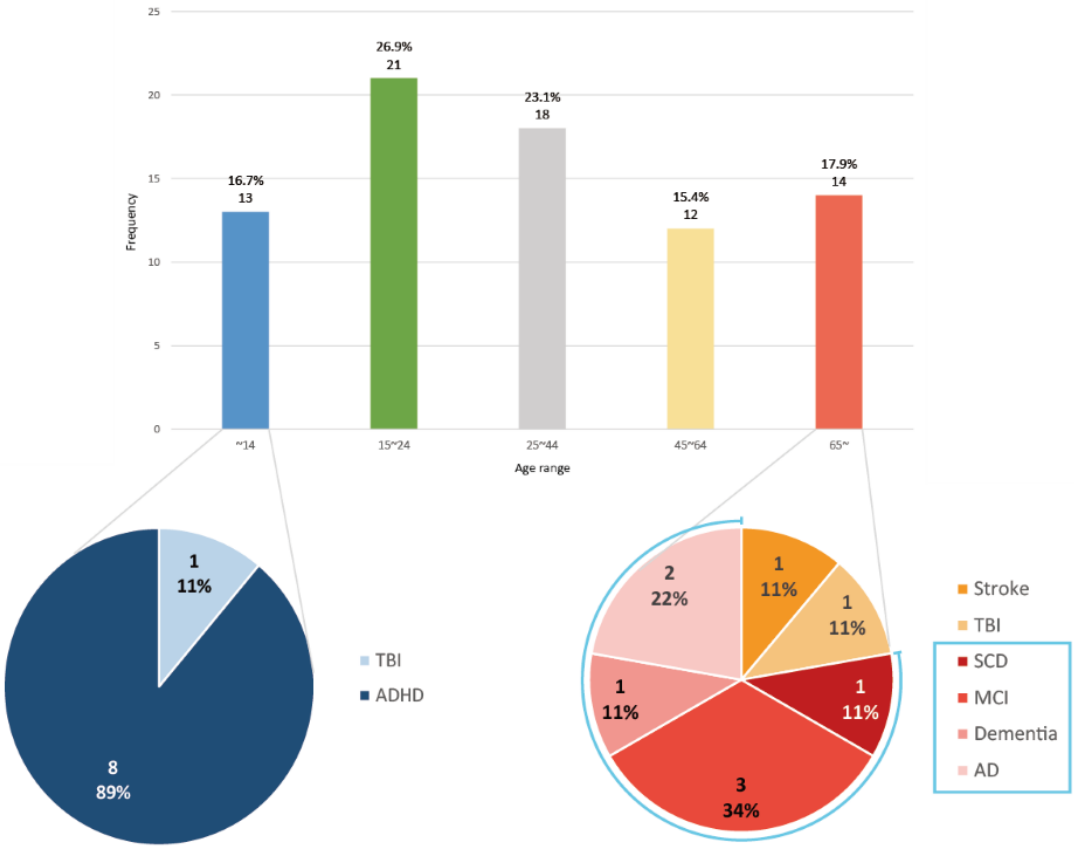
Inhibition and attention were targeted the most, as they are deeply related functionally and conceptually. They were studied 27 and 24 times, respectively. For inhibition and attention, virtual classroom environments were mostly used, and virtual apartments were the next most common type of environment. The next most targeted EF was working memory, which was studied 21 times. For working memory, various VEs were used, but markets were the most common, with eight studies. Next, planning was studied 10 times, shifting was studied 9 times, spatial and motor were examined 6 times each, and decision making, updating, and problem solving were studied 3, 2, and 1 times, respectively.

Description of Participants

An overall distribution of the participants' age is presented in Figure 7. For age sorting, the United Nations' age classification guideline for health, health services, and nutrition was applied [88]. Some articles that did not perfectly fit the guideline were checked multiple times according to their range of participants.

XSL•FO
RenderX

Figure 7. Research based on the age of the participant and symptoms of two age groups, under 14 and over 65 years. AD: Alzheimer's disease; ADHD: attention deficit hyperactivity disorder; MCI: mild cognitive impairment; SCD: subjective cognitive decline; TBI: traumatic brain injury.



The age distribution of participants in all the articles was similar in all periods. The proportion of 15-24-year-olds was the highest, and the proportion of those 65 years and was the third. In the under-14 age group, there is just one article that studied traumatic brain injury (TBI), and the others examined ADHD. Among those above 65 years, the age-related cognitive decline corresponded to nearly 80% of all articles. In addition, there were many studies with only a healthy group.

The Emphasis on Real-Time Walking

The use of a multidirectional treadmill to allow a subject the possibility of walking by themselves first appeared in period 3. Research using treadmills usually targeted patients with TBI and MCI to rehabilitate their gate and motor ability. Entering period 4, some VR applications began to provide real-time walking without any equipment. In fully immersive VR, supermarkets, kitchens, laboratories, and small squares were used as VEs, but supermarkets corresponded to half of all environments.

Fully immersive VR provides a sense of reality and of presence in virtual spaces with a special apparatus, such as an HMD [21]. Furthermore, another merit of fully immersive VR is that it has a high level of ecological validity by allowing free movement in a certain space. The real-time walking studies can be separated into two categories, real-time walking and walking on a treadmill. In the case of a treadmill, fully immersive environments are built with a 180°/360° or a dome-shaped screen [42,76,83]. For real-time walking, VR can function as a tool allowing ADL.

Discussion

Expansion of Research: From Assessment to Treatment and From Simple Tasks to ADL

The most noticeable phenomenon is the appearance of treatment studies in period 4. In the early periods, when VR was first used in neuropsychological tests, it was only used for assessing cognitive states. However, high ecological validity, reproducibility of daily life, and the possibility of measuring behavior responses that were not available with traditional neuropsychological tests opened a new venue for cognitive rehabilitation [42,64,89,90]. Therefore, the research trend changed from being assessment oriented toward an additional interest for treatments. The possibility of treatment of cognitive problems using VR seems to expand continuously. In these studies, researchers have tried to train or rehabilitate the participants' cognitive abilities (mainly inhibition and working memory) through regular and repetitive VR tasks that consisted of daily-living scenarios or game-like applications [42,58,62,64]. Some research argues that the biomarker and the degree of brain activation obtained as kinematic outcomes are indicators of treatment response [91,92].

From the perspective of changes in validation, it seems that there are two main streams. The first perspective is about

XSL•FO
RenderX

transplanting existing traditional tests into VR, such as the Stroop test [29,50,57,67,69,70,73] and the trail-making task [76]. This includes most of the usability validation studies in data set 2. Because the Stroop test is a simple and ocular task, it has been considered suitable to convert into a VE. In addition, there is one study that brings the trail-making test into the VE. Second, an interesting thing is that most of the structure validation studies were conducted in supermarket environments based on the multiple-errand test [55,56,68,71]. Supermarket environments are popular due to their deep relation with daily living. Researchers have also tried to design various scenarios that are related to daily living, such as parking simulation [51] and pathfinding with Google Street View [59]. However, continuing to adapt parameters or structural properties to suit VR neuropsychological tests remains a challenge.

The increase in uncategorized studies in period 4 could also be seen in the same context. Most of the uncategorized studies examined the sense of presence and cybersickness [44,47,54,58,63,67,75,77]. These are indeed the two major problems for wider acceptance of VR for more age groups and for obtaining more ecological validity. The increase in studies to solve these problems aims for participants to feel more comfortable, which could further lead to the development of daily-living contents and expand the area of treatment. Therefore, solving these problems seems a key point to improve fully immersive neuropsychological testing in VR, but there is still a long way to go.

Diversification of Targeted EFs and VEs

As shown in Figure 3, diversification started in line with the popularization of VR in the late 2000s, and the research has grown quantitatively and qualitatively ever since. Initially, studies mostly targeted low-order EFs, such as inhibition, attention, and working memory [61,66,78]. However, in recent years, VR technology has evolved to overcome the problem of presence and cybersickness. This enabled researchers to implement higher ecological ADL with various practical scenarios and VEs. Consequently, it has led to studies that deal with higher-order EFs, such as decision making, planning, and problem solving.

Inhibition and attention are functions deeply related to ADHD. Because most subjects with ADHD are children and youth, it seems that researchers have actively used classroom and apartment (more accurately, the living room) environments that those ages are familiar with [43,45,53,54,65,78]. As the Stroop test is famous for assessing one's inhibition, many studies on inhibition and attention have translated it into VR. Especially, most studies conducted under virtual apartment environments have developed their application based on the Stroop test [50,57,73,74]. Since familiarity with the participants' daily-living situation induces psychological comfortability, application can sophisticatedly measure the natural state of cognitive abilities.

The supermarket environment is most frequently used for measuring working memory. Planning [61,63,87] and decision making [75] are also associated with the supermarket environment. For higher-order EFs, various other environments and daily-living scenarios have been tried, such as laboratories [51,77], kitchens [47], and parking simulation [51]. It is unrealistic to precisely measure one isolated EF due to its interdependent property. Higher-order EFs should be comprehensively measured. Only one case measured planning alone, a study that was the first attempt to apply the trail-making test to immersive VR [76]. Lastly, studies on spatial and motor abilities were tried after 2012 using a treadmill or motion-tracking systems.

Since one of the EFs closely connected to other EFs we defined, EFs in ADL tend to be treated comprehensively rather than alone. These deal with specific situations, the emergence of distractors, and well-organized scenarios that assist the original task.

Diversification of Target Symptoms and Expansion of the Age Group

As shown in Figure 4, target symptoms began to diversify after period 2. Stroke and brain injury accounted for the largest percentage until period 3, but in period 4, this changed. This indicates that VR studies on neuropsychology are being conducted in a wide variety of conditions. The thing to keep an eye on is the conspicuous increase of age-related cognitive decline and a relatively high proportion of ADHD.

The many VR studies have focused on ADHD, which is a popular symptom of young age in the early stage of VR research. In Figure 7, the under-14 age group is composed of TBI and ADHD, and ADHD represents the majority of these studies. It is not surprising as almost 5% of children are diagnosed with ADHD [86]. Many previous studies have reported the positive effect of VR on children who have a mental disorder [93-95] and surely on ADHD [43,78,86]. In addition, the ease of converting the Stroop test into immersive VR would have played a role, since the Stroop test is a neuropsychological assessment to measure attention and inhibition, which are major deficit features of ADHD [43]. If we regard under 25 years as one unique group, the activeness of related research can be seen as familiarity and accessibility to VR and resistance to cybersickness of adolescents. Young people are more adaptable to VR, even immersive VR [96].

As VR technology advances and related research accumulates, obstacles that blocked the application of VR neuropsychological tests to the elderly are being resolved [97]. Figure 7 also shows the expansion of the research to the elderly. In the 65+ age group, about 80% of papers are related to age-related cognitive decline. As mentioned before, we consider SCD, MCI, AD, and dementia together as age-related cognitive decline. MCI is a symptom that decreases one's memory, attention, and cognitive function, and it could lead to dementia or AD [98]. SCD is an early state when someone's cognitive ability is beginning to decline [99]. In other words, SCD may indicate a pre-MCI condition, even an early marker for the symptomatic manifestation of dementia. Age-related cognitive decline symptoms are common in the elderly. In addition, depression is a quite under-investigated area despite a huge impact on daily-living executive dysfunction in 60% of depressive states in the elderly, with an exponential growth of prevalence with age, a poorer response to antidepressants, and more evolution

XSL•FO
RenderX

toward dementia compared to depression without dysexecutive syndrome [100-102].

Investigating VR as a neuropsychological tool for the elderly seems increasingly realistic since VR has become more familiar even for seniors. As we have seen before, there has been a lot of research on and technological advancement in presence and cybersickness in VR, and it is possible that the public and even the elderly could comfortably access it [103-105], which enables expansion to the elderly and causes rapid growth of the research that targets age-related cognitive decline symptoms after period 4.

Overall, immersive VR offers a high number of possibilities to deal with neuropsychological symptoms in various age groups. It shows the versatility of VR, which is applicable from youth with ADHD to the elderly with age-related cognitive decline.

Combining Behavioral and Physiological Data Measurement

It seems that measuring behavioral data and physiological data is becoming increasingly important. These kinds of data are used to extend the concept of embodied cognition to measure the neuropsychological status and to overcome the weaknesses of VR.

Embodied cognition is a broad notion from embodiment thesis in philosophy, and it states that an agent's cognition is powerfully affected by aspects of their body beyond the brain itself. This means that sensorimotor experiences and actions are crucial to cognitive processing [106]. These aspects are thought to be crucial for cognitive performances in the elderly [107]. In this context, measuring behavioral and physiological variables in immersive VR would be a key method of assessing the neuropsychological status. Thus, the growth of associated research seems to be increasing because of the ecological characteristics of VR environments. In addition, when we looked deeper into gait or walking, there was a large difference in the appearance of walking on a treadmill and real-time walking in periods 3 and 4. With real walking in a VR neuropsychological test, researchers were able to apply the concept of embodied cognition and measure it in ADL.

Furthermore, an increase in physiological data measurements should be considered as these can provide auxiliary ways to assess the neuropsychological status of a patient. In addition, it is also considered to be a helpful way to overcome cybersickness [108-110]. Cybersickness is a major obstacle in applying VR to older people and people who have difficulties adapting to VR because they feel dizzy. In addition, research on emotion recognition and presence in VR uses physiological data [111-114]. With these efforts added, VR neuropsychological tests' versatility for all ages continues to improve.

Overall, by using behavioral and physiological data measurement, VR neuropsychological tests are not only getting over existing drawbacks but also expanding to the concept of embodied cognition to improve its measurement capability.

Limitations

This review had a few limitations. First, keyword analysis was conducted by only using titles and abstracts of the articles. This means that there is a possibility of a difference with the articles' main texts. Second, we only included articles that described research conducted with real participants, so conceptual and theoretical trends in the area were not covered. Finally, not every article markedly described the target EFs. Likewise, there were some ambiguities in clearly assigning the purpose of each article to assessment, treatment, structure validation, or usability validation. Even though we attempted to judge based on the description of what EFs the articles targeted, and what their intended purpose was, there was still room for error. Future research could include broader studies in the area and use more rigorous methods to analyze the trends.

Conclusion

This review will assist researchers in understanding the trends in VR neuropsychological tests over the past 20 years. Associated research has been on the rise and has sharply increased in recent years. In this process, the advancements in technology and various approaches have led to diversified target cognitive abilities, including EFs, as well as target symptoms. Moreover, collecting behavioral and physiological data enables a wide understanding about treating EFs by using the concept of embodied cognition. As a result, VR neuropsychological tests now cover a wide range of age groups and extend beyond assessment tools to treatment tools. This review shows that there is a continuously increased interest in dealing with neuropsychology by using fully immersive VR, and it is expected that this will help advance research in this area.

Acknowledgments

This work was supported by the Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korean Government (MSIP; no. 2019-0-01569, Development of VR and IoT-based Systems for Evaluation of Embodied Cognition). In addition, this work has been partly funded by the Pays de la Loire French region through the West Creative Industries program and by the European Union through the Hubert Curien French-Korean program STAR.

Authors' Contributions

GHK and JH conceptualized the study. EK collected data. EK, GHK, and JH conducted screening, selection, and quality assessment of the studies. EK led the analysis and interpretation of data and results. EK, GHK, JH, YP, and TV developed the interpretation. HC and SB advised the medical viewpoint. EK conducted draft writing. All authors have read, revised, and approved the final manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Botella C, Fernández-Álvarez J, Guillén V, García-Palacios A, Baños R. Recent progress in virtual reality exposure therapy for phobias: a systematic review. Curr Psychiatry Rep 2017 Jul 24;19(7):42. [doi: 10.1007/s11920-017-0788-4] [Medline: 28540594]
2. Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Trans Inf Syst 1994;E77-D(12):1321-1329.
3. Ott M, Freina L. A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives. 2015 Presented at: eLearning and Software for Education (eLSE); April 2015; Bucharest (Romania).
4. Mujber T, Szecsi T, Hashmi M. Virtual reality applications in manufacturing process simulation. J Mater Process Technol 2004 Nov;155-156:1834-1838 [FREE Full text] [doi: 10.1016/j.jmatprotec.2004.04.401]
5. Waltemate T, Gall D, Roth D, Botsch M, Latoschik ME. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. IEEE Trans Visual Comput Graphics 2018 Apr;24(4):1643-1652. [doi: 10.1109/tvcg.2018.2794629]
6. Gutiérrez F, Pierce J, Vergara V, Coulter R, Saland L, Caudell T, et al. The effect of degree of immersion upon learning performance in virtual reality simulations for medical education. Stud Health Technol Inform 2007;125:155-160. [Medline: 17377256]
7. Kim O, Pang Y, Kim J. The effectiveness of virtual reality for people with mild cognitive impairment or dementia: a meta-analysis. BMC Psychiatry 2019 Jul 12;19(1):219 [FREE Full text] [doi: 10.1186/s12888-019-2180-x] [Medline: 31299921]
8. Lezak M. Neuropsychological Assessment. 5th ed. Oxford: Oxford University Press; 2012.
9. 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]
10. Partington J, Leiter R. Partington's pathways test. Psychol Serv Center J 1949;1:11-20. [doi: 10.1037/t66320-000]
11. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935;18(6):643-662. [doi: 10.1037/h0054651]
12. Howieson D. Current limitations of neuropsychological tests and assessment procedures. Clin Neuropsychol 2019 Feb 04;33(2):200-208. [doi: 10.1080/13854046.2018.1552762] [Medline: 30608020]
13. Nef T, Chesham A, Schütz N, Botros AA, Vanbellingen T, Burgunder J, et al. Development and evaluation of maze-like puzzle games to assess cognitive and motor function in aging and neurodegenerative diseases. Front Aging Neurosci 2020;12:87 [FREE Full text] [doi: 10.3389/fnagi.2020.00087] [Medline: 32372942]
14. Negu A, Matu S, Sava FA, David D. Virtual reality measures in neuropsychological assessment: a meta-analytic review. Clin Neuropsychol 2016 Feb;30(2):165-184. [doi: 10.1080/13854046.2016.1144793] [Medline: 26923937]
15. Alvarez JA, Emory E. Executive function and the frontal lobes: a meta-analytic review. Neuropsychol Rev 2006 Mar;16(1):17-42. [doi: 10.1007/s11065-006-9002-x] [Medline: 16794878]
16. Chaytor N, Schmitter-Edgecombe M. The ecological validity of neuropsychological tests: a review of the literature on everyday cognitive skills. Neuropsychol Rev 2003 Dec;13(4):181-197. [doi: 10.1023/b:nerv.0000009483.91468.fb]
17. Elkind JS, Rubin E, Rosenthal S, Skoff B, Prather P. A simulated reality scenario compared with the computerized Wisconsin card sorting test: an analysis of preliminary results. Cyberpsychol Behav 2001 Aug;4(4):489-496. [doi: 10.1089/109493101750527042] [Medline: 11708728]
18. Schultheis M, Himelstein J, Rizzo A. Virtual reality and neuropsychology: upgrading the current tools. J Head Trauma Rehabil 2002 Oct;17(5):378-394. [doi: 10.1097/00001199-200210000-00002] [Medline: 12802250]
19. Maggio MG, De Cola MC, Latella D, Maresca G, Finocchiaro C, La Rosa G, et al. What about the role of virtual reality in Parkinson disease's cognitive rehabilitation? Preliminary findings from a randomized clinical trial. J Geriatr Psychiatry Neurol 2018 Nov 25;31(6):312-318. [doi: 10.1177/0891988718807973] [Medline: 30360679]
20. Maggio M, De Luca R, Molonia F, Porcari B, Destro M, Casella C, et al. Cognitive rehabilitation in patients with traumatic brain injury: a narrative review on the emerging use of virtual reality. J Clin Neurosci 2019 Mar;61:1-4. [doi: 10.1016/j.jocn.2018.12.020] [Medline: 30616874]
21. Valmaggia LR, Latif L, Kempton MJ, Rus-Calafell M. Virtual reality in the psychological treatment for mental health problems: a systematic review of recent evidence. Psychiatry Res 2016 Feb 28;236:189-195. [doi: 10.1016/j.psychres.2016.01.015] [Medline: 26795129]
22. Knight RG, Titov N. Use of virtual reality tasks to assess prospective memory: applicability and evidence. Brain Impair 2012 Feb 21;10(1):3-13. [doi: 10.1375/brim.10.1.3]
23. Pollak Y, Weiss P, Rizzo A, Weizer M, Shriki L, Shalev R. The utility of a continuous performance test embedded in virtual reality in measuring adhd-related deficits. J Dev Behav Pediatr 2009;30(1):2-6. [doi: 10.1097/dbp.0b013e3181969b22]

XSL•FO
RenderX

24. Overdorp E, Kessels R, Claassen J, Oosterman J. The combined effect of neuropsychological and neuropathological deficits on instrumental activities of daily living in older adults: a systematic review. Neuropsychol Rev 2016 Mar;26(1):92-106 [FREE Full text] [doi: 10.1007/s11065-015-9312-y] [Medline: 26732392]
25. Tomaszewski Farias S, Cahn-Weiner DA, Harvey DJ, Reed BR, Mungas D, Kramer JH, et al. Longitudinal changes in memory and executive functioning are associated with longitudinal change in instrumental activities of daily living in older adults. Clin Neuropsychol 2009 Apr;23(3):446-461 [FREE Full text] [doi: 10.1080/13854040802360558] [Medline: 18821181]
26. Pérès K, Helmer C, Amieva H, Orgogozo J, Rouch I, Dartigues J, et al. Natural history of decline in instrumental activities of daily living performance over the 10 years preceding the clinical diagnosis of dementia: a prospective population-based study. J Am Geriatr Soc 2008 Jan;56(1):37-44. [doi: 10.1111/j.1532-5415.2007.01499.x] [Medline: 18028344]
27. Wicklund A, Johnson N, Rademaker A, Weitner B, Weintraub S. Profiles of decline in activities of daily living in non-Alzheimer dementia. Alzheimer Dis Assoc Disord 2007;21(1):8-13. [doi: 10.1097/wad.0b013e3180324549]
28. Schmeidler J, Mohs RC, Aryan M. Relationship of disease severity to decline on specific cognitive and functional measures in Alzheimer disease. Alzheimer Dis Assoc Disord 1998 Sep;12(3):146-151. [doi: 10.1097/00002093-199809000-00005] [Medline: 9772016]
29. Lalonde G, Henry M, Drouin-Germain A, Nolin PA, Beauchamp MH. Assessment of executive function in adolescence: a comparison of traditional and virtual reality tools. J Neurosci Methods 2013 Sep 30;219(1):76-82. [doi: 10.1016/j.jneumeth.2013.07.005] [Medline: 23867080]
30. Lezak MD. The problem of assessing executive functions. Int J Psychol 1982 Jan 21;17(1-4):281-297. [doi: 10.1080/00207598208247445] [Medline: 5637250]
31. Anderson P. Assessment and development of executive function (EF) during childhood. Child Neuropsychol 2002 Jun 09;8(2):71-82. [doi: 10.1076/chin.8.2.71.8724] [Medline: 12638061]
32. Hughes C. Executive functions and development: why the interest? Inf Child Develop 2002 Jun;11(2):69-71. [doi: 10.1002/icd.296]
33. Lehto J, Juujärvi P, Kooistra L, Pulkkinen L. Dimensions of executive functioningvidence from children. Br J Dev Psychol 2003:59-80 [FREE Full text] [doi: 10.1348/026151003321164627]
34. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. Cogn Psychol 2000 Aug;41(1):49-100. [doi: 10.1006/cogp.1999.0734] [Medline: 10945922]
35. Lunt L, Bramham J, Morris R, Bullock P, Selway R, Xenitidis K, et al. Prefrontal cortex dysfunction and 'jumping to conclusions': bias or deficit? J Neuropsychol 2012 Mar;6(1):65-78. [doi: 10.1111/j.1748-6653.2011.02005.x] [Medline: 22257612]
36. Collins A, Koechlin E. Reasoning, learning, and creativity: frontal lobe function and human decision-making. PLoS Biol 2012;10(3):e1001293 [FREE Full text] [doi: 10.1371/journal.pbio.1001293] [Medline: 22479152]
37. Diamond A. Executive functions. Annu Rev Psychol 2013 Jan 03;64(1):135-168 [FREE Full text] [doi: 10.1146/annurev-psych-113011-143750] [Medline: 23020641]
38. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009 Jul 21;339(jul21 1):b2700-b2700 [FREE Full text] [doi: 10.1136/bmj.b2700] [Medline: 19622552]
39. Gupta M, Musilek P. Fuzzy neural networks and cognitive modeling. Int J Gen Syst 2000 Jan;29(1):7-28. [doi: 10.1080/03081070008960922]
40. Cartwright K. Insights from cognitive neuroscience: the importance of executive function for early reading development and education. Early Educ Dev 2012 Jan;23(1):24-36. [doi: 10.1080/10409289.2011.615025]
41. Zygouris S, Tsolaki M. Handbook of Research on Innovations in the Diagnosis and Treatment of Dementia. PA: IGI Global; 2018:1762-1779.
42. De Luca R, Portaro S, Le Cause M, De Domenico C, Maggio M, Cristina Ferrera M, et al. Cognitive rehabilitation using immersive virtual reality at young age: a case report on traumatic brain injury. Appl Neuropsychol Child 2020;9(3):282-287. [doi: 10.1080/21622965.2019.1576525] [Medline: 30838889]
43. Areces D, Dockrell J, García T, González-Castro P, Rodríguez C. Analysis of cognitive and attentional profiles in children with and without ADHD using an innovative virtual reality tool. PLoS One 2018 Aug 15;13(8):e0201039 [FREE Full text] [doi: 10.1371/journal.pone.0201039] [Medline: 30110334]
44. Bailey JO, Bailenson JN, Obradović J, Aguiar NR. Virtual reality's effect on children's inhibitory control, social compliance, and sharing. J Appl Dev Psychol 2019 Jul;64:101052. [doi: 10.1016/j.appdev.2019.101052]
45. Blume F, Hudak J, Dresler T, Ehlis A, Kühnhausen J, Renner TJ, et al. NIRS-based neurofeedback training in a virtual reality classroom for children with attention-deficit/hyperactivity disorder: study protocol for a randomized controlled trial. Trials 2017 Jan 24;18(1):41 [FREE Full text] [doi: 10.1186/s13063-016-1769-3] [Medline: 28118856]
46. Chicchi Giglioli IA, de Juan Ripoll C, Parra E, Alcañiz Raya M. Are 3D virtual environments better than 2D interfaces in serious games performance? An explorative study for the assessment of executive functions. Appl Neuropsychol Adult 2021;28(2):148-157. [doi: 10.1080/23279095.2019.1607735] [Medline: 31070055]

XSL•FO
RenderX

47. Chicchi Giglioli IA, Bermejo Vidal C, Alcañiz Raya M. A virtual versus an augmented reality cooking task based-tools: a behavioral and physiological study on the assessment of executive functions. Front Psychol 2019;10:2529 [FREE Full text] [doi: 10.3389/fpsyg.2019.02529] [Medline: 31798497]
48. Climent G, Rodríguez C, García T, Areces D, Mejías M, Aierbe A, et al. New virtual reality tool (Nesplora Aquarium) for assessing attention and working memory in adults: a normative study. Appl Neuropsychol Adult 2021;28(4):403-415. [doi: 10.1080/23279095.2019.1646745] [Medline: 31382773]
49. Coleman B, Marion S, Rizzo A, Turnbull J, Nolty A. Virtual reality assessment of classroom - related attention: an ecologically relevant approach to evaluating the effectiveness of working memory training. Front Psychol 2019;10:1851 [FREE Full text] [doi: 10.3389/fpsyg.2019.01851] [Medline: 31481911]
50. Dahdah MN, Bennett M, Prajapati P, Parsons TD, Sullivan E, Driver S. Application of virtual environments in a multi-disciplinary day neurorehabilitation program to improve executive functioning using the Stroop task. NRE 2017 Dec 15;41(4):721-734. [doi: 10.3233/nre-172183]
51. Davison SMC, Deeprose C, Terbeck S. A comparison of immersive virtual reality with traditional neuropsychological measures in the assessment of executive functions. Acta Neuropsychiatr 2018 Apr 09;30(2):79-89. [doi: 10.1017/neu.2017.14] [Medline: 28482936]
52. De Lillo C, Kirby M, James F. Spatial working memory in immersive virtual reality foraging: path organization, traveling distance and search efficiency in humans (Homo sapiens). Am J Primatol 2014 May;76(5):436-446. [doi: 10.1002/ajp.22195] [Medline: 24038208]
53. Díaz-Orueta U, Garcia-López C, Crespo-Eguílaz N, Sánchez-Carpintero R, Climent G, Narbona J. AULA virtual reality test as an attention measure: convergent validity with Conners' Continuous Performance Test. Child Neuropsychol 2014;20(3):328-342. [doi: 10.1080/09297049.2013.792332] [Medline: 23638628]
54. Eom H, Kim KK, Lee S, Hong Y, Heo J, Kim J, et al. Development of virtual reality continuous performance test utilizing social cues for children and adolescents with attention-deficit/hyperactivity disorder. Cyberpsychol Behav Soc Netw 2019 Mar;22(3):198-204. [doi: 10.1089/cyber.2018.0377] [Medline: 30672714]
55. Grewe P, Kohsik A, Flentge D, Dyck E, Botsch M, Winter Y, et al. Learning real-life cognitive abilities in a novel 360°-virtual reality supermarket: a neuropsychological study of healthy participants and patients with epilepsy. J Neuroeng Rehabil 2013 Apr 23;10:42 [FREE Full text] [doi: 10.1186/1743-0003-10-42] [Medline: 23618596]
56. Grewe P, Lahr D, Kohsik A, Dyck E, Markowitsch H, Bien C, et al. Real-life memory and spatial navigation in patients with focal epilepsy: ecological validity of a virtual reality supermarket task. Epilepsy Behav 2014 Feb;31:57-66. [doi: 10.1016/j.yebeh.2013.11.014] [Medline: 24361763]
57. Henry M, Joyal C, Nolin P. Development and initial assessment of a new paradigm for assessing cognitive and motor inhibition: the bimodal virtual-reality Stroop. J Neurosci Methods 2012 Sep 30;210(2):125-131. [doi: 10.1016/j.jneumeth.2012.07.025] [Medline: 22897988]
58. Huang K. Exergaming executive functions: an immersive virtual reality-based cognitive training for adults aged 50 and older. Cyberpsychol Behav Soc Netw 2020 Mar;23(3):143-149. [doi: 10.1089/cyber.2019.0269] [Medline: 31794673]
59. Ijaz K, Ahmadpour N, Naismith SL, Calvo RA. An immersive virtual reality platform for assessing spatial navigation memory in predementia screening: feasibility and usability study. JMIR Ment Health 2019 Sep 03;6(9):e13887 [FREE Full text] [doi: 10.2196/13887] [Medline: 31482851]
60. Iriarte Y, Diaz-Orueta U, Cueto E, Irazustabarrena P, Banterla F, Climent G. AULA-advanced virtual reality tool for the assessment of attention: normative study in Spain. J Atten Disord 2016 Jun;20(6):542-568. [doi: 10.1177/1087054712465335] [Medline: 23239784]
61. Kang YJ, Ku J, Han K, Kim SI, Yu TW, Lee JH, et al. Development and clinical trial of virtual reality-based cognitive assessment in people with stroke: preliminary study. Cyberpsychol Behav 2008 Jun;11(3):329-339. [doi: 10.1089/cpb.2007.0116] [Medline: 18537503]
62. Liao Y, Chen I, Lin Y, Chen Y, Hsu W. Effects of virtual reality-based physical and cognitive training on executive function and dual-task gait performance in older adults with mild cognitive impairment: a randomized control trial. Front Aging Neurosci 2019 Jul 16;11:162 [FREE Full text] [doi: 10.3389/fnagi.2019.00162] [Medline: 31379553]
63. Lo Priore C, Castelnuovo G, Liccione D, Liccione D. Experience with V-STORE: considerations on presence in virtual environments for effective neuropsychological rehabilitation of executive functions. Cyberpsychol Behav 2003 Jun;6(3):281-287. [doi: 10.1089/109493103322011579] [Medline: 12855084]
64. Mondellini M, Arlati S, Pizzagalli S, Greci L, Sacco M, Arlati S. Assessment of the Usability of an Immersive Virtual Supermarket for the Cognitive Rehabilitation of Elderly Patients: A Pilot Study on Young Adults. 2018 Presented at: 201IEEE 6th international conference on serious games and applications for health (SeGAH); 2018; TU Wien p. A. [doi: 10.1109/segah.2018.8401313]
65. Negu A, Jurma A, David D. Virtual-reality-based attention assessment of ADHD: ClinicaVR: Classroom-CPT versus a traditional continuous performance test. Child Neuropsychol 2017 Aug;23(6):692-712. [doi: 10.1080/09297049.2016.1186617] [Medline: 27258210]
66. Nolin P, Martin C, Bouchard S. Assessment of inhibition deficits with the virtual classroom in children with traumatic brain injury: a pilot-study. Stud Health Technol Inform 2009;144:240-242. [Medline: 19592772]

XSL•FO
RenderX

67. Nolin P, Stipanicic A, Henry M, Lachapelle Y, Lussier-Desrochers D, Rizzo A. ClinicaVR: Classroom-CPT: a virtual reality tool for assessing attention and inhibition in children and adolescents. Comput Hum Behav 2016:A [FREE Full text] [doi: 10.1016/j.chb.2016.02.023]
68. Ouellet, Boller B, Corriveau-Lecavalier N, Cloutier S, Belleville S. The virtual shop: a new immersive virtual reality environment and scenario for the assessment of everyday memory. J Neurosci Methods 2018 Jun 01;303:126-135 [FREE Full text] [doi: 10.1016/j.jneumeth.2018.03.010] [Medline: 29581009]
69. Parsons TD, Courtney CG, Dawson ME. Virtual reality Stroop task for assessment of supervisory attentional processing. J Clin Exp Neuropsychol 2013 Oct;35(8):812-826. [doi: 10.1080/13803395.2013.824556] [Medline: 23961959]
70. Parsons TD, Carlew AR. Bimodal virtual reality Stroop for assessing distractor inhibition in autism spectrum disorders. J Autism Dev Disord 2016 Apr 27;46(4):1255-1267. [doi: 10.1007/s10803-015-2663-7] [Medline: 26614084]
71. Parsons T, McMahan T. An initial validation of the Virtual Environment Grocery Store. J Neurosci Methods 2017 Nov 01;291:13-19 [FREE Full text] [doi: 10.1016/j.jneumeth.2017.07.027] [Medline: 28782630]
72. Parsons TD, Courtney CG. Interactions between threat and executive control in a virtual reality Stroop task. IEEE Trans Affective Comput 2018 Jan 1;9(1):66-75. [doi: 10.1109/taffc.2016.2569086]
73. Parsons T, Barnett M. Virtual apartment Stroop task: Comparison with computerized and traditional Stroop tasks. J Neurosci Methods 2018 Nov 01;309:35-40. [doi: 10.1016/j.jneumeth.2018.08.022] [Medline: 30144496]
74. Parsons TD, Barnett M. Virtual apartment-based Stroop for assessing distractor inhibition in healthy aging. Appl Neuropsychol Adult 2019 Oct 04;26(2):144-154. [doi: 10.1080/23279095.2017.1373281] [Medline: 28976213]
75. Plechatá A, Sahula V, Fayette D, Fajnerová I. Age-related differences with immersive and non-immersive virtual reality in memory assessment. Front Psychol 2019 Jun 11;10:1330 [FREE Full text] [doi: 10.3389/fpsyg.2019.01330] [Medline: 31244729]
76. Plotnik M, Doniger G, Bahat Y, Gottleib A, Gal O, Arad E. Immersive Trail Making: Construct Validity of an Ecological Neuropsychological Test. 2017 Presented at: 2017 International Conference on Virtual Rehabilitation (ICVR); 2017; Canada. [doi: 10.1109/icvr.2017.8007501]
77. Proulx C, Cabral A, Choudhury N, Debergue P. Acceptability study of a novel immersive cognitive care platform for remediation of cognitive deficits. Annu Rev CyberTherapy Telemed 2018:43-49.
78. Rizzo A, Buckwalter J, Bowerly T, Van Der Zaag C, Humphrey L, Neumann U, et al. The virtual classroom: a virtual reality environment for the assessment and rehabilitation of attention deficits. CyberPsychol Behav 2000 Jun;3(3):483-499. [doi: 10.1089/10949310050078940]
79. Robitaille N, Jackson P, Hébert LJ, Mercier C, Bouyer L, Fecteau S, et al. A virtual reality avatar interaction (VRai) platform to assess residual executive dysfunction in active military personnel with previous mild traumatic brain injury: proof of concept. Disabil Rehabil Assist Technol 2017 Oct;12(7):758-764. [doi: 10.1080/17483107.2016.1229048] [Medline: 27677827]
80. Serino S, Morganti F, Colombo D, Pedroli E, Cipresso P, Riva G. Disentangling the contribution of spatial reference frames to executive functioning in healthy and pathological aging: an experimental study with virtual reality. Sensors (Basel) 2018 Jun 01;18(6):A [FREE Full text] [doi: 10.3390/s18061783] [Medline: 29865196]
81. Tarnanas I, Mouzakidis C, Schlee W. Functional Impairment in Virtual-Reality-Daily-Living-Activities as a Defining Feature of Amnestic MCI: Cognitive and Psychomotor Correlates. 2013 Presented at: 2013 International Conference on Virtual Rehabilitation (ICVR); 2013; Philadelphia. [doi: 10.1109/icvr.2013.6662099]
82. Tarnanas I, Schlee W, Tsolaki M, Müri R, Mosimann U, Nef T. Ecological validity of virtual reality daily living activities screening for early dementia: longitudinal study. JMIR Serious Games 2013 Aug 06;1(1):e1 [FREE Full text] [doi: 10.2196/games.2778] [Medline: 25658491]
83. Tarnanas I, Tsolaki M, Nef T, M Müri R, Mosimann UP. Can a novel computerized cognitive screening test provide additional information for early detection of Alzheimer's disease? Alzheimers Dement 2014 Nov 20;10(6):790-798. [doi: 10.1016/j.jalz.2014.01.002] [Medline: 24656838]
84. Voinescu A, Fodor L, Fraser D, Mejías M, David D. Exploring the Usability of Nesplora Aquarium, a Virtual Reality System for Neuropsychological Assessment of Attention and Executive Functioning. 2019 Presented at: IEEE Conference on Virtual Reality and 3D User Interfaces (VR); 2019; Japan. [doi: 10.1109/vr.2019.8798191]
85. Yasuda K, Muroi D, Ohira M, Iwata H. Validation of an immersive virtual reality system for training near and far space neglect in individuals with stroke: a pilot study. Top Stroke Rehabil 2017 Oct;24(7):533-538. [doi: 10.1080/10749357.2017.1351069] [Medline: 28701101]
86. Yeh S, Tsai C, Fan Y, Liu P, Rizzo A. An Innovative ADHD Assessment System Using Virtual Reality. 2012 Presented at: IEEE-EMBS Conference on Biomedical Engineering and Sciences; 2012; Malaysia. [doi: 10.1109/iecbes.2012.6498026]
87. Yeh S, Chen Y, Tsai C, Rizzo A. An Innovative Virtual Reality System for Mild Cognitive Impairment: Diagnosis and Evaluation. 2012 Presented at: 2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences; 2012; Malaysia. [doi: 10.1109/iecbes.2012.6498023]
88. United Nations Statistical Office. Provisional Guidelines on Standard International Age Classification. NY: United Nations; 1982.

XSL•FO
RenderX

89. Kolk A, Saard M, Pertens L, Kallakas T, Sepp K, Kornet K. Structured model of neurorehab: a pilot study of modern multitouch technology and virtual reality platforms for training sociocognitive deficit in children with acquired brain injury. Appl Neuropsychol Child 2019;8(4):326-332. [doi: 10.1080/21622965.2018.1486193] [Medline: 30295544]
90. Rand D, Weiss P, Katz N. Training multitasking in a virtual supermarket: a novel intervention after stroke. Am J Occup Ther 2009;63(5):535-542. [doi: 10.5014/ajot.63.5.535] [Medline: 19785252]
91. Howett D, Castegnaro A, Krzywicka K, Hagman J, Marchment D, Henson R, et al. Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation. Brain 2019 Jun 01;142(6):1751-1766 [FREE Full text] [doi: 10.1093/brain/awz116] [Medline: 31121601]
92. Liao Y, Tseng H, Lin Y, Wang C, Hsu W. Using virtual reality-based training to improve cognitive function, instrumental activities of daily living and neural efficiency in older adults with mild cognitive impairment. Eur J Phys Rehabil Med 2020 Feb;56(1):47. [doi: 10.23736/s1973-9087.19.05899-4]
93. Yoo S, Weatherall A, Wong G, Scott S, Menezes M, Wood N. Clinician Perspective on VR Games for Managing Periprocedural Anxiety in Children. 2019 Presented at: OZCHI'19: Proceedings of the 31st Australian Conference on Human-Computer-Interaction; December 2019; Fremantle, WA. [doi: 10.1145/3369457.3369522]
94. Biffi E, Maghini C, Marelli A, Diella E, Panzeri D, Cesareo A. Immersive Virtual Reality Platform for Cerebral Palsy Rehabilitation. 2016 Presented at: REHAB '16: Proceedings of the 4th Workshop on ICTs for improving Patients Rehabilitation Research Techniques; October 2016; Lisbon. [doi: 10.1145/3051488.3051497]
95. Pas E, Johnson S, Larson K, Brandenburg L, Church R, Bradshaw C. Reducing behavior problems among students with autism spectrum disorder: coaching teachers in a mixed-reality setting. J Autism Dev Disord 2016 Dec;46(12):3640-3652. [doi: 10.1007/s10803-016-2898-y] [Medline: 27590663]
96. Yalon-Chamovitz S, Weiss P. Virtual reality as a leisure activity for young adults with physical and intellectual disabilities. Res Dev Disabil 2008;29(3):273-287 [FREE Full text] [doi: 10.1016/j.ridd.2007.05.004] [Medline: 17590313]
97. Oliveira CR, Lopes Filho BJP, Esteves CS, Rossi T, Nunes DS, Lima MMBMP, et al. Neuropsychological assessment of older adults with virtual reality: association of age, schooling, and general cognitive status. Front Psychol 2018 Jun 29;9:1085 [FREE Full text] [doi: 10.3389/fpsyg.2018.01085] [Medline: 30008689]
98. Liu L, Yu B, Han M, Yuan S, Wang N. Mild cognitive impairment understanding: an empirical study by data-driven approach. BMC Bioinformatics 2019 Dec 24;20(Suppl 15):481 [FREE Full text] [doi: 10.1186/s12859-019-3057-1] [Medline: 31874606]
99. Ávila-Villanueva M, Maestú F, Fernández-Blázquez M. Internal consistency over time of subjective cognitive decline: drawing preclinical Alzheimer's disease trajectories. JAD 2018 Oct 16;66(1):173-183. [doi: 10.3233/jad-180307]
100. Lauriola M, Mangiacotti A, D'Onofrio G, Cascavilla L, Paris F, Ciccone F, et al. Late-life depression versus amnestic mild cognitive impairment: Alzheimer's disease incidence in 4 years of follow-up. Dement Geriatr Cogn Disord 2018;46(3-4):140-153. [doi: 10.1159/000492489] [Medline: 30199883]
101. Pimontel MA, Rindskopf D, Rutherford BR, Brown PJ, Roose SP, Sneed JR. A meta-analysis of executive dysfunction and antidepressant treatment response in late-life depression. Am J Geriatr Psychiatry 2016 Jan;24(1):31-41 [FREE Full text] [doi: 10.1016/j.jagp.2015.05.010] [Medline: 26282222]
102. Rane LJ, Fekadu A, Papadopoulos AS, Wooderson SC, Poon L, Markopoulou K, et al. Psychological and physiological effects of caring for patients with treatment-resistant depression. Psychol Med 2012 Jan 17;42(9):1825-1833. [doi: 10.1017/s0033291711003035]
103. Huygelier H, Schraepen B, van Ee R, Vanden Abeele V, Gillebert CR. Acceptance of immersive head-mounted virtual reality in older adults. Sci Rep 2019 Mar 14;9(1):4519 [FREE Full text] [doi: 10.1038/s41598-019-41200-6] [Medline: 30872760]
104. Roberts AR, De Schutter B, Franks K, Radina ME. Older adults' experiences with audiovisual virtual reality: perceived usefulness and other factors influencing technology acceptance. Clin Gerontol 2019 Mar 05;42(1):27-33. [doi: 10.1080/07317115.2018.1442380] [Medline: 29505343]
105. Appel L, Appel E, Bogler O, Wiseman M, Cohen L, Ein N, et al. Older adults with cognitive and/or physical impairments can benefit from immersive virtual reality experiences: a feasibility study. Front Med (Lausanne) 2019 Jan 15;6:329 [FREE Full text] [doi: 10.3389/fmed.2019.00329] [Medline: 32010701]
106. Wellsby M, Pexman PM. Developing embodied cognition: insights from children's concepts and language processing. Front Psychol 2014 May 28;5:506 [FREE Full text] [doi: 10.3389/fpsyg.2014.00506] [Medline: 24904513]
107. Kuehn E, Perez-Lopez MB, Diersch N, Döhler J, Wolbers T, Riemer M. Embodiment in the aging mind. Neurosci Biobehav Rev 2018 Mar;86:207-225 [FREE Full text] [doi: 10.1016/j.neubiorev.2017.11.016] [Medline: 29175306]
108. Islam R, Lee Y, Jaloli M, Muhammad I, Zhu D, Quarles J. Automatic Detection of Cybersickness from Physiological Signal in a Virtual Roller Coaster Simulation. 2020 Presented at: Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VRW); March 2020; Atlanta. [doi: 10.1109/vrw50115.2020.00175]
109. Magaki T, Vallance M. Developing an Accessible Evaluation Method of VR Cybersickness. 2019 Presented at: 26th IEEE Conference on Virtual Reality and 3D User Interfaces; March 2019; Osaka. [doi: 10.1109/vr.2019.8797748]
110. Kim J, Son J, Leem H, Lee S. Psychophysiological alteration after virtual reality experiences using smartphone-assisted head mount displays: an EEG-based source localization study. Appl Sci 2019 Jun 19;9(12):2501. [doi: 10.3390/app9122501]

XSL•FO
RenderX

111. Houzangbe S, Christmann O, Gorisse G, Richir S. Effects of voluntary heart rate control on user engagement and agency in a virtual reality game. Virtual Reality 2020 Mar 04;24(4):665-681. [doi: 10.1007/s10055-020-00429-7]
112. Patrão B, Pedro S, Menezes P. Human emotions and physiological signals: a classroom experiment. Int J Onl Eng 2016 Apr 28;12(04):37. [doi: 10.3991/ijoe.v12i04.5098]
113. Hinkle L, Khoshhal K, Metsis V. Physiological Measurement for Emotion Recognition in Virtual Reality. 2019 Presented at: 2nd International Conference on Data Intelligence and Security; June 2019; Texas. [doi: 10.1109/icdis.2019.00028]
114. Pinto M, Melo M, Bessa M. Use of the Physiological Response to Improve the Gaming Experience. 2018 Presented at: ICGI 2018: International Conference on Graphics and Interaction; November 2018; Lisbon. [doi: 10.1109/itcgi.2018.8602977]

Abbreviations

AD: Alzheimer's disease
ADHD: attention deficit hyperactivity disorder
ADL: activities of daily living
BE: behavior or ecological measures
CAVE: Cave Automatic Virtual Environment
EDA: electrodermal activity
EF: executive function
HMD: head-mounted display
HRV: heart rate variability
IADL: instrumental activities of daily living
MCI: mild cognitive impairment
NIRS: near-infrared spectroscopy
NLTK: Natural Language Toolkit
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SCD: subjective cognitive decline
TBI: traumatic brain injury
VE: virtual environment
VR: virtual reality

Edited by G Eysenbach; submitted 09.05.21; peer-reviewed by J Hamid, H Lukosch; comments to author 31.05.21; revised version received 31.08.21; accepted 19.09.21; published 24.11.21.

Please cite as:
Kim E, Han J, Choi H, Prié Y, Vigier T, Bulteau S, Kwon GH
Examining the Academic Trends in Neuropsychological Tests for Executive Functions Using Virtual Reality: Systematic Literature Review
JMIR Serious Games 2021;9(4):e30249
URL: https://games.jmir.org/2021/4/e30249
doi:10.2196/30249
PMID:34822341

©Euisung Kim, Jieun Han, Hojin Choi, Yannick Prié, Toinon Vigier, Samuel Bulteau, Gyu Hyun Kwon. Originally published in JMIR Serious Games (https://games.jmir.org), 24.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Virtual Supermarket Program for the Screening of Mild Cognitive Impairment in Older Adults: Diagnostic Accuracy Study

Mingli Yan[1], MSN; Huiru Yin[1], Prof Dr; Qiuyan Meng[1], MSN; Shuo Wang[1], PhD; Yiwen Ding[1], MSN; Guichen Li[1], PhD; Chunyan Wang[2], Prof Dr; Li Chen[1], Prof Dr

[1]School of Nursing, Jilin University, Changchun, China

[2]Senior Officials Inpatient Ward, First Hospital of Jilin University, Changchun, China

Corresponding Author:
Li Chen, Prof Dr
School of Nursing
Jilin University
965 Xinjiang Street
Changchun, 130021
China
Phone: 86 043185619366
Email: chen_care@126.com

Abstract

Background: Mild cognitive impairment (MCI) is often a precursor of dementia, and patients with MCI develop dementia at a higher rate than healthy older adults. Early detection of cognitive decline at the MCI stage supports better planning of care and interventions. At present, the use of virtual reality (VR) in screening for MCI in older adults is promising, but there is little evidence regarding the use of virtual supermarkets to screen for MCI.

Objective: The objectives of this study are to validate a VR game–based test, namely, the Virtual Supermarket Program (VSP), for differentiating patients with MCI and healthy controls and to identify cutoff scores for different age levels.

Methods: Subjects were recruited from several nursing homes and communities in Changchun, China. They were divided into a healthy control group (n=64) and an MCI group (n=62). All subjects were administered the VSP and a series of neuropsychological examinations. The study determined the optimal cutoff, discriminating validity, concurrent validity, and retest reliability of the VSP. We used the area under the receiver operating characteristic curve (AUC) to evaluate the discriminating validity and obtain the optimal cutoff values. Pearson correlation analysis and the intraclass correlation coefficient were used to evaluate the concurrent validity and retest reliability, respectively.

Results: A cutoff score of 46.4 was optimal for the entire sample, yielding a sensitivity of 85.9% and specificity of 79.0% for differentiating individuals with MCI and healthy controls, and the AUC was 0.870 (95% CI 0.799-0.924). The median index of VSP score was 51.1 (range 42.6-60.0). There was a moderate positive correlation between the VSP total score and Mini-Mental State Examination score (r=0.429, P<.001). There was a strong positive correlation between VSP total score and Montreal Cognitive Assessment score (r=0.645, P<.001). The retest reliability of the VSP was feasible (r=0.588, P=.048).

Conclusions: The VSP is interesting and feasible for subjects. It shows high sensitivity and specificity for the identification of MCI in older adults, which makes it a promising screening method. The VSP may be generalized to older adults in other countries, although some cultural adaptation may be necessary.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2000040074; https://www.chictr.org.cn/showprojen.aspx?proj=64639

(JMIR Serious Games 2021;9(4):e30919) doi:10.2196/30919

KEYWORDS

virtual reality; mild cognitive impairment; dementia; ambient intelligence; digital health; elderly population; aging

Introduction

Dementia is becoming more prevalent worldwide. According to the “World Alzheimer Report 2018” [1], 50 million people worldwide had dementia in 2018. As the most populous country, China has experienced the greatest burden of dementia. The total number of people with dementia in China will reach 23.3 million by 2030 [2], and in 2020, the burden of dementia-related

XSL•FO
RenderX

disability and care in China reached US $250 billion, accounting for nearly one-fifth of the global dementia-related costs [3]. However, there is no effective treatment for dementia; therefore, increasing attention has been given to the precursors of dementia. Mild cognitive impairment (MCI) is often a precursor of dementia, but it can also be due to other pathologies (such as Parkinson disease and tumors) [4,5]. Although not all forms of MCI lead to dementia, the mean annual conversion rate of MCI to dementia is approximately 10%, which is far higher than the annual incidence (1%-2%) in healthy older adults [6,7]. Relevant studies have shown that interventions for MCI can effectively delay or even reverse cognitive decline [8]. Therefore, early detection of cognitive decline at the MCI stage is vital for better planning of care and interventions.

The commonly used cognitive screening methods are paper-and-pencil neuropsychological testing [9] and computerized cognitive testing [10,11]. Compared to neuropsychological testing, the advantages of computerized cognitive testing include standardized administration procedures and presentation of stimuli and accurate measurement of response time [12,13]. Virtual reality (VR) is widely used in research with computerized cognitive testing for MCI [14,15]. VR is an ecologically valid and informative evaluation instrument and provides an opportunity to improve cognitive screening [16-18]. Studies have shown that VR is recommended for the screening of older adults with MCI or dementia [15,19]. Among older adults with cognitive impairment screened through VR, virtual supermarket has been accepted [20] and could be used to detect cognitive impairment in older adults [21,22].

Virtual supermarkets are a novel method for screening for MCI and assessing cognition through shopping-related activities. Related research using virtual supermarkets is effective in screening for MCI [20,22,23]. Atkins et al [24] developed the Virtual Reality Functional Capacity Assessment Tool, which assesses functional abilities related to shopping, and found a significant relationship between task performance and cognitive performance. Werner et al [22] developed the Virtual Action Planning-Supermarket, which allows subjects to navigate the store, purchase 7 items, and then pay for them, and they distinguished healthy older adults and MCI patients. Significant differences were found between healthy older adults and patients with MCI in the Virtual Action Planning-Supermarket measures. However, they assessed few cognitive domains and screened for cognitive impairment by assessing only executive function. Zygouris et al [20] developed a screening tool called the Virtual Supermarket Test to detect MCI and assessed 4 cognitive abilities in older adults through four shopping-related tasks. The program can operate on Android-based tablet devices or computers. The discriminative validity of the program was good in older adults over 70 years of age. However, Zygouris et al [20] developed virtual supermarket as a cognitive training application that does not specifically screen for MCI, and it screens few cognitive domains. Nevertheless, the literature supports virtual supermarket in screening for MCI [23]. However, the virtual supermarkets developed in other countries may not be suitable screens for older Chinese adults, and there is no virtual supermarket software to screen for MCI in China.

The Virtual Supermarket Program (VSP) developed in this study screened subjects with MCI by assessing their ability to shop in conditions relevant to daily life. Subjects could actively explore the VSP environment while performing the task. Compared with other versions of the virtual supermarket, VSP provides a more comprehensive screening, involves a rich environment with VR technology, and is more in line with the living and cultural habits of older adults in China. Therefore, we aimed to verify the feasibility of VSP for screening for MCI in older adults in China.

Methods

Study Design

We evaluated the feasibility of VSP for screening patients with MCI (diagnosed in accordance with Petersen criteria) [25] over 60 years of age in China. The study followed the STARD checklist for the reporting of studies of diagnostic test accuracy [26]. This study was conducted between December 2020 and February 2021. The study protocol was approved by the School of Nursing (approval 2020082805), Jilin University, and it has been registered in the Chinese Clinical Trial Registry (registration ChiCTR2000040074). All subjects were informed about the purpose of the study before providing their verbal consent. The VSP used in this study was a new tool developed by the authors' team at the School of Nursing, Jilin University, China, and a company was entrusted to provide technical support.

Population

Subjects were from several nursing homes and communities in Changchun, China. The inclusion criteria of the subjects with MCI were as follows: (1) MCI diagnosis in accordance with the Petersen criteria [25]; (2) age≥60 years; (3) a Montreal Cognitive Assessment (MoCA) scale [27] score of <26; (4) a Mini-Mental State Examination (MMSE) [28] score of ≥24; (5) no impairment in functional daily living activities (Activities of Daily Living [ADL] scale [29] score of ≤26); (6) absence of psychiatric illnesses, with particular reference to depressive symptoms (15-item Geriatric Depression Scale [GDS-15] [30] score of ≤8); and (7) absence of severe auditory/visual loss that can prevent the use of technological devices and from executing the VSP. The inclusion criteria of subjects in the healthy control group were inclusion criteria (2) and (4)-(7) from the aforementioned list and an MoCA score of ≥26.

Sample Size

We calculated the sample size using equation 1 [31], and we set the confidence level (1-α) at 95% and the allowable error (δ) of sensitivity and specificity at 10%. We estimated the sample size required for sensitivity (the sample size for the MCI group) and the sample size required for specificity (the sample size for the healthy control group).

The sensitivity and specificity of the tools developed by Zygouris et al [21] were 82.35% and 95.24%, respectively, which were used to calculate the sample size for the MCI group and healthy control group in this study. We predicted a dropout rate of 10%. The sample size of the MCI group was 62, and that

XSL•FO
RenderX

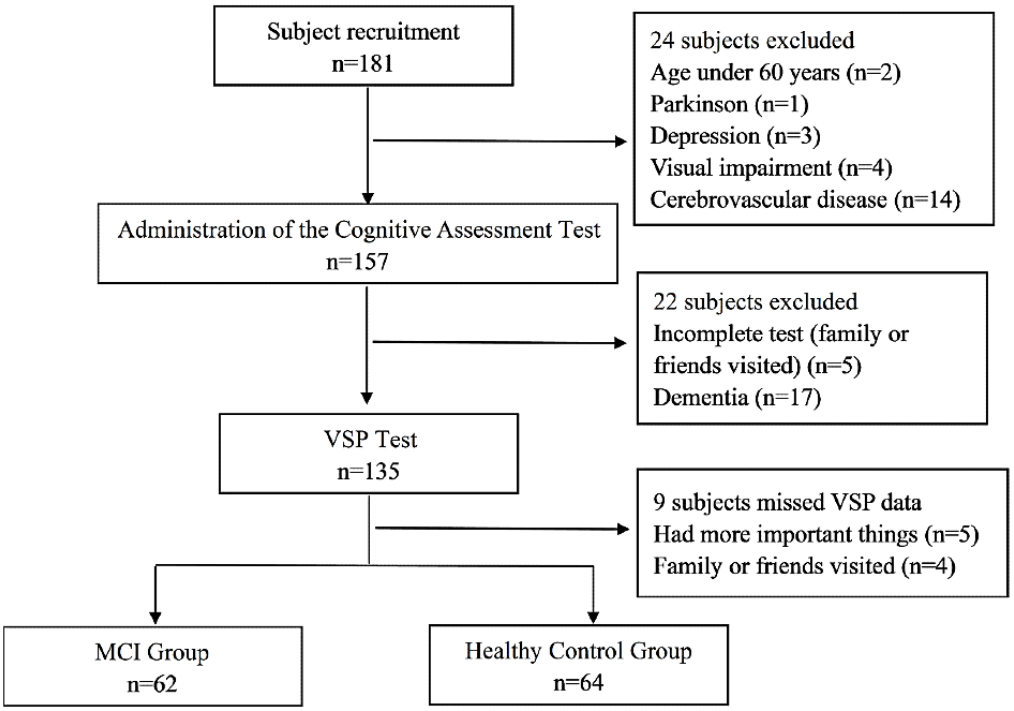
of the healthy control group was 20 in this study. Therefore, we needed to enroll a minimum of 82 subjects.

Study Design

This study recruited subjects over the age of 60 years in nursing homes and communities. The subjects who volunteered for the study were first administered a neuropsychological test to screen for MCI in accordance with the Peterson criteria [25]. The test took place in a quiet and bright room. First, baseline characteristics (age, education level, height and weight, and previous computer use before VSP) and past medical history were collected. Older adults with depression were excluded using the GDS-15. Then, the Clinical Dementia Rating (CDR) and MMSE were used to exclude patients with dementia. Thereafter, MoCA and functional status (ADL scale) measures were collected. Last, VSP tests were performed for all subjects. Prior to each task in the VSP, the subjects were repeatedly instructed on the task, and operation instructions were available throughout the test through the user interface. The VSP was managed by independent researchers, who were blinded to the subjects' cognitive status. All subjects used the VSP application under the same circumstances.

All data were obtained directly from the subjects and included neuropsychological test scores, VSP scores, and general information. The assessors were blinded to the cognitive status or grouping of the subjects. Figure 1 shows a flowchart of subject recruitment and testing.

Figure 1. Flowchart of subject recruitment and testing. MCI: mild cognitive impairment; VSP: Virtual Supermarket Program.



Neuropsychological Assessments

Subjects were administered a battery of neuropsychological tests, including the CDR, MMSE, GDS-15, ADL, and MoCA. The reference standard results were available to the testers. The neuropsychological tests used in this study were as follows.

CDR

The CDR assesses 6 abilities on a 5-point scale [32]: no dementia (CDR=0), questionable dementia (CDR=0.5), mild dementia (CDR=1.0), moderate dementia (CDR=2.0), and severe dementia (CDR=3.0). Finally, the results of the 6 ability ratings are combined into one rating on the basis of the scoring standard, and there is no dementia when the CDR was zero [33].

MMSE

The MMSE has 30 questions, and the highest potential score is 30. The optimal cutoff values differ for subjects with different educational levels. Subjects who are illiterate and scored ≤17 points, those with a primary school education who scored ≤20 points, and those with a high school education or above who scored ≤24 points are classified as having dementia [34].

GDS-15

The GDS-15 comprises 15 questions to which subjects are asked to respond with “yes” or “no.” The total score can range from 0 to 15, and a total score of >8 is associated with depressive symptoms [30].

MoCA Scale

The MoCA scale comprises 32 questions, and the highest potential score is 30. If the number of years of education is ≤12 years, 1 point is added to the actual total score, and a final total score of >26 is considered normal [27].

ADL Scale

The ADL scale comprises 14 items that are scored at 4 levels. The total score can range from 14 to 56, and a total ADL score of >26 points indicate an impaired ability to perform activities of daily living [29].

XSL•FO
RenderX

Virtual Supermarket Program

We developed the VSP to be in line with Chinese cultural habits and with reference to the virtual supermarket of Zygouris et al [20,21]. We mimicked the MoCA and designed the VSP with VR technology. For instance, for the assessment of temporal orientation, the MoCA scale asks the subjects to identify the current date. In the VSP, we designed the link to input the payment password, and the password was the current date. All the tasks in the VSP were designed in this manner. Shopping-related tasks were used to assess learning and memory, executive functions, language, time orientation, and complex attention of older adults and to explore the feasibility of the VSP for screening MCI in older Chinese adults. There were 9 tasks in the VSP, and it was run on a computer (CPU, i5 or higher; memory, ≥16 GB; independent video memory, ≥2 GB; storage capacity, ≥200 GB) for which subjects did not need VR glasses and a handle to operate. The VSP was operated with a computer mouse and a keyboard (Multimedia Appendix 1). The following tasks were included in the VSP: task 1: memorize shopping list; task 2: look at the map to get directions to the supermarket and make purchases on the list; task 3: recall the sales announced by voice in the supermarket; task 4: label the item that have lost their labels; task 5: determine whether the correct items were purchased; task 6: calculate the amount for the purchased items; task 7: enter the correct payment code (current date); task 8: compare lucky draw numbers at the service center; and task 9: choose the correct bus route home.

The VSP assessed five cognitive domains [35]: learning and memory (tasks 1 and 3), executive functions (tasks 2, 6, and 9), language (tasks 4 and 5), time orientation (task 7), and complex attention (task 8). The VSP score ranges from 0 to 60 (Multimedia Appendix 2). The technician devised a fixed algorithm to calculate the score for each task and the total VSP score on the basis of the accuracy with which the subject completed the task. There was no time limit for subjects to perform each VR task.

Statistical Analysis

Statistical analyses were performed using SPSS Statistics (version 23.0, IBM Corp) and MedCalc (version 19.2, MedCalc Software Ltd). The normality of data distributions was assessed using the Shapiro–Wilk test. Mean (SD) values were used to describe continuous variables that were normally distributed, and median (IQR) values were used to describe the continuous variables that were not normally distributed. Descriptive statistics included n (%) values for categorical variables. Independent samples *t* tests were used to compare continuous variables that were normally distributed between the two groups, nonparametric Mann-Whitney *U* rank sum tests were used to compare continuous variables that were not normally distributed between the two groups, and chi-square tests were used for the categorical variables. The feasibility of VSP screening for MCI was verified by measures of optimal cutoff, discriminating validity, concurrent validity and test-retest reliability for the VSP. Discriminating validity was examined through receiver operating characteristic (ROC) analysis of VSP total scores. The area under the receiver operating characteristic curve (AUC) can range between 0 and 1. The closer the AUC to 1, the better the diagnosis [36]. Concurrent validity was examined through Pearson correlation analyses of VSP total scores against MMSE scores [34] and MoCA scores [27]. Test-retest reliability was assessed with an intraclass correlation coefficient (ICC). A 2-sided *P* value of <.05 indicated statistical significance.

Results

Demographics and Characteristics by Cognitive Status

The demographics and cognitive characteristics of the subjects are presented in Table 1. In the actual process of recruitment, the healthy control group collected more than the minimum sample size we previously calculated. Considering that a large number of samples can improve the test efficiency, we expanded the sample size. Initially, 181 participants were recruited (Figure 1). No adverse events occurred during the study. Finally, the study included 126 older adults (45 male and 91 female). The mean age was 77.12 years (range 61-94 years), and years of education ranged from 0 to 16 years. In total, 62 older adults were enrolled in the MCI group and 64 were enrolled in the healthy control group. There was a significant difference in age between the 2 groups (*P*=.02). No between-group differences in sex distribution (*P*=.06), years of education, or computer experience (having used a computer at least once) were found (*P*=.08). As expected, the 2 groups had significantly different MMSE scores and MoCA scores (*P*<.001).

XSL•FO
RenderX

Table 1. Demographics and cognitive characteristics of subjects (N=126).

Characteristics	Healthy controls (n=64)	Patients with mild cognitive impairment (n=62)	*T*, *z*, or chi-square test	*P* value
Age (years), mean (SD)	73.5 (16.0)	82.00 (15.50)	–2.486[a]	.01
Sex, n (%)			3.613[b]	.06
Female	13.0 (20.3)	22.0 (35.5)		
Male	51.0 (79.7)	40.0 (64.5)		
Years of education, n (%)			6.875[b]	.08
Primary school	10.0 (15.6)	4.0 (6.5)		
Middle school	13.0 (20.3)	19.0 (30.6)		
High school	13.0 (20.3)	20.0 (32.3)		
University	28.0 (43.8)	19.0 (30.6)		
Computer experience (having used a computer at least once), n (%)			2.873[b]	.09
Yes	29.0 (45.3)	19.0 (30.6)		
No	35.0 (54.7)	43.0 (64.9)		
Cognitive assessment score, mean (SD)				
Mini-Mental State Examination score (maximum score=30)	28.5 (1.3)	27.0 (1.6)	5.668[c]	<.001
Montreal Cognitive Assessment score (maximum score=30)	26.6 (1.0)	20.8 (2.7)	16.106[c]	<.001

[a]Mann-Whitney *U* test.

[b]Chi-square test.

[c]Independent samples *t* tests.

Performance in the VSP Between Patients With MCI and Healthy Controls

The performance scores for the individual tasks on the VSP are summarized in Table 2. The mean scores in each VSP task in the healthy control group were higher than those in the MCI group. With the exception of the difference for task 6, these differences were significant. The total VSP score in the MCI group was significantly lower than that in the healthy control group (*P*<.001).

Table 2. Performance scores in the Virtual Supermarket Program.

Task score	Healthy controls (n=64), mean (SD)	Patients with mild cognitive impairment (n=62), mean (SD)	*t* test (independent samples)	*P* value
Task 1	9.9 (0.5)	9.6 (1.0)	1.999	.048
Task 2	4.0 (2.0)	2.3 (2.5)	4.265	<.001
Task 3	6.0 (3.5)	4.1 (3.8)	2.940	.004
Task 4	6.0 (0.5)	5.6 (1.2)	2.087	.04
Task 5	5.6 (0.8)	4.5 (2.0)	4.379	<.001
Task 6	5.8 (0.9)	5.8 (1.1)	0.213	.83
Task 7	5.5 (1.6)	4.7 (2.5)	2.328	.02
Task 8	4.0 (1.4)	2.0 (1.6)	7.407	<.001
Task 9	4.8 (1.1)	2.8 (2.5)	5.708	<.001
Total	51.4 (4.8)	41.3 (7.7)	8.894	<.001

Potential of the VSP to Distinguish Between Patients With MCI and Healthy Controls

An ROC curve analysis was performed with the total VSP scores that were used to distinguish between patients with MCI and healthy controls (Figure 2). The AUC was found to be 0.870 (95% CI 0.799-0.924, *P*<.001). An optimal statistical cutoff was achieved at a cutoff score of 46.4 points (85.9% sensitivity, 79.0% specificity).

Among subjects aged under 85 years, the optimal cutoff was 46.5 points, and the sensitivity and specificity were 72.73% and

XSL•FO
RenderX

86.79%, respectively (AUC=0.840, *P*<.001). Among the subjects aged over 85 years, the optimal cutoff was 47.3 points, yielding a sensitivity of 94.4% and specificity of 80.0% (AUC=0.936, *P*<.001). The ROC curve stratified by age is shown in Figure 3, and the discriminative validity is shown in Table 3.

Figure 2. Receiver operating characteristic curve analysis for the total score in the Virtual Supermarket Program.

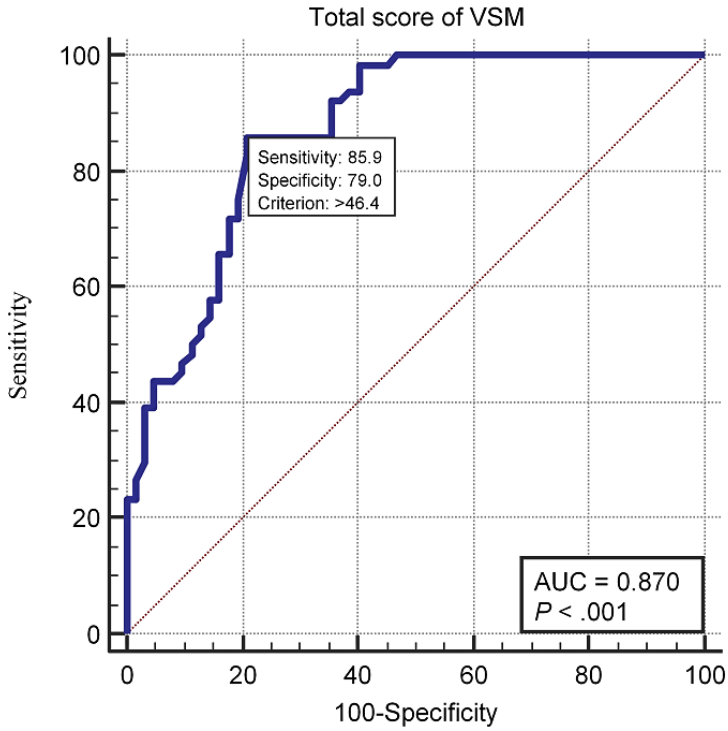
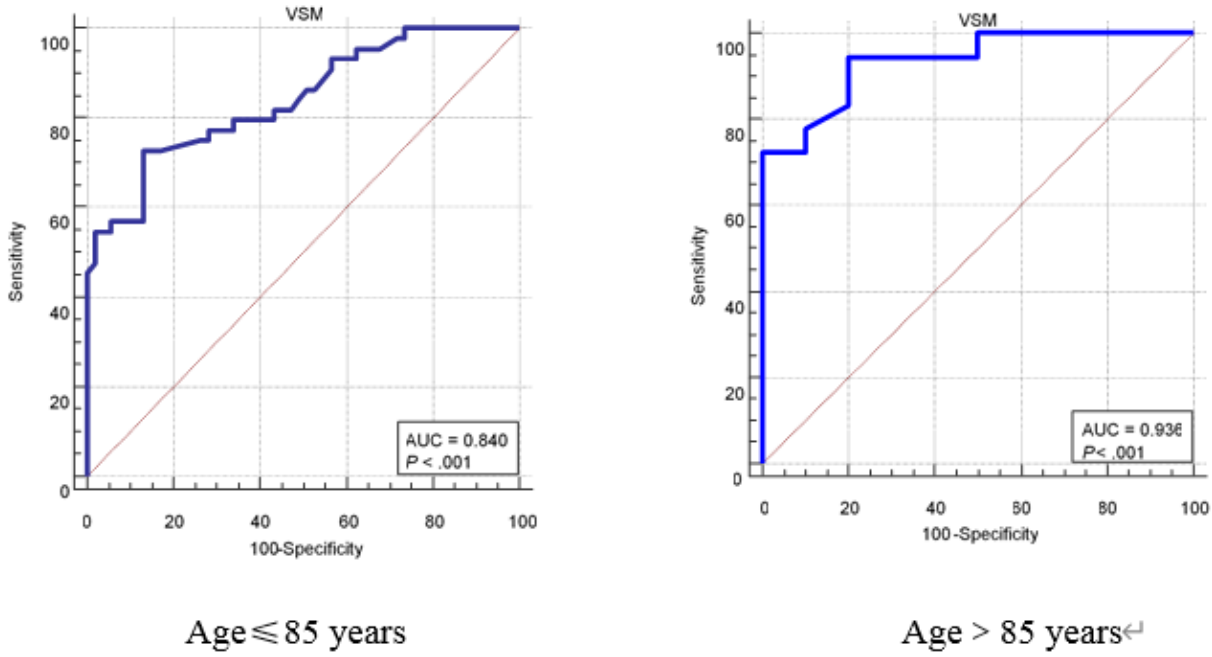


Figure 3. Receiver operating characteristic curve analysis for the total score in the Virtual Supermarket Program among different age groups.



XSL•FO
RenderX

Table 3. The total score in the Virtual Supermarket Program and the discriminant results of age stratification.

	Optimal cutoff	Sensitivity, %	Specificity, %	Area under the receiver operation characteristic curve	95% CI	*P* value	Youden index J
Entire sample	46.4	85.90	79.00	0.870	0.799-0.924	<.001	0.650
Age (years)							
≤85	46.5	72.73	86.79	0.840	0.752-0.907	<.001	0.600
>85	47.3	94.44	80.00	0.936	0.775-0.993	<.001	0.744

Correlations Between the VSP and Routine Cognitive Screening Assessments

There was a moderate positive correlation between the total VSP scores and MMSE scores (r=0.429, P<.001). There was a strong positive correlation between the total VSP scores and MoCA scores (r=0.645, P<.001). In addition, we found a negative correlation of the total VSP scores with age (r=–0.308, P<.001). Table 4 shows the correlation matrix between the different variables.

Table 4. Correlation matrix demonstrating the degree of associations among the Mini-Mental State Examination score, Montreal Cognitive Assessment score, age, education, and total performance scores.

Variable	Age	Education	Mini-Mental State Examination Score	Montreal Cognitive Assessment score	Score in the Virtual Supermarket Program
Age					
r	—[a]	–0.119	–0.342	–0.247	–0.308
P value	—	0.186	<.001	.005	<.001
Education					
r	–0.119	—	0.269	0.110	0.098
P value	.19	—	.002	.22	.27
Mini-Mental State Examination score					
r	–0.342	0.269	—	0.592	0.429
P value	<.001	.002	—	<.001	<.001
Montreal Cognitive Assessment score					
r	–0.247	0.110	0.592	—	0.645
P value	.005	0.22	<.001	—	<.001
Score in the Virtual Supermarket Program					
r	–0.308	0.098	0.429	0.645	—
P value	<.001	0.27	<.001	<.001	—

[a]—: not determined.

Retest Reliability of the VSP

A month after the first test, 9 healthy controls were randomly selected to receive the VSP again, and the results showed that the ICC between total VSP scores was feasible (r=0.588, P=.048).

Discussion

Principal Findings

To our knowledge, this is the first study to use a locally developed virtual supermarket task to screen for MCI among older adults in China. The VSP we designed in this study was a computerized application running on a computer to screen for MCI. Through verification of its screening effectiveness, the VSP had the highest sensitivity and specificity when the optimal cutoff was 46.4. The VSP can effectively distinguish older adults with MCI from healthy controls, and it was more suitable for older adults aged 85 years and above. Performance in the VSP had a strong positive correlation with MoCA scores but a moderate positive correlation with MMSE scores. The retest reliability of the VSP scores was feasible.

Comparing Performance Between Patients With MCI and Healthy Controls

We found that the overall scores and scores for each task were significantly higher in the healthy control group, but task 6 did not show a significant difference between the 2 groups. In task 6, the subjects were asked to calculate the total cost of the items they had purchased. Given the situation of older adults, the settings related to the task list in the system were set to be relatively simple, which is probably why task 6 did not show a

XSL•FO
RenderX

significant difference between the MCI subjects and healthy controls. This is a solvable problem. In the future, we will adjust the difficulty of task 6 by increasing the number of items purchased or adjusting the price of items.

Compared to the healthy controls, the subjects in the MCI group performed worse on tasks (tasks 1, 2, 3, 5, and 9) that involved learning and memory and executive function. The MCI subjects were more likely to forget or buy the wrong item on their shopping list, and they could not perform the shopping tasks very well. This is evidence that these tasks were better at distinguishing between the 2 groups. Task 5 and task 8 represent the ingenuity in our task design. Task 5 involved an assessment of learning and memory, executive function and language that comprehensively evaluated the subjects' cognition. Task 8 evaluated the subjects' complex attention. The subjects needed to use their attention, hearing, and finger reflexes at the same time to complete the task. This task could distinguish the MCI subjects and healthy controls well. Above all, these tasks highlight the effectiveness of the VSP in distinguishing between the MCI subjects and healthy controls.

Discriminating the Validity of the VSP Total Score

In this study, the VSP score could effectively distinguish between MCI subjects and healthy controls. The sensitivity and specificity were 85.90% and 79.00%, respectively. The AUC based on the VSP scores was 0.870 in this study. Compared with other studies, this level of accuracy is relatively high. Boz et al [23] used the Virtual Supermarket, which was developed in Greece [21], to screen for MCI among older Turkish adults. They reported that the sensitivity and specificity were 74.00% and 85.00%, respectively; they did not report the AUC. The sensitivity and specificity of the virtual supermarket task were better in our study than in that by Boz et al [23]. This may have resulted from using Euros on the payment screen, which adds additional complexity and cognitive burden [23]. The system of Cognitive Assessment by Virtual Reality [35] is adapted from the RE@CH module, and Chua et al [37] reported an AUC of 0.821, a sensitivity of 78.2%, and a specificity of 75.7% in the Singaporean population. Compared to the task used by Chua et al [37], the VSP is more in line with the culture and living habits of older Chinese adults, and the sample size in our study was larger.

Age is an important factor in cognition. Among subjects younger than 85 years, the AUC, sensitivity, and specificity were 0.840, 72.73%, and 86.79%, respectively. Among subjects over 85 years of age, the AUC, sensitivity, and specificity were 0.903, 94.44%, and 80.00%, respectively. As people age, their cognitive performance declines. Compared with younger subjects, older subjects committed more task errors and had a harder time concentrating [38]. This may have led to a better discriminative validity of the VSP among individuals over 85 years of age. This finding indicates that our VSP has a wider range of uses and is suitable for older adults of different ages.

Associations Among VSP Scores, Age, Education, and Neuropsychological Testing

VSP scores were negatively correlated with age, which is consistent with the work of Chua et al [37,39]. Considered a task involving multiple demands, the VSP was formed by a series of activities consisting of goals during its execution, and it requires activation and cooperation of multiple cognitive domains [40]. The older the subjects, worse the VSP performance and scores. There was no correlation between VSP scores and education level. The literature is not conclusive regarding the impact of education level on cognition [41,42]. The VSP that was developed and designed in this study was suitable for older adults with different educational levels and was intended to have wide applicability. The tasks closely reflect to the habits of older adults, and successful performance does not necessarily depend on education. Kang et al [43] did not report an effect of education levels among older adults with a multiple-order VR task for neuropsychological assessment.

The total VSP performance scores showed a moderate positive correlation with MMSE scores (r=0.429, P<.001) and a strong positive correlation with MoCA scores (r=0.645, P<.001). The correlation between VSP scores and MoCA scores was significantly higher than that between VSP scores and MMSE scores, which was consistent with our expectation. The VSP used in this study was specifically developed for detecting MCI, and the design of VSP mimicked the MoCA scale. Research has shown that the MoCA scale has more advantages than the MMSE scale in detecting visual executive dysfunction [44]. Therefore, the correlation between VSP and MoCA scores may be stronger than that between VSP and MMSE scores. In addition, although the MMSE has a modest specificity for screening MCI, its sensitivity for screening MCI is low, and the capacity to detect MCI converters is poor [28]. This outcome is consistent with that of Oliveira et al, who also showed a weak correlation between MMSE scores and performance in VR research [39]. Chan et al [45] and Chua et al [37] also conducted correlation analyses with MoCA scores when verifying the screening application for MCI, which they had developed, and showed strong correlations between the measures.

Clinical Implications

Our findings provide preliminary evidence that the VSP can be used for MCI screening in Chinese communities and nursing homes, providing a more convenient, concise, and effective tool for MCI screening in China. We introduced several short activities based on VR to assess different cognitive domains, rather than using a single game, which allowed us to assess impairments in the major cognitive domains. The subjects were screened using shopping tasks based on the abovementioned cognitive domains. After completing the task, the obtained scores were compared with the cutoff value of 46.4, which could preliminarily screen whether the subjects had MCI. This screening program could provide rapid screening in Chinese communities and nursing homes. It expands this new method of screening for MCI and provides new ideas and methods for the development of MCI screening software in China. In addition, compared to traditional neuropsychological tests, the VSP has an automatic scoring function and allows data to be extracted and interpreted more easily. The VSP has other advantages, including administration cost savings and better ecological relevance, and is suitable for unsupervised use in a home or clinical setting. We observed that owing to the visual content and interaction style through VR, most subjects found

XSL•FO
RenderX

the task more interesting than traditional screening tools and were more willing to take the test. Yun et al [46] also reported elderly people's interest in VR. Implementation studies are needed in the future to evaluate its rollout in clinical practice.

Study Limitations and Future Research

There are several limitations in the study, which limit the generalizability of the results. First, subjects in the MCI group were older than those in the healthy control group, which may have affected their performance. Age is an important factor in cognition. To further evaluate the use of VSP in cognitive assessment, demographic data matching and expanding the sample size can be incorporated in future studies.

Second, even if some of the subjects had previously used a computer, many of the subjects had limited familiarity with new technologies, particularly computer equipment and operations. Oliveira et al [39] showed that experience with the computer was not a relevant predictor for any of the models investigated. We also need to further verify the impact of new technology on the ability of older adults to engage in the task.

Third, the correlations between VSP subtests assessing specific cognitive domains and existing neuropsychological tests that assess the same domains were not verified. In future research, we need to verify the correlation between the two to verify the feasibility of VSP subtests for assessing specific cognitive domains.

Owing to technical limitations, this study did not verify the feasibility of the time to complete the VSP in screening for MCI. In the future, it will be necessary to improve the assessment technology of VSP stay time in each area to verify the feasibility of the total time of VSP to detect older adults with MCI. In addition, we need to further verify the retest reliability of the VSP in future research. We tested the VSP again a month after the first test, but the community and nursing home were closed owing to the COVID-19 pandemic, and we could not enter and leave freely. Finally, only nine subjects were retested on the VSP. In the future, it will be necessary to further expand the sample size to verify the retest reliability of the VSP.

Conclusions

The VSP we designed in this study is a computerized application running on a computer to identify MCI. This study demonstrated the feasibility of the VSP in distinguishing between MCI and healthy adults among older Chinese adults. The results of this feasibility study are invaluable. Future studies should focus on testing and validating longitudinal data for the ability to track the progression of cognitive decline.

Acknowledgments

We would like to extend our gratitude to Ms Xu Yan and Ms Wan En of the Boyuan·Xiangzhiyuan nursing home, Ms Gao Jie of the Wenxin nursing home, and Ms Xu Jun of the Duoen nursing home, who supported us in various ways, including screening and referring patients for the study. We are thankful to Ms Zhao Lijing of Jilin University, Mr Chen Yanqiang of the Jilin University Xinmin Campus from the retirement office, Ms Han Song of the community health service center of Yong Chang, and Ms Huang Xiaoxia of the Changchun city community for their assistance. This study was supported by National Natural Science Foundation of China (grant 81902295) and Graduate Innovation Found of Jilin University, China (grant 101832020CX321). This study also received funding from “13th Five-Year” Science and Technology Research Project of Education Department of Jilin Province, China (grant JJKH20190006KJ) and Sanitation and Health technology Innovation Project of Health Department of Jilin Province, China (grant 2020J036).

Authors' Contributions

MY and HY contributed equally as first authors. CW and LC contributed equally as coauthors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the VSP.
[DOC File , 7619 KB - games_v9i4e30919_app1.doc]

Multimedia Appendix 2
VSP score of each task and total score.
[DOC File , 33 KB - games_v9i4e30919_app2.doc]

References

1. Patterson C, Lynch C, Bliss A, Lefevre M. World Alzheimer Report 2018. Alzheimer's Disease International. URL: https://www.alzint.org/resource/world-alzheimer-report-2018/ [accessed 2021-11-25]

XSL•FO
RenderX

2. Xu J, Wang J, Wimo A, Fratiglioni L, Qiu C. The economic burden of dementia in China, 1990-2030: implications for health policy. Bull World Health Organ 2017 Jan 01;95(1):18-26 [FREE Full text] [doi: 10.2471/BLT.15.167726] [Medline: 28053361]
3. Jia J, Wei C, Chen S, Li F, Tang Y, Qin W, et al. The cost of Alzheimer's disease in China and re-estimation of costs worldwide. Alzheimers Dement 2018 Apr;14(4):483-491. [doi: 10.1016/j.jalz.2017.12.006] [Medline: 29433981]
4. Weil RS, Costantini AA, Schrag AE. Mild Cognitive Impairment in Parkinson's Disease-What Is It? Curr Neurol Neurosci Rep 2018 Mar 10;18(4):17 [FREE Full text] [doi: 10.1007/s11910-018-0823-9] [Medline: 29525906]
5. Bennett DA, Schneider JA, Bienias JL, Evans DA, Wilson RS. Mild cognitive impairment is related to Alzheimer disease pathology and cerebral infarctions. Neurology 2005 Mar 08;64(5):834-841. [doi: 10.1212/01.WNL.0000152982.47274.9E] [Medline: 15753419]
6. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Author response: Practice guideline update summary: Mild cognitive impairment: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 2018 Aug 21;91(8):373-374. [doi: 10.1212/WNL.0000000000006042] [Medline: 30126885]
7. Roberts RO, Knopman DS, Mielke MM, Cha RH, Pankratz VS, Christianson TJH, et al. Higher risk of progression to dementia in mild cognitive impairment cases who revert to normal. Neurology 2014 Jan 28;82(4):317-325 [FREE Full text] [doi: 10.1212/WNL.0000000000000055] [Medline: 24353333]
8. Delbroek T, Vermeylen W, Spildooren J. The effect of cognitive-motor dual task training with the biorescue force platform on cognition, balance and dual task performance in institutionalized older adults: a randomized controlled trial. J Phys Ther Sci 2017 Jul;29(7):1137-1143 [FREE Full text] [doi: 10.1589/jpts.29.1137] [Medline: 28744033]
9. Knopman DS, Petersen RC. Mild cognitive impairment and mild dementia: a clinical perspective. Mayo Clin Proc 2014 Oct;89(10):1452-1459 [FREE Full text] [doi: 10.1016/j.mayocp.2014.06.019] [Medline: 25282431]
10. Lim YY, Ellis KA, Harrington K, Ames D, Martins RN, Masters CL, The Aibl Research Group. Use of the CogState Brief Battery in the assessment of Alzheimer's disease related cognitive impairment in the Australian Imaging, Biomarkers and Lifestyle (AIBL) study. J Clin Exp Neuropsychol 2012;34(4):345-358. [doi: 10.1080/13803395.2011.643227] [Medline: 22248010]
11. 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]
12. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: a review. Am J Alzheimers Dis Other Demen 2015 Feb;30(1):13-28. [doi: 10.1177/1533317514522852] [Medline: 24526761]
13. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov;4(6):428-437 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
14. Tuena C, Pedroli E, Trimarchi PD, Gallucci A, Chiappini M, Goulene K, et al. Usability Issues of Clinical and Research Applications of Virtual Reality in Older People: A Systematic Review. Front Hum Neurosci 2020;14:93 [FREE Full text] [doi: 10.3389/fnhum.2020.00093] [Medline: 32322194]
15. Wang S, Yin H, Li G, Jia Y, Leng M, Meng Q, et al. Detection of Mild Cognitive Impairment Based on Virtual Reality: A Scoping Review. Curr Alzheimer Res 2020;17(2):126-140. [doi: 10.2174/1567205017666200317100421] [Medline: 32183674]
16. Lecouvey G, Morand A, Gonneaud J, Piolino P, Orriols E, Pélerin A, et al. An Impairment of Prospective Memory in Mild Alzheimer's Disease: A Ride in a Virtual Town. Front Psychol 2019;10:241 [FREE Full text] [doi: 10.3389/fpsyg.2019.00241] [Medline: 30809174]
17. Giovannetti T, Yamaguchi T, Roll E, Harada T, Rycroft SS, Divers R, et al. The Virtual Kitchen Challenge: preliminary data from a novel virtual reality test of mild difficulties in everyday functioning. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2019 Nov;26(6):823-841. [doi: 10.1080/13825585.2018.1536774] [Medline: 30370825]
18. Rizzo A, Gambino G, Sardo P, Rizzo V. Being in the Past and Perform the Future in a Virtual World: VR Applications to Assess and Enhance Episodic and Prospective Memory in Normal and Pathological Aging. Front Hum Neurosci 2020;14:297 [FREE Full text] [doi: 10.3389/fnhum.2020.00297] [Medline: 32848672]
19. Negu A, Matu S, Sava FA, David D. Virtual reality measures in neuropsychological assessment: a meta-analytic review. Clin Neuropsychol 2016 Feb;30(2):165-184. [doi: 10.1080/13854046.2016.1144793] [Medline: 26923937]
20. Zygouris S, Iliadou P, Lazarou E, Giakoumis D, Votis K, Alexiadis A, et al. Detection of Mild Cognitive Impairment in an At-Risk Group of Older Adults: Can a Novel Self-Administered Serious Game-Based Screening Test Improve Diagnostic Accuracy? J Alzheimers Dis 2020;78(1):405-412 [FREE Full text] [doi: 10.3233/JAD-200880] [Medline: 32986676]
21. Zygouris S, Giakoumis D, Votis K, Doumpoulakis S, Ntovas K, Segkouli S, et al. Can a virtual reality cognitive training application fulfill a dual role? Using the virtual supermarket cognitive training application as a screening tool for mild cognitive impairment. J Alzheimers Dis 2015;44(4):1333-1347. [doi: 10.3233/JAD-141260] [Medline: 25428251]
22. Werner P, Rabinowitz S, Klinger E, Korczyn AD, Josman N. Use of the virtual action planning supermarket for the diagnosis of mild cognitive impairment: a preliminary study. Dement Geriatr Cogn Disord 2009;27(4):301-309. [doi: 10.1159/000204915] [Medline: 19252401]

XSL•FO
RenderX

23. Eraslan Boz H, Limoncu H, Zygouris S, Tsolaki M, Giakoumis D, Votis K, et al. A new tool to assess amnestic mild cognitive impairment in Turkish older adults: virtual supermarket (VSM). Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2020 Sep;27(5):639-653. [doi: 10.1080/13825585.2019.1663146] [Medline: 31482749]
24. Atkins AS, Khan A, Ulshen D, Vaughan A, Balentin D, Dickerson H, et al. Assessment of Instrumental Activities of Daily Living in Older Adults with Subjective Cognitive Decline Using the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). J Prev Alzheimers Dis 2018;5(4):216-234. [doi: 10.14283/jpad.2018.28] [Medline: 30298179]
25. 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]
26. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open 2016 Nov 14;6(11):e012799 [FREE Full text] [doi: 10.1136/bmjopen-2016-012799] [Medline: 28137831]
27. 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]
28. Arevalo-Rodriguez I, Smailagic N, Roqué I Figuls M, Ciapponi A, Sanchez-Perez E, Giannakou A, et al. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI). Cochrane Database Syst Rev 2015 Mar 05(3):CD010783 [FREE Full text] [doi: 10.1002/14651858.CD010783.pub2] [Medline: 25740785]
29. Graf C, Hartford Institute for Geriatric Nursing. The Lawton instrumental activities of daily living (IADL) scale. Medsurg Nurs 2008 Oct;17(5):343-344. [Medline: 19051984]
30. Cullum S, Tucker S, Todd C, Brayne C. Screening for depression in older medical inpatients. Int J Geriatr Psychiatry 2006 May;21(5):469-476. [doi: 10.1002/gps.1497] [Medline: 16676293]
31. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. J Biomed Inform 2014 Apr;48:193-204 [FREE Full text] [doi: 10.1016/j.jbi.2014.02.013] [Medline: 24582925]
32. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993 Nov;43(11):2412-2414. [doi: 10.1212/wnl.43.11.2412-a] [Medline: 8232972]
33. Lim WS, Chong MS, Sahadevan S. Utility of the clinical dementia rating in Asian populations. Clin Med Res 2007 Mar;5(1):61-70 [FREE Full text] [doi: 10.3121/cmr.2007.693] [Medline: 17456836]
34. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
35. Lim JE, Wong WT, Teh TA, Lim SH, Allen JC, Quah JHM, et al. A Fully-Immersive and Automated Virtual Reality System to Assess the Six Domains of Cognition: Protocol for a Feasibility Study. Front Aging Neurosci 2020;12:604670 [FREE Full text] [doi: 10.3389/fnagi.2020.604670] [Medline: 33488382]
36. Park SH, Goo JM, Jo C. Receiver operating characteristic (ROC) curve: practical review for radiologists. Korean J Radiol 2004;5(1):11-18 [FREE Full text] [doi: 10.3348/kjr.2004.5.1.11] [Medline: 15064554]
37. Chua SIL, Tan NC, Wong WT, Allen JC, Quah JHM, Malhotra R, et al. Virtual Reality for Screening of Cognitive Function in Older Persons: Comparative Study. J Med Internet Res 2019 Aug 01;21(8):e14821 [FREE Full text] [doi: 10.2196/14821] [Medline: 31373274]
38. Ben-Sadoun G, Sacco G, Manera V, Bourgeois J, König A, Foulon P, et al. Physical and Cognitive Stimulation Using an Exergame in Subjects with Normal Aging, Mild and Moderate Cognitive Impairment. J Alzheimers Dis 2016 Jun 30;53(4):1299-1314. [doi: 10.3233/JAD-160268] [Medline: 27372645]
39. Oliveira CR, Lopes Filho BJP, Esteves CS, Rossi T, Nunes DS, Lima MMBMP, et al. Neuropsychological Assessment of Older Adults With Virtual Reality: Association of Age, Schooling, and General Cognitive Status. Front Psychol 2018;9:1085 [FREE Full text] [doi: 10.3389/fpsyg.2018.01085] [Medline: 30008689]
40. Burgess PW, Veitch E, de Lacy Costello A, Shallice T. The cognitive and neuroanatomical correlates of multitasking. Neuropsychologia 2000;38(6):848-863. [doi: 10.1016/s0028-3932(99)00134-7] [Medline: 10689059]
41. Farina M, Paloski LH, de Oliveira CR, de Lima Argimon II, Irigaray TQ. Cognitive Reserve in Elderly and Its Connection with Cognitive Performance: A Systematic Review. Ageing Int 2017 May 23;43(4):496-507. [doi: 10.1007/s12126-017-9295-5]
42. Salthouse TA. Correlates of cognitive change. J Exp Psychol Gen 2014 Jun;143(3):1026-1048 [FREE Full text] [doi: 10.1037/a0034847] [Medline: 24219021]
43. Kang YJ, Ku J, Han K, Kim SI, Yu TW, Lee JH, et al. Development and clinical trial of virtual reality-based cognitive assessment in people with stroke: preliminary study. Cyberpsychol Behav 2008 Jun;11(3):329-339. [doi: 10.1089/cpb.2007.0116] [Medline: 18537503]
44. Mai LM, Sposato LA, Rothwell PM, Hachinski V, Pendlebury ST. A comparison between the MoCA and the MMSE visuoexecutive sub-tests in detecting abnormalities in TIA/stroke patients. Int J Stroke 2016 Jun;11(4):420-424. [doi: 10.1177/1747493016632238] [Medline: 26865154]

XSL•FO
RenderX

45. Chan J, Wong A, Yiu B, Mok H, Lam P, Kwan P, et al. Correction: Electronic Cognitive Screen Technology for Screening Older Adults With Dementia and Mild Cognitive Impairment in a Community Setting: Development and Validation Study. J Med Internet Res 2021 Jan 19;23(1):e26724 [FREE Full text] [doi: 10.2196/26724] [Medline: 33465038]
46. Yun SJ, Kang M, Yang D, Choi Y, Kim H, Oh B, et al. Cognitive Training Using Fully Immersive, Enriched Environment Virtual Reality for Patients With Mild Cognitive Impairment and Mild Dementia: Feasibility and Usability Study. JMIR Serious Games 2020 Oct 14;8(4):e18127 [FREE Full text] [doi: 10.2196/18127] [Medline: 33052115]

Abbreviations

ADL: Activities of Daily Living
CDR: Clinical Dementia Rating
GDS-15: 15-item Geriatric Depression Scale
ICC: intraclass correlation coefficient
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
ROC: receiver operating characteristic
AUC: Area under the receiver operating characteristic curve
VR: virtual reality
VSP: Virtual Supermarket Program

Edited by N Zary; submitted 04.06.21; peer-reviewed by HG Seo, NC Tan, S Zygouris; comments to author 22.07.21; revised version received 29.08.21; accepted 10.09.21; published 03.12.21.

Please cite as:
Yan M, Yin H, Meng Q, Wang S, Ding Y, Li G, Wang C, Chen L
A Virtual Supermarket Program for the Screening of Mild Cognitive Impairment in Older Adults: Diagnostic Accuracy Study
JMIR Serious Games 2021;9(4):e30919
URL: https://games.jmir.org/2021/4/e30919
doi:10.2196/30919
PMID:34870610

©Mingli Yan, Huiru Yin, Qiuyan Meng, Shuo Wang, Yiwen Ding, Guichen Li, Chunyan Wang, Li Chen. Originally published in JMIR Serious Games (https://games.jmir.org), 03.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effectiveness and Utility of Virtual Reality Simulation as an Educational Tool for Safe Performance of COVID-19 Diagnostics: Prospective, Randomized Pilot Trial

Tanja Birrenbach[1,2], MME, MD; Josua Zbinden[1]; George Papagiannakis[3,4,5], PhD; Aristomenis K Exadaktylos[1], MD, PhD; Martin Müller[1], MD, PhD; Wolf E Hautz[1], MME, MD; Thomas Christian Sauter[1], MME, MD

[1]Department of Emergency Medicine, Inselspital, University Hospital Bern, Bern, Switzerland
[2]Centre for Health Sciences Education, Faculty of Medicine, University of Oslo, Oslo, Norway
[3]ORamaVR SA, Geneva, Switzerland
[4]Institute of Computer Science, Foundation for Research and Technology, Hellas, Heraklion, Greece
[5]Department of Computer Science, University of Crete, Heraklion, Greece

Corresponding Author:
Tanja Birrenbach, MME, MD
Department of Emergency Medicine, Inselspital
University Hospital Bern
Freiburgstrasse 16C
Bern, 3010
Switzerland
Phone: 41 31 6322111
Email: tanja.birrenbach@insel.ch

Abstract

Background: Although the proper use of hygiene and personal protective equipment (PPE) is paramount for preventing the spread of diseases such as COVID-19, health care personnel have been shown to use incorrect techniques for donning/doffing of PPE and hand hygiene, leading to a large number of infections among health professionals. Education and training are difficult owing to the social distancing restrictions in place, shortages of PPE and testing material, and lack of evidence on optimal training. Virtual reality (VR) simulation can offer a multisensory, 3-D, fully immersive, and safe training opportunity that addresses these obstacles.

Objective: The aim of this study is to explore the short- and long-term effectiveness of a fully immersive VR simulation versus a traditional learning method regarding a COVID-19–related skill set and media-specific variables influencing training outcomes.

Methods: This was a prospective, randomized controlled pilot study on medical students (N=29; intervention VR training, n=15, vs control video-based instruction, n=14) to compare the performance of hand disinfection, nasopharyngeal swab taking, and donning/doffing of PPE before and after training and 1 month later as well as variables of media use.

Results: Both groups performed significantly better after training, with the effect sustained over one month. After training, the VR group performed significantly better in taking a nasopharyngeal swab, scoring a median of 14 out of 17 points (IQR 13-15) versus 12 out of 17 points (IQR 11-14) in the control group, *P*=.03. With good immersion and tolerability of the VR simulation, satisfaction was significantly higher in the VR group compared to the control group (median score of User Satisfaction Evaluation Questionnaire 27/30, IQR 23-28, vs 22/30, IQR 20-24, in the control group; *P*=.01).

Conclusions: VR simulation was at least as effective as traditional learning methods in training medical students while providing benefits regarding user satisfaction. These results add to the growing body of evidence that VR is a useful tool for acquiring simple and complex clinical skills.

(JMIR Serious Games 2021;9(4):e29586) doi:10.2196/29586

KEYWORDS

virtual reality; VR; simulation; medical education; hand hygiene; COVID-19; PPE; nasopharyngeal swab; protection; effectiveness; utility; diagnostic; testing; pilot study

XSL•FO
RenderX

Introduction

The COVID-19 pandemic is a global health emergency that places massive demands on health systems and health care workers [1]. Proper use of hygiene and personal protective equipment (PPE) is paramount to prevent spreading of disease and contamination of health care workers. One possible reason for the high infection rate of COVID-19 is ineffective use of PPE. In Italy, up to 20% of health care workers were initially infected with the disease [2].

PPE recommendations from international organizations are largely consistent (eg, those from the US Centers for Disease Control and Prevention [CDC] and World Health Organization [WHO]) [3-6]; however, the actual use of PPE is not. Health care personnel in all professions and at all levels of training have been shown to use incorrect techniques for donning and doffing of PPE and hand hygiene [7-11]. The main reason appears to be inadequate training in correct PPE technique and lack of assessment of proficiency [7,11,12].

Simulation proves to be a powerful tool to test the accurate use of hygiene skills that are relevant for the treatment of patients with COVID-19 (ie, PPE and hand hygiene) [13]; nevertheless, there is still ambiguity regarding which training method works best. A recent Cochrane review of evidence relating to PPE and protection of health care staff exposed to contaminated body fluids highlights the lack of robust evidence in this area [11].

Furthermore, education and training of health care personnel is difficult with social distancing restrictions in place and shortages of both PPE [2] and testing materials.

Virtual reality (VR) uses computer systems to generate realistic pseudoenvironments that provide users with visual, tactile, and auditory sensations, with the possibility of realistic interaction with the virtual environment [14]. Milgram and Kishino [15] referred to mixed reality (MR) as the technologies which involve the merging of real and virtual worlds.

VR simulation with the use of head-mounted devices (HMDs) can offer a multisensory, 3-D, fully immersive, and safe training opportunity, avoiding the restrictions of social distancing and material shortages [16,17]. Through the concept of immersion, sense of presence, and interaction with the virtual environment in a real-time and realistic manner, VR simulation can create emotional experiences that facilitate experiential learning, exceeding other 2-D learning modalities [18].

The value of VR in medical education has already been demonstrated for various tasks [16,19-30]. VR is often used for training skills of varying complexity, ranging from simple nursing skills (Foley catheter placement, gaining venous access [31]) to laparoscopic/endoscopic/endovascular skills [23,32] or complex surgical procedures [26-29,33,34]. It is suggested that skills training as a first step to acquiring a competency can be better taught with VR than with traditional learning methods, because VR allows for a more active and immersive learning experience [21].

Recent studies suggest that VR improves postintervention knowledge and skills of health professionals better than traditional education or other types of digital education [21,31,35]. VR offers several advantages, as it provides possibilities for flexible learning and self-learning, providing standardization, reproducibility, and stimuli control; it enables automated generation of data about the details of simulations, including performance measurements that can be used for research or to provide automated individualized feedback [24]. The simple novelty of interactive technologies themselves, such as VR, can improve student motivation [36]. The initial cost and effort of creating the program can easily be compensated by broad distribution [16,37], as VR training is gradually finding its way into the medical curriculum [38].

Only a few virtual or mixed reality simulations exist for training hand hygiene [39-42]. In addition, high-quality studies evaluating the effectiveness and long-term retention of a VR simulation compared to conventional training methods are lacking.

Therefore, we hypothesized that a VR simulation could be an effective and useful tool with high student satisfaction to teach COVID-19 diagnostics, and we performed a randomized pilot study in medical students to explore (1) the effectiveness of a fully immersive VR simulation versus a traditional learning method regarding a COVID-19–related skill set (ie, proper hand hygiene, proficiency in PPE use, and correct acquisition of a nasopharyngeal specimen) tested in a simulated clinical scenario, and, as a secondary outcome, its long-term effectiveness 1 month after training; and (2) media-specific variables influencing training outcomes, such as usability, satisfaction, simulator sickness, and the experience of presence and immersion.

Methods

Study Design, Setting, and Ethical Approval

This is a prospective randomized controlled pilot study, taking place at the emergency department of the Inselspital, University Hospital Bern, Switzerland [43], from September to November 2020.

The study population consisted of a convenience sample of medical students at the University of Bern. All participants attended on a voluntary basis; no remuneration was provided. Informed consent was obtained. Data were collected, analyzed, and stored in anonymized form.

The local ethics committee deemed our study exempt from full ethical approval (Business Administration System for Ethics Committees Req-2020-00889).

Inclusion/Exclusion

The inclusion criteria were as follows: medical students (years 3-6 out of a 6-year curriculum) at the University of Bern.

The exclusion criteria consisted of unwillingness to participate or to provide informed consent.

Baseline Investigations

Baseline Survey

A brief survey about sociodemographic factors; prior training; and experience in hand hygiene and PPE use, taking of

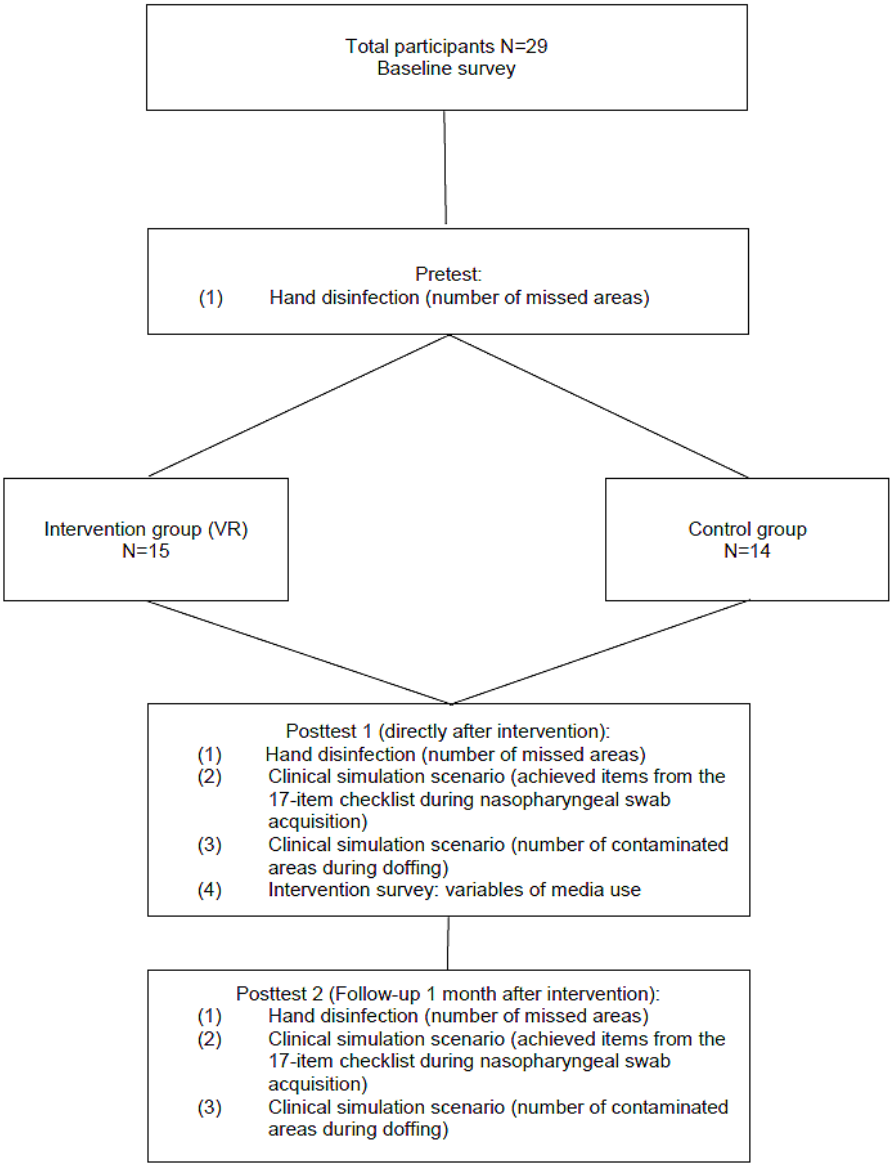
XSL•FO
RenderX

respiratory samples (nasopharyngeal swab), and prior experience with VR was performed after enrollment.

Assessments/Measurements

We evaluated the performance of hand disinfection, taking a nasopharyngeal swab on a manikin, and contamination during doffing of PPE (Figure 1).

Figure 1. Flowchart of the study. VR: virtual reality.



Hand Disinfection

Hand disinfection performance was evaluated using a fluorescent marker (Visirub conc and Sterillium, Hartmann AG) and UV-light scanning was performed using the Derma Litecheck UV Multimedia device (KBD Ltd) at the time of enrollment (pretest), directly after the intervention (posttest 1), and 1 month after the intervention (posttest 2). Participants were blindfolded during the assessment and were unable to assess their results.

A performance analysis scheme (documentation of missed locations; n=38 locations for each hand, with a total of 76 areas investigated) was developed by the institution's infection control and medical educator team as adapted from Pan et al [44]. The outcome was the number of missed locations (range from 0 to 76; a lower number is better). Performance was supervised by an independent and trained rater, and the images were electronically recorded and analyzed according to the predefined scheme by a rater blinded to the intervention (Figures 2 and 3).

XSL•FO
RenderX

Figure 2. UV-light scanning of a perfectly disinfected palmar surface of the right hand.

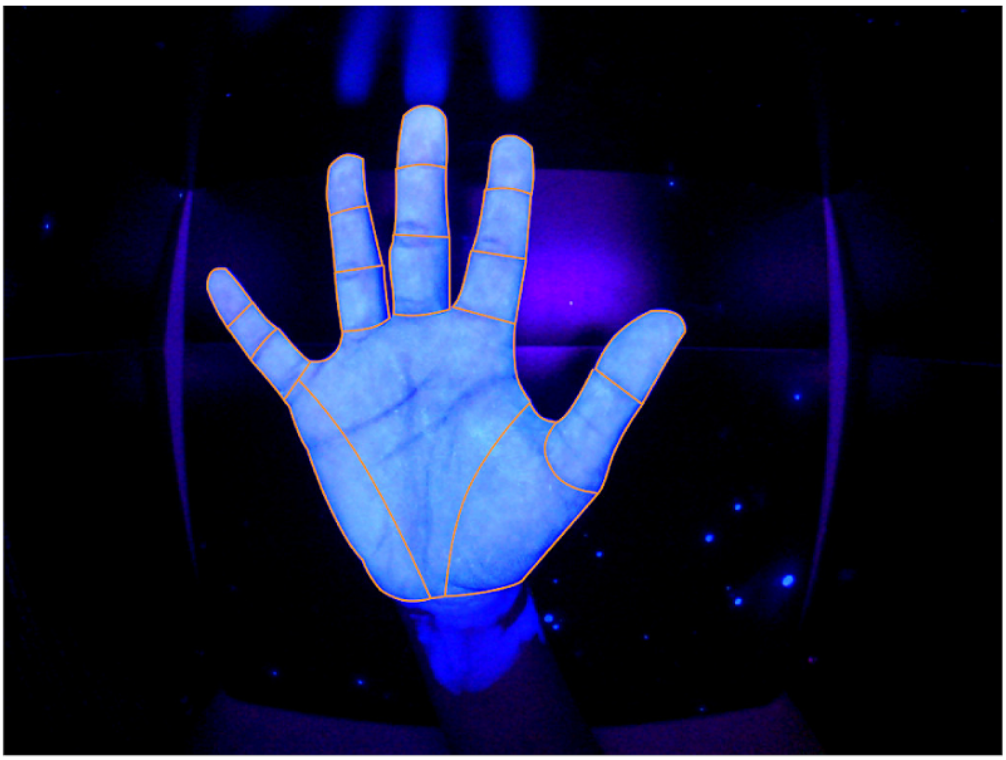
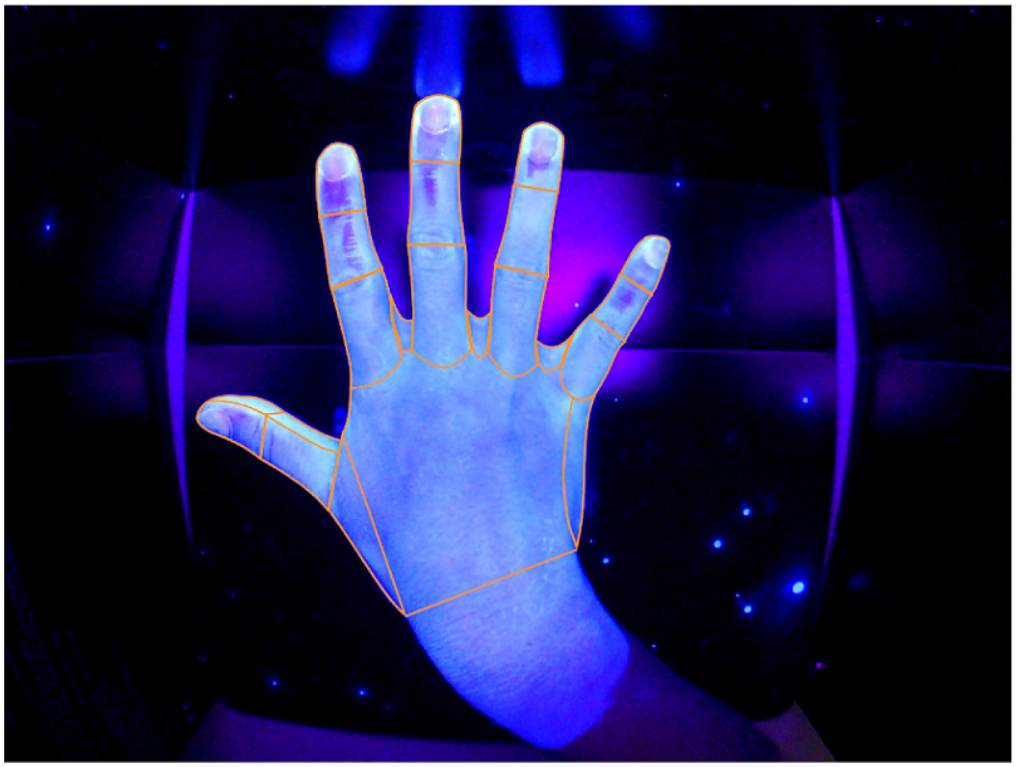


Figure 3. UV-light scanning of right hand dorsal surface. Missing areas of hand disinfection, right hand: digitus I, dorsal, distal phalanx; digitus II, dorsal, distal, and middle phalanx; digitus III, dorsal, distal, and middle phalanx; digitus IV, dorsal, distal phalanx; digitus V, dorsal, distal, and middle phalanx (total number of missed locations=8).



XSL•FO
RenderX

Obtaining the Nasopharyngeal Swab and Evaluating Contamination During Doffing

A simulation setup for conducting a nasopharyngeal swab for COVID-19 testing on a manikin (Little Anne, Laerdal Medical) using proper hand hygiene and PPE was installed.

The correct procedure of taking a nasopharyngeal swab sample as well as possible contamination while doffing were evaluated directly after the intervention (posttest 1) and 1 month after the intervention (posttest 2).

An independent and trained rater blinded to the intervention assessed each participant's performance using a 17-item checklist adapted from [8,10] based on the CDC guidelines for PPE [3,6], WHO guidelines for hand hygiene [5], and international recommendations for taking a nasopharyngeal swab [6]; this checklist was developed by the institution's infection control and medical educator team (Multimedia Appendix 1). The outcome was the number of points achieved on the checklist (range 0-17; a higher number of points indicated a better result).

Contamination during the procedure was evaluated using fluorescent lotion (Dermalux Testlotion S, KBD Ltd), which was applied to the participants' hands, forearms, and torso before the doffing of PPE. After doffing, 10 areas (right hand, right forearm, right upper arm, left hand, left forearm, left upper arm, torso ventral, torso dorsal, neck, head/ears) were analyzed by UV lighting for contamination by an independent rater. The outcome was the number of contaminated areas (range from 0 to 10; a lower number indicated a better result).

Intervention

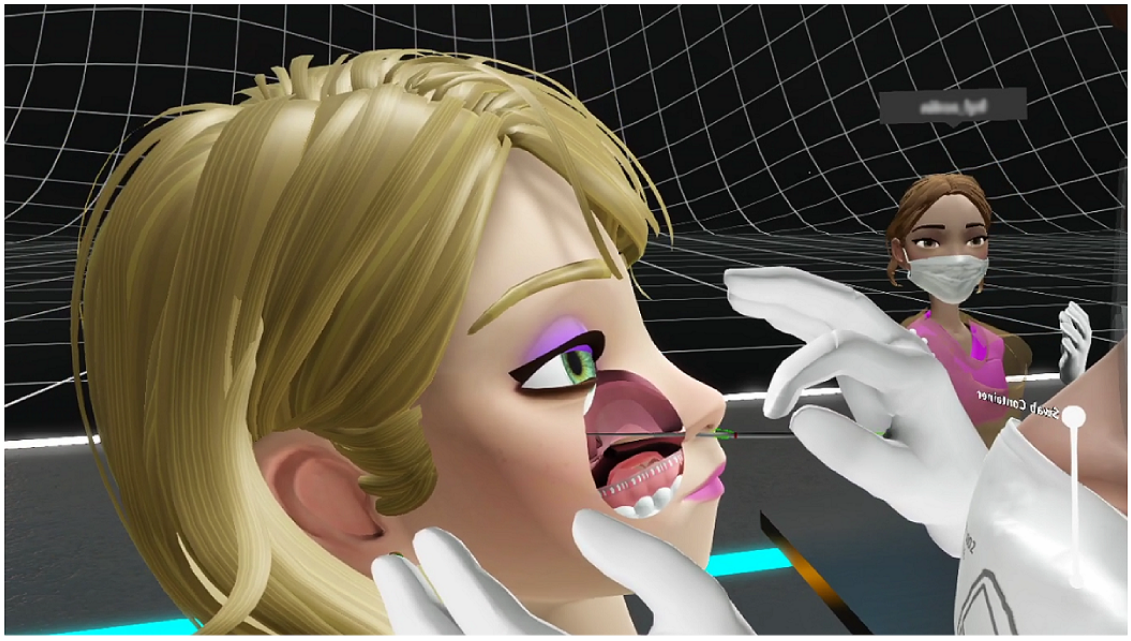
Participants were randomized to either the intervention group (VR simulation) or control group in a 1:1 ratio using a computer-generated system.

VR Simulation

The intervention group was trained in COVID-19–related skills using the VR simulation (the Covid-19 VR Strikes Back (CVRSB) module, version 1.1.6), a software platform developed by ORamaVR SA, and the Oculus Rift S head mounted device and hand controllers (Facebook Inc). The proprietary ORamaVR software medical VR training application is available free of charge for all VR desktop and mobile HMDs [45] (Figure 4).

The participants performed two runs in the simulation using the single player modus.

Figure 4. Screenshot of the virtual reality (VR) application, Covid-19 VR Strikes Back, showing the taking of a nasopharyngeal swab.



Control Group

The control group was trained using traditional learning methods: printed instructions and local instruction videos on COVID-19–related skills, ie, PPE donning and doffing, as well as formal videos on proper hand hygiene according to the WHO and on taking a correct nasopharyngeal sample [46].

Intervention Survey

Both groups were evaluated regarding variables of media use according to established questionnaires.

Usability for both training modules was assessed using the After-Scenario Questionnaire (ASQ) [47], which assesses the ease of task completion, satisfaction with completion time, and satisfaction with supporting information on a 7-point Likert scale (total score ranges from 1, full satisfaction, to 7, poor satisfaction).

The User Satisfaction Evaluation Questionnaire (USEQ) [48] contains 6 questions with a 5-point Likert scale to evaluate user satisfaction (total score ranges from 6, poor satisfaction, to 30, excellent satisfaction).

XSL•FO
RenderX

For the VR simulation, "visually induced motion sickness" was assessed with 4 items (nausea, headache, blurred vision, dizziness) according to the Simulator Sickness Questionnaire (SSQ) adapted from Kennedy et al [49] (total score ranges from 1, no simulator sickness, to 5, strong simulator sickness).

Presence and immersion in the virtual world was determined according to the 6-item questionnaire developed by Slater-Usoh-Steed [50] (total score ranges from 1, no immersion, to 7, full immersion).

Statistical Analysis

Data were analyzed using SPSS, version 22 (IBM Corporation), and Stata 16.1 (StataCorp).

Baseline characteristics are presented as numbers and percentages or medians and interquartile ranges using descriptive statistics as appropriate.

The intervention and control group were compared regarding the baseline characteristics by chi-square test and Wilcoxon rank sum test as applicable.

The Wilcoxon rank sum test was used at a specific time point between the study groups for the comparison of all four outcome groups: (1) the number of missed areas during hand disinfection, (2) achieved items from the 17-item checklist during nasopharyngeal swab acquisition, (3) the number of contaminated areas during doffing, and (4) variables of media use. Within-group differences for different time points were tested using the Wilcoxon matched-pairs signed-rank test.

For all tests, a P value <.05 was considered significant. For this pilot study, no adjustment for multiple testing was performed. Furthermore, pairwise comparisons were favored over more complex analyses, such as mixed linear regression analysis.

Results

Sample Characteristics

In total, 29 students completed the study (control group, n=14; intervention group, n=15) (Figure 1). All students included completed the whole study. There were no dropouts. The baseline characteristics of the participants are detailed in Table 1. No significant differences were found regarding gender, mean age, educational level in medical school, need to wear glasses, previous experience with computer games, or previous experience with VR. Likewise, previous education and experience regarding hand disinfection, use of PPE, and taking nasopharyngeal swabs did not show any significant differences.

XSL•FO
RenderX

Table 1. Baseline characteristics of the study sample (N=29).

	VR[a] group (n=15)	Control group (n=14)	*P* value
Sociodemographic factors			
Female gender, n (%)	9 (60)	9 (64)	.81
Age (years), median (IQR)	23 (22-25)	22.5 (22-24)	.56
Year of medical school, n (%)			.44
3	1 (7)	2 (14)	
4	14 (93)	11 (79)	
5	0 (0)	1 (7)	
Wears glasses, n (%)	8 (53)	8 (57)	.84
Plays computer games regularly (Likert scale response[b]), n (%)			.95
1	10 (67)	10 (71)	
2	4 (27)	3 (21)	
3	0 (0)	0 (0)	
4	0 (0)	0 (0)	
5	1 (7)	1 (7)	
Uses VR regularly (Likert scale response[b]), n (%)			
1	15 (100)	14 (100)	
Previous education and experience			
Previous education in hand disinfection, n (%)	14 (93)	11 (79)	.25
Previous education in PPE,[c] n (%)	6 (40)	3 (21)	.28
Previous no. of swabs, median (IQR)	0 (0-0)	0 (0-0)	.54
Uses PPE regularly (Likert scale response[b]), n (%)			.36
1	10 (67)	8 (57)	
2	3 (20)	2 (14)	
3	0 (0)	3 (21)	
4	1 (7)	0 (0)	
5	1 (7)	1 (7)	

[a]VR: virtual reality.

[b]Likert scale: 1, completely disagree, to 5, completely agree.

[c]PPE: personal protective equipment.

Hand Disinfection

There was no significant difference in the number of missed areas during hand disinfection at baseline (intervention group: median 21, IQR 11-27, vs control group: median 20, IQR 14-21; *P*=.47) (Table 2). Both groups performed significantly better after training without a significant group difference (posttest 1) (median 7, IQR 4-14, in the intervention group vs median 10, IQR 6-14, in the control group; *P*=.34). For the secondary outcome, at posttest 2, again, no significant difference was noted between the intervention and control groups (median 14, IQR 8-17, in the intervention group vs median 11, IQR 7-16, in the control group; *P*=.74). In both groups, no significant difference was found between posttests 1 and 2 (intervention group, *P*=.11; control group, *P*=.25) (Table 3).

XSL•FO
RenderX

Table 2. Comparison between the VR group and the control group regarding hand disinfection, nasopharyngeal swab testing, and contamination during doffing.

	Values, median (IQR)		*P* value
	VR[a] group (n=15)	Control group (n=14)	
Number of missing areas during hand disinfection (out of 76 possible areas)			
Pretest	21 (11-27)	20 (14-21)	.47
Posttest 1	7 (4-14)	10 (6-14)	.34
Posttest 2	14 (8-17)	11 (7-16)	.74
Nasopharyngeal swab test score (out of 17 points)			
Posttest 1	14 (13-15)	12 (11-14)	.03
Posttest 2	14 (14-16)	14 (14-15)	.79
Number of contaminated body areas during doffing (out of 10 possible areas)			
Posttest 1	2 (2-4)	3 (1-4)	.64
Posttest 2	1 (0-2)	0 (0-1)	.18

[a]VR: virtual reality.

Table 3. Comparison between posttests 1 and 2 for the VR group and the control group.

	Values, median (IQR)		*P* value
	Posttest 1	Posttest 2	
Hand disinfection (number of missing areas)			
VR[a] group	7 (4-14)	14 (8-17)	.11
Control group	10 (6-14)	11 (7-16)	.25
Swab test (score out of 17)			
VR group	14 (13-15)	14 (14-16)	.28
Control group	12 (11-14)	14 (14-15)	.02
Doffing (number of contaminated areas)			
VR group	2 (2-4)	1 (0-2)	.005
Control group	3 (1-4)	0 (0-1)	.003

[a]VR: virtual reality.

Nasopharyngeal Swab Acquisition

At posttest 1, the intervention group performed significantly better in taking a nasopharyngeal swab, scoring a median of 14 points on the 17-item checklist (IQR 13-15) versus 12 points (IQR 11-14) in the control group (*P*=.03) (Table 2). No significant differences between the groups were found after 1 month at posttest 2 (*P*=.79) for long-term retention as the secondary outcome.

The number of actual nasopharyngeal swabs performed in real life between posttests 1 and 2 did not differ between the groups (VR group, median no. of swabs 0, IQR 0-0; control group, median swabs 0, IQR 0-0; *P*=.56).

Contamination During Doffing

No significant difference between the number of contaminated areas during doffing was found between the groups at both time points (Table 2). However, in both groups, a significant reduction of contamination was noted at posttest 2 compared to posttest 1 (intervention group: posttest 1, median contaminated areas 2, IQR 2-4; posttest 2, median contaminated areas 1, IQR 0-2, *P*=.005; control group: posttest 1, median contaminated areas 3, IQR 1-4; posttest 2, median contaminated areas 0, IQR 0-1, *P*=.003).

Variables of Media Use

The results of the intervention survey regarding usability, satisfaction, simulator sickness, sense of presence, and immersion are detailed in Table 4.

XSL•FO
RenderX

Table 4. Variables related to media use.

	Values, median (IQR)		*P* value
	VR[a] group (n=15)	Control group (n=14)	
After-Scenario Questionnaire score[b]	1 (1-3)	3 (2-4)	.002
User Satisfaction Evaluation Questionnaire[c]	27 (23-28)	22 (20-24)	.01
Simulator Sickness Questionnaire, abbreviated[d]	1 (1-3)	N/A[e]	N/A
Presence and immersion according to Slater-Usoh-Steed[f]	5 (5-5)	N/A	N/A

[a]VR: virtual reality.

[b]Range 1 to 7 (1=full satisfaction).

[c]Range 6 to 30 (30=optimal satisfaction).

[d]Range 1 to 5 (1=no simulator sickness).

[e]N/A: not applicable.

[f]Range 1 to 7 (7=full presence and immersion).

The ASQ revealed a significantly better result for the VR module (median score in the intervention group: 1, IQR 1-3; median score in the control group: 3, IQR 2-4; *P*=.002), as did the USEQ (median score in the intervention group: 27/30, IQR 23-28; median score in the control group: 22/30, IQR 0-24; *P*=.01).

The median score in the 4-item SSQ in the intervention group was 1 (IQR 1-3), thus revealing good tolerability of the VR simulation.

Presence and immersion in the virtual world according to the questionnaire of Slater, Usoh, and Steed was high (median 5 on a scale from 1 to 7, IQR 5-5).

Discussion

In this study, we demonstrated that our VR simulation was at least as effective as traditional learning methods (video and written instruction) in training medical students in COVID-19–related skills—namely, the correct performance of hand hygiene, use of PPE, and taking of a nasopharyngeal swab specimen—and that it provides a benefit in user satisfaction.

Effectiveness of Training

For the most investigated steps of the training, both educational methods improved performance to a similar extent without a significant difference. However, students in the VR group performed significantly better in acquiring a nasopharyngeal swab specimen on a manikin directly after the intervention than the control group, but this finding may be a result of multiple testing. Furthermore, the medians of the two groups only differed by 2 points on a 17-point outcome scale, and the IQRs overlapped.

There are few VR and MR simulations for training hand hygiene, and high quality evidence of their effectiveness is limited [39,40,42]. Shimada and colleagues [41] targeted preschool children, using a data glove instead of hand-held controllers to obtain the posture of user's hand as a VR device. They used Leap Motion to obtain the posture of a hand, in contrast to our setting, which used only commercially available standard hardware. They found that their VR system was more effective than conventional hand hygiene instruction in a small group (n=12) of young children.

Performance of correct hand hygiene in our study was poor at baseline, in accordance with the existing evidence, stressing the need for effective instruction [7,11]. However, the performance significantly improved in both groups after the instruction, despite the limited technical possibilities to simulate the hand movements of disinfection with the need to hold the standard controllers in both hands while performing the correct movements. To minimize this limitation, participants could see the correct movements of their avatar's hands (without the controllers) in a mirror.

This study adds to the body of evidence that VR simulation can help with the acquisition of simple skills, in combination with increased user satisfaction. Furthermore, our study highlights the necessity of strong collaborations between developers, users, and educators to ensure that these new technologies can complement and enhance existing educational curricula.

Variables of Media Use

Satisfaction is considered to be one of the key components of usability [48]. The satisfaction of participants in the VR group measured by the USEQ was significantly higher than that of the participants in the control group. As most of our students were inexperienced in VR, the novelty effect may have added to the results. This effect consists of an increase in perceived usability of a technology due to its newness or the tendency for performance to initially improve when new technology is instituted, not because of any actual improvement in learning or achievement but in response to increased interest in the new technology. However, Huang and colleagues [51] state that novelty does not necessarily increase learning achievement. According to them, the increase of learning achievement is more dependent on a match between the learning content and the learning method. The embodied learning method in VR is particularly appropriate for instructing difficult knowledge and spatial knowledge [51]. It remains to be elucidated whether satisfaction with and efficiency of VR simulations will decrease over time as the technology becomes more widespread.

XSL•FO
RenderX

However, with potentially increasing technological advances, a certain novelty effect will remain.

To make our VR training available to as many users as possible, we avoided using specialized hardware such as that used in many other studies of VR skill training. It could be speculated that this use of off-the-shelf controllers may reduce the realism of the VR simulation. However, we were able to demonstrate that a high degree of immersion and satisfaction could be achieved with our simulation even with standard hardware. This may be even more pronounced with future developments, such as hand tracking without the need for traditional controllers.

This "experience of presence" in VR, which could be demonstrated in our study, is known to correlate positively with training effectiveness [52,53].

Skill Retention

In the follow-up after 1 month, there were no significant differences between the groups regarding any outcome. This cannot be explained by different exposures, as there was no difference in the mean number of swabs performed in real life in the meantime. One possible explanation is that the participants prepared more deliberately for the second appointment, as they might have suspected that it would involve a repetition of the first assessment ("assessment drives learning") [54]. Most motor skills are lost over time, or at least the level of performance deteriorates, starting as soon as 1 day after training [55]. Maagaard et al [56] detected that the laparoscopic skills of novices acquired in VR simulator training deteriorated in a period between 6 and 18 months without further training.

We were able to show that the observed learning effect was maintained over the observed time frame of 1 month in both learning groups. Whether there will be a difference in skill decay between the two learning methods in the long run remains an important open question.

Strengths and Limitations

Our study has several *strengths* to point out. First, the study assesses outcomes of direct clinical relevance, not surrogates such as performance on multiple-choice tests or user satisfaction only. Second, the study not only assesses the effect of training on performance gains but also includes a quantification of skill retention over time. Third, we compared the novel VR intervention to an established educational alternative rather than to no intervention.

In addition, attrition bias was nonexistent because all participants completed the study protocol without dropping out.

As a practical benefit, our VR simulation program is available for free; thus, program directors and educators are able to enhance their existing curricula with an effective novel adjunct or alternative, or replicate the study setting.

This study has several *limitations*, including its single center design, which restricts external validity. The number of participants in our study was limited due to the large logistical and human resources required to conduct the study during a pandemic. Therefore, the detection of small differences between training modalities is not possible with our study design. There is the possibility of selection bias, based on volunteer convenience sampling of medical students, as well as a possible performance bias, with allocation to the interventional group leading to higher motivation, satisfaction, and performance. The need to use hand-held controllers instead of hand-tracking might have further impacted the efficiency of the VR simulation; however, we wanted to apply technological equipment that is widely available.

Furthermore, the correlation of these findings to clinical, patient-oriented outcomes remains to be validated.

Conclusion

To our knowledge, this is the first study using a VR simulation to train health care personnel in the correct use of hand hygiene, PPE, and taking a nasopharyngeal swab specimen and to compare the effectiveness to established traditional training (video and written instructions). VR simulation was at least as effective as traditional learning methods in training medical students while providing a benefit in user satisfaction. These results add to the growing body of evidence that VR is a useful tool for acquiring and maintaining simple and complex clinical skills.

Conflicts of Interest

WEH has received research funding from the European Union, the Swiss National Science Foundation, the Zoll Foundation, Dräger Medical Germany, Mundipharma Research UK, MDI International Australia, and Roche Diagnostics Germany, all outside the submitted work. WEH has provided paid consultancies to AO Foundation Switzerland, outside the submitted work. WEH received financial support for a congress he chaired from EBSCO Germany, Isabel Healthcare UK, Mundipharma Medical Switzerland, and VisualDx USA, all outside the submitted work. TCS holds an endowed professorship supported by the Touring Club Switzerland. The sponsor has no influence on the research conducted, in particular on the results or the decision to publish. GP is Chief Technical Officer of ORamaVR, the company that developed the software for this study. The company had no influence on the study protocol, the analysis, the results, or the decision to publish. This study was funded in part by a grant from the Swiss National Science foundation (grant number 31CA30_196615) for the project "Mixed-method evaluation of an online forward triage tool within the COVID-19 pandemic" to TCS and WEH. The other authors report no conflict of interest.

Editorial note: This randomized study was not registered, explained by authors as follows: "As our trial was not a clinical trial and did not involve patients or patient data, but instead was a pilot study involving medical students, it was not registered as a clinical trial." The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, given

XSL•FO
RenderX

that the subjects were medical students. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness.

Multimedia Appendix 1
Checklist used for performance assessment.
[DOCX File , 13 KB - games_v9i4e29586_app1.docx]

Multimedia Appendix 2
CONSORT-EHEALTH checklist (V. 1.6.1).
[PDF File (Adobe PDF File), 693 KB - games_v9i4e29586_app2.pdf]

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 12;92(4):401-402 [FREE Full text] [doi: 10.1002/jmv.25678] [Medline: 31950516]
2. The Lancet. COVID-19: protecting health-care workers. Lancet 2020 Mar;395(10228):922. [doi: 10.1016/s0140-6736(20)30644-9]
3. Using personal protective equipment (PPE). US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/coronavirus/2019-ncov/hcp/using-ppe.html [accessed 2020-05-11]
4. Infection prevention and control during health care when novel coronavirus (nCoV) infection is suspected. World Health Organization. URL: https://www.who.int/publications-detail/infection-prevention-and-control-during-health-care-when-novel-coronavirus-(ncov)-infection-is-suspected-20200125 [accessed 2020-05-12]
5. WHO guidelines on hand hygiene in health care. World Health Organization. 2009 Jan 15. URL: https://www.who.int/gpsc/5may/tools/9789241597906/en/ [accessed 2020-05-12]
6. Protecting healthcare personnel. US Centers for Disease Control and Prevention. URL: https://www.cdc.gov/hai/prevent/ppe.html [accessed 2020-05-12]
7. John A, Tomas ME, Hari A, Wilson BM, Donskey CJ. Do medical students receive training in correct use of personal protective equipment? Med Educ Online 2017 Jan 04;22(1):1264125 [FREE Full text] [doi: 10.1080/10872981.2017.1264125] [Medline: 28178912]
8. Nagoshi Y, Cooper LA, Meyer L, Cherabuddi K, Close J, Dow J, et al. Application of an objective structured clinical examination to evaluate and monitor intern's proficiency of hand hygiene and personal protective equipment use in the United States. J Educ Eval Health Prof 2019 Oct 15;16:31 [FREE Full text] [doi: 10.3352/jeehp.2019.16.31] [Medline: 31614408]
9. Hung P, Choi K, Chiang V. Using interactive computer simulation for teaching the proper use of personal protective equipment. Comput Inform Nurs 2015 Feb;33(2):57. [doi: 10.1097/cin.0000000000000125]
10. Beam EL, Gibbs SG, Boulter KC, Beckerdite ME, Smith PW. A method for evaluating health care workers' personal protective equipment technique. Am J Infect Control 2011 Jun;39(5):415-420 [FREE Full text] [doi: 10.1016/j.ajic.2010.07.009] [Medline: 21255874]
11. Verbeek J, Rajamaki B, Ijaz S, Sauni R, Toomey E, Blackwood B. Personal protective equipment for preventing highly infectious diseases due to exposure to contaminated body fluids in healthcare staff. Cochrane Database Syst Rev 2020;4:1-150. [doi: 10.1002/14651858.cd011621.pub4]
12. John A, Tomas ME, Cadnum JL, Mana TS, Jencson A, Shaikh A, et al. Are health care personnel trained in correct use of personal protective equipment? Am J Infect Control 2016 Jul 01;44(7):840-842. [doi: 10.1016/j.ajic.2016.03.031] [Medline: 27181222]
13. Lockhart SL, Naidu JJ, Badh CS, Duggan LV. Simulation as a tool for assessing and evolving your current personal protective equipment: lessons learned during the coronavirus disease (COVID-19) pandemic. Can J Anaesth 2020 Jul 27;67(7):895-896 [FREE Full text] [doi: 10.1007/s12630-020-01638-z] [Medline: 32221852]
14. Pensieri C, Pennacchini M. Overview: virtual reality in medicine. J Virtual Worlds Res 2014 Jan 19;7(1):1-34 [FREE Full text] [doi: 10.4101/jvwr.v7i1.6364]
15. Milgram P, Kishino F. A taxonomy of mixed reality visual displays. IEICE Trans Inf Syst 1994 Dec 24;E77-D(12):9 [FREE Full text]
16. Li L, Yu F, Shi D, Shi J, Tian Z, Yang J, et al. Application of virtual reality technology in clinical medicine. Am J Transl Res 2017;9(9):3867-3880 [FREE Full text] [Medline: 28979666]
17. Slater M, Sanchez-Vives M. Enhancing our lives with immersive virtual reality. Front Robot AI 2016 Dec 19;3 [FREE Full text] [doi: 10.3389/frobt.2016.00074]
18. Babini MH, Kulish VV, Namazi H. Physiological state and learning ability of students in normal and virtual reality conditions: complexity-based analysis. J Med Internet Res 2020 Jun 01;22(6):e17945 [FREE Full text] [doi: 10.2196/17945] [Medline: 32478661]

XSL•FO
RenderX

19. Gerup J, Soerensen CB, Dieckmann P. Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review. Int J Med Educ 2020 Jan 18;11:1-18 [FREE Full text] [doi: 10.5116/ijme.5e01.eb1a] [Medline: 31955150]
20. Bin KM, Zary N. Augmented reality, virtual reality and mixed reality in medical education: a comparative Web of Science scoping review. Preprints.org. Preprint posted online on April 29, 2019. [FREE Full text] [doi: 10.20944/preprints201904.0323.v1]
21. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
22. Bracq M, Michinov E, Jannin P. Virtual reality simulation in nontechnical skills training for healthcare professionals. Sim Healthcare 2019 Jan 2;14(3):188-194. [doi: 10.1097/sih.0000000000000347]
23. Pantelidis P, Chorti A, Papagiouvanni I, Paparoidamis G, Drosos C, Panagiotakopoulos T. Virtual and augmented reality in medical education. In: Medical and Surgical Education - Past, Present and Future. London, UK: IntechOpen; 2017:77-97.
24. Ustun AB, Yilmaz R, Yilmaz FGK. Virtual reality in medical education. In: Mobile Devices and Smart Gadgets in Medical Sciences. Hershey, PA: IGI Global; 2020:73.
25. Semeraro F, Scapigliati A, Ristagno G, Luciani A, Gandolfi S, Lockey A, et al. Virtual reality for CPR training: how cool is that? Dedicated to the "next generation". Resuscitation 2017 Dec;121:e1-e2. [doi: 10.1016/j.resuscitation.2017.09.024] [Medline: 28951295]
26. Clarke E. Virtual reality simulation-the future of orthopaedic training? A systematic review and narrative analysis. Adv Simul (Lond) 2021 Jan 13;6(1):2 [FREE Full text] [doi: 10.1186/s41077-020-00153-x] [Medline: 33441190]
27. Lee R, Raison N, Lau WY, Aydin A, Dasgupta P, Ahmed K, et al. A systematic review of simulation-based training tools for technical and non-technical skills in ophthalmology. Eye (Lond) 2020 Oct 13;34(10):1737-1759 [FREE Full text] [doi: 10.1038/s41433-020-0832-1] [Medline: 32203241]
28. Lohre R, Bois AJ, Pollock JW, Lapner P, McIlquham K, Athwal GS, et al. Effectiveness of immersive virtual reality on orthopedic surgical skills and knowledge acquisition among senior surgical residents: a randomized clinical trial. JAMA Netw Open 2020 Dec 01;3(12):e2031217 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.31217] [Medline: 33369660]
29. Polce EM, Kunze KN, Williams BT, Krivicich LM, Maheshwer B, Beletsky A, et al. Efficacy and validity of orthopaedic simulators in surgical training: a systematic review and meta-analysis of randomized controlled trials. J Am Acad Orthop Surg 2020 May 8;28(24):1027-1040. [doi: 10.5435/jaaos-d-19-00839]
30. Hassandra M, Galanis E, Hatzigeorgiadis A, Goudas M, Mouzakidis C, Karathanasi EM, et al. A virtual reality app for physical and cognitive training of older people with mild cognitive impairment: mixed methods feasibility study. JMIR Serious Games 2021 Mar 24;9(1):e24170 [FREE Full text] [doi: 10.2196/24170] [Medline: 33759797]
31. Rourke S. How does virtual reality simulation compare to simulated practice in the acquisition of clinical psychomotor skills for pre-registration student nurses? A systematic review. Int J Nurs Stud 2020 Feb;102:103466. [doi: 10.1016/j.ijnurstu.2019.103466] [Medline: 31783192]
32. Khan R, Plahouras J, Johnston BC, Scaffidi MA, Grover SC, Walsh CM. Virtual reality simulation training in endoscopy: a Cochrane review and meta-analysis. Endoscopy 2019 Jul 09;51(7):653-664. [doi: 10.1055/a-0894-4400] [Medline: 31071757]
33. Hooper J, Tsiridis E, Feng JE, Schwarzkopf R, Waren D, Long WJ, NYU Virtual Reality Consortium. Virtual reality simulation facilitates resident training in total hip arthroplasty: a randomized controlled trial. J Arthroplasty 2019 Oct;34(10):2278-2283. [doi: 10.1016/j.arth.2019.04.002] [Medline: 31056442]
34. Grantcharov TP. Is virtual reality simulation an effective training method in surgery? Nat Clin Pract Gastroenterol Hepatol 2008 May 1;5(5):232-233. [doi: 10.1038/ncpgasthep1101] [Medline: 18382434]
35. Zhao G, Fan M, Yuan Y, Zhao F, Huang H. The comparison of teaching efficiency between virtual reality and traditional education in medical education: a systematic review and meta-analysis. Ann Transl Med 2021 Feb;9(3):252-252 [FREE Full text] [doi: 10.21037/atm-20-2785] [Medline: 33708879]
36. Kavanagh S, Luxton-Reilly A, Wuensche B, Plimmer B. A systematic review of virtual reality in education. Themes Sci Technol Educ 2017;10(2):119 [FREE Full text]
37. King D, Tee S, Falconer L, Angell C, Holley D, Mills A. Virtual health education: scaling practice to transform student learning: using virtual reality learning environments in healthcare education to bridge the theory/practice gap and improve patient safety. Nurse Educ Today 2018 Dec;71:7-9. [doi: 10.1016/j.nedt.2018.08.002] [Medline: 30205259]
38. Hautz S, Hoffmann M, Exadaktylos A, Hautz W, Sauter T. Digital competencies in medical education in Switzerland: an overview of the current situation. GMS J Med Educ 2020;37(6):Doc62 [FREE Full text] [doi: 10.3205/zma001355] [Medline: 33225054]
39. Gill B, Tsang Y. Technological innovation in nursing education and practice. The Open University of Hong Kong. 2019. URL: http://www.ouhk.edu.hk/NHS/Conference&Symposium/STINE/eBook/Technological%20Innovation%20in%20Nursing%20Education%20and%20Practice.pdf [accessed 2021-09-28]
40. Clack L, Hirt C, Wenger M, Saleschus D, Kunz A, Sax H. VIRTUE - a virtual reality trainer for hand hygiene. 2018 Presented at: 9th International Conference on Information, Intelligence, Systems and Applications (IISA); July 23-25, 2018; Zakynthos, Greece. [doi: 10.1109/iisa.2018.8633588]

XSL•FO
RenderX

41. Shimada S, Funahashi K, Ito K, Tanase Y, Iwazaki K. VR hand hygiene training system that visualizes germs to be washed and removed. 2017 Presented at: InIEEE 6th Global Conference on Consumer Electronics (GCCE); October 24-27, 2017; Nagoya, Japan. [doi: 10.1109/gcce.2017.8229333]
42. Kanazawa A, Asai T, Minazuki A, Hayashi H. Construction of a MR space using projection mapping to promote hand hygiene. 2014 Presented at: InIIAI 3rd International Conference on Advanced Applied Informatics; Kokura Kita-ku, Japan; Aug 31-Sept 4, 2014. [doi: 10.1109/iiai-aai.2014.59]
43. Exadaktylos A, Hautz WE. Emergency medicine in Switzerland. ICU Manag Pract 2015;15:3 [FREE Full text]
44. Pan S, Chen E, Tien K, Hung I, Sheng W, Chen Y, et al. Assessing the thoroughness of hand hygiene: "seeing is believing". Am J Infect Control 2014 Jul;42(7):799-801. [doi: 10.1016/j.ajic.2014.03.003] [Medline: 24767574]
45. Project Covid-19 VR Strikes Back. ORamaVR. URL: http://elearn.oramavr.com/cvrsb/ [accessed 2021-04-13]
46. Marty FM, Chen K, Verrill KA. How to obtain a nasopharyngeal swab specimen. N Engl J Med 2020 May 28;382(22):e76. [doi: 10.1056/nejmvcm2010260]
47. Lewis JR. Psychometric evaluation of an after-scenario questionnaire for computer usability studies. SIGCHI Bull 1991 Jan;23(1):78-81. [doi: 10.1145/122672.122692]
48. Gil-Gómez JA, Manzano-Hernández P, Albiol-Pérez S, Aula-Valero C, Gil-Gómez H, Lozano-Quilis J. USEQ: a short questionnaire for satisfaction evaluation of virtual rehabilitation systems. Sensors (Basel) 2017 Jul 07;17(7):1589 [FREE Full text] [doi: 10.3390/s17071589] [Medline: 28686174]
49. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]
50. Usoh M, Catena E, Arman S, Slater M. Using presence questionnaires in reality. Presence: Teleoperators Virtual Environ 2000 Oct;9(5):497-503. [doi: 10.1162/105474600566989]
51. Huang W. Investigating the Novelty Effect in Virtual Reality on STEM Learning. Thesis. Arizona State University. 2020. URL: http://hdl.handle.net/2286/R.I.57391 [accessed 2021-06-02]
52. Lerner D, Mohr S, Schild J, Göring M, Luiz T. An immersive multi-user virtual reality for emergency simulation training: usability study. JMIR Serious Games 2020 Jul 31;8(3):e18822 [FREE Full text] [doi: 10.2196/18822] [Medline: 32735548]
53. Mantovani F, Castelnuovo G. The sense of presence in virtual training: enhancing skills acquisition and transfer of knowledge through learning experience in virtual environments. In: Riva G, Davide F, Ijsselsteijn WA, editors. Being There: Concepts, Effects and Measurement of User Presence in Synthetic Environments. Amsterdam, the Netherlands: IOS Press; 2003:167-181.
54. McLachlan JC. The relationship between assessment and learning. Med Educ 2006 Aug;40(8):716-717. [doi: 10.1111/j.1365-2929.2006.02518.x] [Medline: 16869912]
55. Arthur Jr W, Bennett Jr W, Stanush PL, McNelly TL. Factors that influence skill decay and retention: a quantitative review and analysis. Hum Perform 1998 Mar;11(1):57-101. [doi: 10.1207/s15327043hup1101_3]
56. Maagaard M, Sorensen JL, Oestergaard J, Dalsgaard T, Grantcharov TP, Ottesen BS, et al. Retention of laparoscopic procedural skills acquired on a virtual-reality surgical trainer. Surg Endosc 2011 Mar 7;25(3):722-727. [doi: 10.1007/s00464-010-1233-5] [Medline: 20927549]

Abbreviations

ASQ: After-Scenario Questionnaire
CDC: US Centers for Disease Control
CVRSB: Covid-19 VR Strikes Back
HMD: head-mounted device
MR: mixed reality
PPE: personal protective equipment
SSQ: Simulator-Sickness Questionnaire
USEQ: User Satisfaction Evaluation Questionnaire
VR: virtual reality
WHO: World Health Organization

XSL•FO
RenderX

Edited by N Zary; submitted 26.04.21; peer-reviewed by R Lundin, S Gallagher; comments to author 26.05.21; revised version received 12.06.21; accepted 09.09.21; published 08.10.21.

Please cite as:
Birrenbach T, Zbinden J, Papagiannakis G, Exadaktylos AK, Müller M, Hautz WE, Sauter TC
Effectiveness and Utility of Virtual Reality Simulation as an Educational Tool for Safe Performance of COVID-19 Diagnostics: Prospective, Randomized Pilot Trial
JMIR Serious Games 2021;9(4):e29586
URL: https://games.jmir.org/2021/4/e29586
doi:10.2196/29586
PMID:34623315

©Tanja Birrenbach, Josua Zbinden, George Papagiannakis, Aristomenis K Exadaktylos, Martin Müller, Wolf E Hautz, Thomas Christian Sauter. Originally published in JMIR Serious Games (https://games.jmir.org), 08.10.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Software Engineering Frameworks Used for Serious Games Development in Physical Rehabilitation: Systematic Review

Jorge Fernando Ambros-Antemate[1*], MSc; María Del Pilar Beristain-Colorado[1*], MSc; Marciano Vargas-Treviño[2*], PhD; Jaime Gutiérrez-Gutiérrez[2*], PhD; Pedro Antonio Hernández-Cruz[3*], PhD; Itandehui Belem Gallegos-Velasco[3*], PhD; Adriana Moreno-Rodríguez[4*], 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]Laboratorio de genómica y proteómica, Centro de investigación UNAM-UABJO, Facultad de Medicina y Cirugía UABJO, Oaxaca de Juárez, Mexico
[4]Facultad de Ciencias Químicas, Universidad Autónoma "Benito Juárez" de Oaxaca, Oaxaca de Juárez, Mexico
[*]all authors contributed equally

Corresponding Author:
Jorge Fernando Ambros-Antemate, 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: jambros.cat@uabjo.mx

Abstract

Background: Serious games are a support in the rehabilitation process for treating people with physical disabilities. However, many of these serious games are not adapted to the patient's needs because they are not developed with a software engineering framework with a set of activities, actions, and tasks that must be executed when creating a software product. Better serious games for rehabilitation will be developed if the patient and therapist requirements are identified, the development is planned, and system improvements and feedback are involved. The goal is that the serious game must offer a more attractive environment, while maintaining patient interest in the rehabilitation process.

Objective: This paper submits the results of a systematic review of serious games in physical rehabilitation identifying the benefits of using a software engineering framework.

Methods: A systematic research was conducted using PubMed, PEDro (Physiotherapy Evidence Database), IEEE Xplore, ScienceDirect, ACM Digital Library, Mary Ann Liebert, Taylor & Francis Online, Wiley Online Library, and Springer databases. The initial search resulted in 701 papers. After assessing the results according to the inclusion criteria, 83 papers were selected for this study.

Results: From the 83 papers reviewed, 8 used a software engineering framework for its development. Most of them focused their efforts on 1 or more aspects, such as data acquisition and processing, game levels, motivation, therapist supervision.

Conclusions: This systematic review proves that most of the serious games do not use a software engineering framework for their development. As a result, development systems overlook several aspects and do not have a standardized process, eventually omitting important implementation aspects, which impact the patient's recovery time.

(JMIR Serious Games 2021;9(4):e25831) doi:10.2196/25831

KEYWORDS

serious game; physical rehabilitation; framework; methodology

XSL•FO
RenderX

Introduction

Overview

According to the World Health Organization, over 1 billion people have some form of disability [1], with up to 200 million people having loss or decrease in movement, which limits their ability to perform activities of daily living. To overcome it, they must undergo a rehabilitation program to gradually regain movement and consequently, improve their quality of life.

However, the traditional rehabilitation process is often slow and presents problems such as lack of motivation, boredom, and others; as a result, many patients consider the exercises stressful, and therefore abandon the therapy [2].

To avoid these situations, new ways of conventional therapy support have been used in recent years, such as medicinal treatments, robotics, video games (known as serious games), and others [3], which have contributed to faster rehabilitation when performing exercises in a fun way, allowing the patients to forget their conditions and concentrate on the game.

For this reason, new interaction modes, such as serious games [4], have the potential to provide more attractive, motivating, and enriching experiences for patients who suffer from decreases in movement. Currently, serious game–based physical rehabilitation is an area of research in constant evolution, and therefore, there is the need for developing guidelines adapted from other research fields.

Despite the potential benefits of serious games in physical rehabilitation, many available platforms are inflexible and limited in their scope. Many developments do not follow a process involving a set of activities, actions, or tasks that must be executed when a software product is to be created. As a result, essential elements to the patient's improvement process are ignored within the video game. Some of these elements are motivation, play levels, player commitment, challenges according to the patient's level, clinical evaluation, assessment scales, among others [5,6].

This work aims to describe the software engineering frameworks used in serious games development and their benefits in the physical rehabilitation process.

Background

A Note on Frameworks

The term *framework* has several meanings depending on the field. For example, it may refer to a model, prescription, guidelines underlying a design and analysis, among others.

The concept of framework is widely used in the field of computer science. However, there is some confusion between the software engineering framework and the application framework. The former provides a skeletal abstraction of a solution to several problems that have some similarities. A software engineering framework will generally outline the steps or phases that must be followed in implementing a solution without getting into the details of what activities are done in each phase [7]. The goal is for developers to use the framework as a guide to creating software systems by applying "building blocks" depending on the problem domain; by contrast, application framework is an integrated set of software artifacts (such as classes, objects, and components) that collaborate to provide a reusable architecture for a family of related applications [8]. They are used to facilitate the development process of applications, reducing time, effort, and costs.

Software engineering framework and application framework should not be confused. The latter is composed of pre-established source codes (eg, data access routines, form validation, templates) that the programmer uses to reduce workload and do not start the project from scratch.

One of the main motivations for applying a software engineering framework in serious game development is to design an efficient and satisfactory system for the patient.

Software Engineering Frameworks and Serious Games

The use of software engineering frameworks for the development of serious games allows the application of a variety of concepts, models, techniques, and artifacts at a high level of abstraction. Being an interdisciplinary field, an orientation on the developed tasks is required. Besides, it is flexible to adapt to changing conditions or personalization according to the final approach of the video game (rehabilitation, education, etc.).

Serious games like other software developments require a "systematic, disciplined, and quantifiable" approach. Every aspect of production, from early stages of system specification to maintenance after its operation, must be established. Below is a set of related activities that lead to the development of a software product [9-12].

Structural Activities in Software Development

In software engineering, 5 generic structural activities are used during software development [9-12]: communication, planning, modeling, construction, and deployment. The software process details will be different in each case, but the structural activities are the same. The definitions of the structural activities are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. Definitions of the structural activities in software development.

Communication

Defining the software characteristics and functions is particularly important to communicate and collaborate with the client and other participants. This activity aims to understand the project objectives of the participants and meet the requirements.

Planning

Once the requirements are obtained, this activity presents an estimate of the resources; establishes a software project plan; and describes technical tasks, probable risks, and program activities.

Modeling

Its objective is to help understand the requirements through models. The models' aim is to affirm the understanding of the work and give technical guidance to those who will implement the software, establishing, for example, the database model, the software architecture, user screen prototypes, and others. In some developments, this activity is the equivalent of the design stage.

Construction

This activity consists of the code generation and tests required to discover bugs in the software product.

Deployment

Once the software is created (completely or an increment), it is delivered to the client who will evaluate it and give feedback for system improvement.

Gamification

According to Kumar [13] gamification applies game design principles and mechanics to nongame environments. In the rehabilitation process, gamification can increase motivation and engagement through rewards, game levels, accessibility, feedback, and challenge. Therefore, the software engineering framework for serious game development must incorporate gamification. Various gamification elements include immersion, support for different roles, flow enhancement, visual enhancement, support for different learning stages and experience levels, design for interactivity, and progress [14]. By contrast, Vermeir et al [15] identified the following elements: avatar, challenge, competition, difficulty adjustment, feedback loops, levels, progress, rewards, social interaction, sound effects, and story/theme.

Benefits of Gamification in Rehabilitation

de Castro-Cros et al [16] analyzed the effects of gamification on the mental imagery brain–computer interface in rehabilitation functional assessments in 10 patients with stroke with hemiparesis in the upper limb and 6 healthy individuals. The authors concluded that user opinions about the game level of entertainment, clarity of rules, narrative, and visual attractiveness were all positive. The patients were consensus about the interest in gamifying stroke rehabilitation sessions. By contrast, Steiner et al [17] performed a scoping review of gamification in the rehabilitation of patients with musculoskeletal disorders of the shoulder. They concluded that gamification is essential in health care to enhance motivation and support therapy in general, especially in chronic diseases and rehabilitation. Other advantages are motivation, avoiding boredom, and distraction from pain and anxiety.

Related Works

A systematic review of literature is a method to identify, evaluate, and interpret all available and relevant research of a particular research question, subject area, or phenomenon of interest. The individual studies that contribute to the systematic review are called primary studies. A systematic review is also considered a form of secondary study [18].

This systematic review includes literature work on developing serious games in physical rehabilitation using a software engineering framework. To identify existing secondary research in the same field, we searched the following electronic databases: IEEE Xplore, ACM Digital Library, Wiley Digital Library, PubMed, ScienceDirect, Taylor & Francis, Mary Ann Liebert, and Springer. Besides, we used Google Scholar as a web source to broaden our results.

The search was realized using the following search string: A1 AND B1 AND (C1 OR C2 OR C3 OR C4 OR C5 OR C6). Textbox 2 shows the terms included in the search string.

XSL•FO
RenderX

Textbox 2. Search terms to identify related secondary studies.

A term

- A1. Serious games

B term

- B1. Framework

C term

- C1. Review
- C2. Systematic review
- C3. Systematic literature review
- C4. Systematic mapping
- C5. Mapping study
- C6. Systematic mapping study

When this search was performed in the electronic databases, no related secondary studies were identified. Therefore, we sought systematic reviews focused on software engineering frameworks in any field. Table 1 summarizes the secondary studies found.

Mubin et al [19] performed a review on gamification design framework and its application for children with autism. This review aimed to offer gamification solutions for interaction skills. They identified the framework phases in 5 papers and target users/audience/focus. The authors concluded that frameworks have been analyzed from an in-game context but did not emphasize on children with autism. In the literature, studies show that gamification is very effective in the areas of therapy and education for children with autism. The most important contribution of this review is the development of interaction skills. This review identified phases of the development process in some studies (eg, planning, designing). However, it does not explain how users benefit from the process interaction.

Vargas et al [20] developed a systematic mapping study on serious game quality. The aim was to discover the current state of serious games quality initiatives. One of the research questions focused on discovering if quality has been constant throughout the software development cycle or in some stages. The authors showed that 97% of the literature reviewed applied quality in the final phase (product). Only 7.14% focused on quality in the design phase and 1.79% in the requirement phase. This study was included because it identified the phases in which quality was applied: requirement, design, code, and final product.

Tomalá-Gonzáles et al [21] reported on methodologies, game engines currently used in serious games development in various areas (education, cognitive disabilities, and physical rehabilitation), and criteria for game engine selection. From the 27 papers, 8 used a defined methodology such as XP, Cascade, and others, while 3 proposed their own model. The authors concluded that although several software development methodologies can be adapted to serious game development, the best option was the SUM methodology because it is based on Scrum (fast, precise, optimized, and adaptable programming characteristics). However, this review did not make distinctions between framework and methodology. It also did not identify methodology phases nor the benefits of applying a methodology in the learning or rehabilitation process.

Table 1. Summary of secondary studies.

Study	Type	Year of publication	Target users/audience /focus	Benefit of framework	Phases of process development identified?
Mubin et al [19]	Review	2019	Children with autism	Interaction skills in children with autism	Yes
Vargas et al [20]	Systematic mapping study	2014	Serious games quality	Quality	Yes
Tomalá-Gonzáles et al [21]	Review	2020	Identifies methodologies and game engines	__[a]	No

[a]Not available.

Although our work shares similarities with the aforementioned studies, the literature review presented in this paper is different because this review (1) focuses on serious games for physical rehabilitation, (2) identifies the software development stages in each software engineering framework according to the structural activities proposed by Pressman [9], who states that “The software process details will be different in each case, but the structural activities are the same”; (3) identifies contributions of software engineering frameworks to the rehabilitation process; and (4) identifies if the proposed software engineering framework provides objective monitoring of the rehabilitation process.

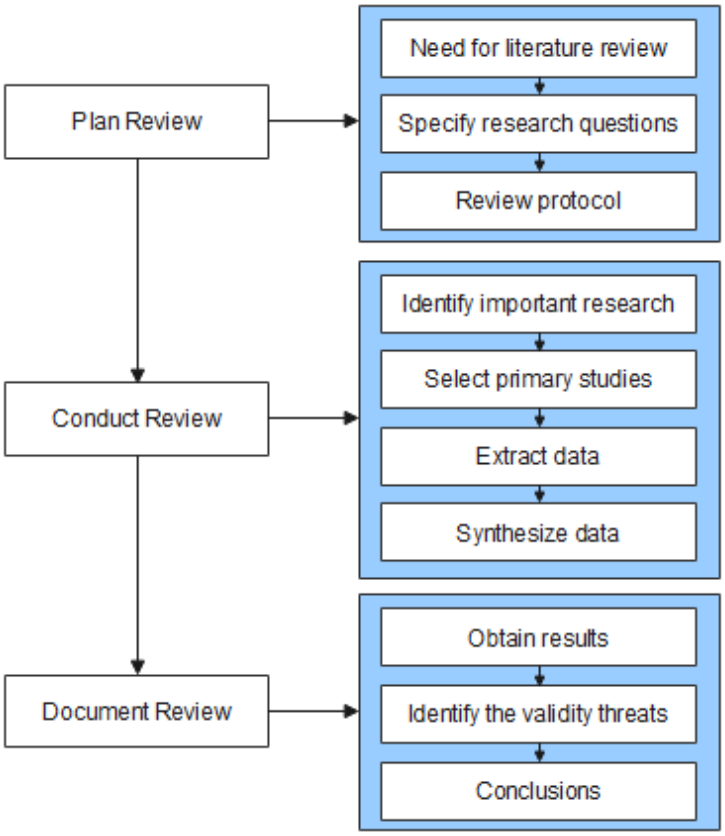
XSL•FO
RenderX

Methods

Research Methodology

The systematic literature review process proposed by Brereton et al [22] was applied for this systematic review. Figure 1 shows the process and steps for each phase. The process consists of 3 main phases: plan review, conduct review, and document review. The first phase consists of the following steps: (1) describe the main reasons for the literature review, (2) specify a set of research questions, and (3) review the protocol. The second phase comprises 4 steps: (1) identify important research, (2) select primary studies, (3) extract data from primary studies, and (4) synthesize data. Finally, the third phase consists of 3 steps: (1) obtain results, (2) identify the validity threats, and (3) conclusions. Figure 1 shows the literature review process. In the following subsections, we describe the activities carried out in each phase of this systematic literature review.

Figure 1. Literature review process.



Research Questions

In this subsection, we present the 9 research questions that guided this study through the investigation to meet the objectives of the systematic review. Table 2 presents these questions.

The research questions can be classified into 4 fields of interest. RQ1 and RQ2 study serious games evaluated in software engineering. These questions identify the number of serious games developed with a software engineering framework and the set of activities, actions, and tasks required.

RQ3 and RQ4 describe framework contributions to the rehabilitation process and implementation of gamification elements. It allows transforming obstacles into positive and fun reinforcements, thereby encouraging patients.

RQ5 and RQ6 are centered on applicability and serious game characteristics for rehabilitation using a software engineering framework. These questions identify relevant data such as target audience, interaction technology for data acquisition, main modalities, among others.

Finally, RQ7, RQ8, and RQ9 studied important aspects to evaluate and provide follow-up of rehabilitation progress depending on the type of exercise.

XSL•FO
RenderX

Table 2. Research questions.

Research question	Question
1	What framework is used in the development of the serious game?
2	What are the generic structural activities used in frameworks?
3	How the framework contributes to the rehabilitation process?
4	What gamification elements does the framework use?
5	What is the targeted disability contemplated in the frameworks?
6	If the framework includes a case study, which part of the body is rehabilitated? What is the modality of the serious game? Which interaction technology is used?
7	What type of evaluation and number of patients are involved in the clinical trials?
8	Does the framework contemplate a standardized scale to evaluate the patient's rehabilitation progress?
9	Does the framework contemplate adaptability?

Search Strategy

The objective of the search strategy was to identify all relevant primary studies. A literature search was conducted to answer the proposed research questions.

The search strategy is an adaptation of Guidelines for Performing Systematic Literature Reviews in Software Engineering [18] and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [23]. Relevant papers were identified by searching in the following databases: PubMed, PEDro (Physiotherapy Evidence Database), IEEE Xplore, ScienceDirect, ACM Digital Library, Mary Ann Liebert, Taylor & Francis Online, Wiley Online Library, and Springer. To build the search string, a list of keywords and their synonyms were identified. Logical operators (AND and OR) and words related to rehabilitation, serious games, and framework were used. The final search strings consisted of the following Boolean expressions: "(A1 AND (B1 OR B2)) AND (C1 OR C2 OR C3) AND D1". The search terms are shown in Textbox 3.

Textbox 3. Search terms for the final search string.

A term

- A1. Serious

B term

- B1. Game
- B2. Games

C term

- C1. Rehabilitation
- C2. Disability
- C3. Disabilities

D term

- D1. Framework

Inclusion Criteria

The systematic review is focused on serious games for physical rehabilitation; clear inclusion criteria were established to determine the eligibility of papers for inclusion in the review. Only studies with the following criteria were considered eligible for inclusion: serious game papers for physical rehabilitation, papers published in English, and all serious games regardless of the year of development.

Exclusion Criteria

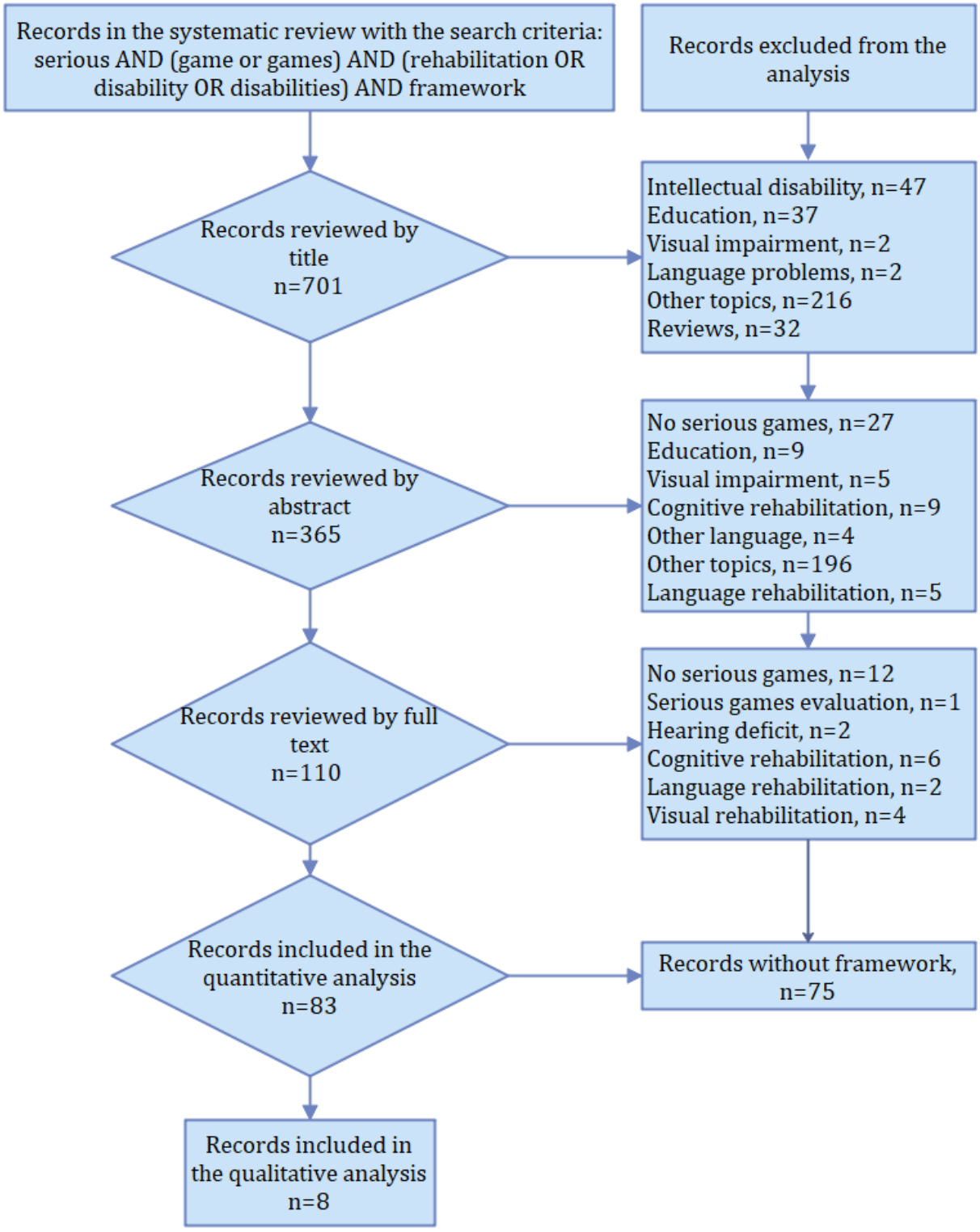
Papers duplicated, papers regarding opinion pieces, existing literature reviews, papers that are not related to rehabilitation using serious games, serious games for educational purposes, and serious games for cognitive rehabilitation were excluded from the study.

Study Selection

First, the search string was used in different databases. Potentially relevant papers were identified after reading the title and abstract. Duplicate papers were removed. Subsequently, an exhaustive verification of compliance with the inclusion and exclusion criteria was carried out to select the papers. Figure 2 shows the item selection process. In the systematic review, 701 papers were included. Table 3 shows the number of documents retrieved from each database.

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)–based flowchart.



XSL•FO
RenderX

Table 3. Search results.

Databases	Results, n
PubMed	14
PEDro	12
IEEE Xplore	103
ScienceDirect	88
ACM Digital Library	166
Mary Ann Liebert	27
Taylor & Francis Online	50
Wiley Online Library	43
Springer	198

Extract Data From Primary Studies

After identification, the primary papers were rigorously analyzed in accordance with the following considerations: (1) only the authors of this review can participate in the data collection process; (2) each primary paper should be reviewed with at least two reviewers; (3) each reviewer will collect a set of data from each primary study, then meet with another reviewer to reach an agreement on the data obtained.

Two types of data were extracted for each study: bibliographic (title, author name, country, year, database) and content data, which are used to answer the research questions. Table 4 shows the concentration of the bibliographic data of the primary papers.

Multimedia Appendix 1 shows the percentage of primary studies from each electronic database. IEEE Xplore presented more primary studies than the rest. The following section presents an analysis of the data collected.

Table 4. Bibliographic data of the primary papers.

Study	Year	Country	Database
Baranyi et al [24]	2013	Austria	IEEE Xplore
Pirovano et al [25]	2016	Italy	ScienceDirect
Amengual Alcover et al [26]	2018	Spain	PubMed
Baranyi et al [27]	2019	Austria	IEEE Xplore
Zain et al [28]	2012	Malaysia	IEEE Xplore
Noveletto et al [29]	2018	Brazil	ScienceDirect
Afyouni et al [30]	2017	Cyprus	ACM Digital Library
Maggiorini et al [31]	2012	Switzerland	ACM Digital Library

Results

RQ1: What Framework Is Used in the Development of the Serious Game?

Only 8 (10%) out of the 83 papers related to physical rehabilitation using a software engineering framework (Multimedia Appendix 2).

In Baranyi et al [24,27], the proposed studies were based on the user-centered design framework. The physiotherapist is important because s/he identifies the needs and limitations of the patients in the rehabilitation process. There are 3 phases: research, design, and evaluation. In research, a physiotherapist conducts brainstorming with the work team and identifies the requirements. Afterward, in the design phase, the team creates mock-ups and a prototype. Finally, the physiotherapist evaluates the application.

Pirovano et al [25] proposed a 4-step procedure to create safe exergames for rehabilitation therapies: exercise definition, virtualization, game design, and secondary goals. In exercise definition, a set of exercises is proposed to cover therapy needs. Each exercise is structured into primary and secondary goals. During virtualization, the team identifies primary goals, and they are implemented into a virtual exercise by defining input (tracking) and output (feedback) requirements through simple graphical elements and specifying interaction mechanisms. Through game design, the virtual exercise is converted into a true exergame. In the last step, there are 2 functionalities. The first is to analyze motion data and identify wrong movements. The second provides feedback to the patients.

In Amengual Alcover et al [26], the serious game development framework follows an iterative process flow structured into 2 dimensions: activities and incremental development. The first dimension is based on 3 approaches: Scrum, the web application development model, and a clinical trial. The activities dimension includes a project initiation activity, an iterative flow composed of 4 basic development activities (planning and control, modeling, construction, and evaluation), and a final clinical

XSL•FO
RenderX

study to evaluate the rehabilitation process of the patient through the serious game. Incremental development includes 3 different increments: interaction mechanism, interaction elements, and serious game. In the first increment, an existing device on the market is identified to capture the movements of patients according to their needs. In the second increment, the development team must design the interaction elements that force patients to perform the therapy correctly. The final increment is aimed at designing a serious game that motivates the patient to perform therapy to obtain the best results.

Zain et al [28] proposed a conceptual framework for people with motor impairment, so they can enjoy the experience of playing serious games. The framework's main elements were player skills, challenge, concentration, feedback, immersion, learning opportunities, accessibility, and adaptivity. The proposed framework will help the game designer and developer create a serious game that combines the game's technology with the learning environment. This framework is based on the game flow model.

Noveletto et al [29] presented a conceptual model for the design or development of serious games to rehabilitate people with stroke. The framework establishes a relationship between experts and patients to obtain the requirements, considering that the biomedical device and the video game score are used to design serious games.

Afyouni et al [30] proposed a framework consisting of a therapy-driven 3D environment augmented with a natural user interface based on movement. The framework incorporated different adaptation techniques to adjust patient's needs. Patient preferences and limitations were considered key parameters for changing the game, thereby creating personalized games for each patient.

Maggiorini et al [31] presented a framework for serious game development that allows the therapist to remotely control the video game home activities. The objective was to create a more attractive game for the elderly with easily adjustable parameters for therapy adaptability. The framework includes 3 phases of serious game development: requirements definition, empirical validation of requirements list, and design and prototyping.

RQ2: What Are the Generic Structural Activities Used in Frameworks?

The objective of this research question was to identify generic structural activities in primary studies (see the "Background" section). Table 5 summarizes the structural activities and Multimedia Appendix 3 shows the frequency of occurrence of each structural activity in primary studies.

Every study established a communication activity to obtain the requirements. Baranyi et al [24] brainstormed with a physiotherapist to identify relevant problems and needs for patients undergoing rehabilitation. Pirovano et al [25] defined exercises addressing the primary and secondary objectives of rehabilitation. To achieve maximum effectiveness, the exercises are defined in collaboration with therapists. In Amengual Alcover et al [26], the communication began by identifying the context, operational objectives, restrictions, and requirements. Baranyi et al [27] established communication with the therapist to obtain the requirements. Zain et al [28] identified the user abilities, limitations, and behavior, which become requirements for the serious game. Noveletto et al [29] considered experts in the field (health personnel, therapists, etc.) and patients to obtain the requirements. Afyouni et al [30] established the type of game through patient needs, preferences, and limitations, allowing custom game features. Finally, Maggiorini et al [31] analyzed the most diffused issues present in elders' homes (eg, size of rooms, habits) to explore requirements and limitations through an immersive approach.

The planning activity was implemented in Amengual Alcover et al [26]. The goal of this activity was to determine the tasks to perform during the development by identifying the end products and the people who will do the work. The activity includes 3 tasks: planning, scheduling, and tracking.

The modeling activity was performed in several papers. For example, Baranyi et al [24] called it design, elaborating basic models discussed with a therapist. Pirovano et al [25] transformed the exercise requirements into a true exergame by adding all the elements and characteristics of a game and a good game design for the patients. Amengual Alcover et al [26] created models that helped the development team to understand the requirements obtained and the game design. By contrast, Baranyi et al [27] contemplated the use of prototypes to refine user requirements. Finally, Maggiorini et al [31] established a list of technical characteristics (desired) for the prototype creation.

The construction activity was implemented in every study. Developments produce executable software units that will be used by users, through the creation of prototypes to improve the software [24-27,30,31], or the final product [28,29].

Finally, the user evaluates and provides feedback on the serious game in the deployment stage. In the primary papers, Pirovano et al [25] and Baranyi et al [27], patients were asked to give their opinion to improve the game design and change some aspects of the application.

XSL•FO
RenderX

Table 5. Structural activities in primary studies.

Study	Communication	Planning	Modeling	Construction	Deploy
Baranyi et al [24]	X	__[a]	X	X	—
Pirovano et al [25]	X	—	X	X	X
Amengual Alcover et al [26]	X	X	X	X	—
Baranyi et al [27]	X	—	X	X	X
Zain et al [28]	X	—	—	X	—
Noveletto et al [29]	X	—	—	X	—
Afyouni et al [30]	X	—	—	X	—
Maggiorini et al [31]	X	—	X	X	—

[a]Not available.

RQ3: How the Framework Contributes to the Rehabilitation Process?

Baranyi et al [24,27] applied a user-centered design approach to establish constant communication with the physiotherapist who has the experience to identify the needs and limitations of the final user. Serious games are developed with entertainment elements such as levels, rewards, challenges, and adaptability to the patient need, considering special conditions.

Pirovano et al [25] proposed the creation of safe exergames, identifying the needs of real exercise besides therapy goals. These needs are incorporated into a video game considering the primary objectives (what a user should do) and secondary objectives (how user actions should be carried out). The former is easily integrated into the gameplay, while the latter aids the patient with corrections or prevention of compensatory movement through analysis of the flow of movement data and wrong movements in real time, thereby providing immediate feedback to patients to correct themselves during the exercise.

Amengual Alcover et al [26] proposed an iterative, prototype-oriented, systematized serious game development process. The proposed process guarantees that products based on this framework are developed and validated by following a coherent and systematic method that leads to high-quality serious games.

For users with motor impairment, Zain et al [28] used flow theory [32] to propose user interface design factors that make their experience enjoyable when playing serious games. This framework includes user interface design factors and aims to establish a conceptual model that can be used by a game designer for efficient game development or an educational practitioner when designing enjoyable serious games for users with motor impairment.

Noveletto et al [29] established a relationship among key stakeholders (experts and patient) and elements (biomedical device and game score) for serious game design. The framework states that a correlation between the game score and clinical tests can aid treatment and evaluation through the biomedical system.

Afyouni et al [30] proposed a framework for video game development with an adaptive and user-centered approach. The framework embeds different adaptation techniques to tailor to patients' needs. The video game adapts to the difficulty level based on the patient's profile and performance in real time. Other aspects such as patient preferences and constraints are considered as key game-changing parameters.

Finally, in Maggiorini et al [31], the framework allowed serious game development with telerehabilitation allowing the therapist to remotely control the video game home activities. It supports parameter adjustments for therapy adaptability. Table 6 summarizes framework contributions.

XSL•FO
RenderX

Table 6. Framework contributions in primary studies.

Framework contribution to rehabilitation	Utility	Primary studies
Communication with health expert	A physiotherapist establishes communications with patients undergoing rehabilitation to identify the problems and needs.	Baranyi et al [24], Pirovano et al [25], Baranyi et al [27], Noveletto et al [29]
Exercise definition	Exercise can be defined as a sequence of different actions needed to complete it to achieve maximum effectiveness.	Pirovano et al [25]
Analyzes the stream of motion data and identifies in real time wrong movements	Provides immediate feedback to the patients for correct exercising.	Pirovano et al [25]
Iterative and prototyping	Visualize prototypes of serious games from early stages. The therapist or patients identify additional requirements or modify them.	Baranyi et al [24], Pirovano et al [25], Baranyi et al [27], Noveletto et al [29], Afyouni et al [30], Maggiorini et al [31]
User interface design factors	Motivation and immersion	Baranyi et al [24], Pirovano et al [25], Amengual Alcover et al [26], Baranyi et al [27], Zain et al [28], Noveletto et al [29]
The correlation between game score and clinical tests	Aids in patient treatment and evaluation	Noveletto et al [29]
Adaptive approach	Adapts difficulty level according to the patient's profile and performance in real time	Baranyi et al [24], Pirovano et al [25], Zain et al [28] Afyouni et al [30], Maggiorini et al [31]
Telerehabilitation	Therapists can remotely control the video game for home activities and provide adjustable parameters to improve therapy	Maggiorini et al [31]

RQ4: What Gamification Elements Does the Framework Use?

Overview

Gamification allows the transformation of obstacles into positive and fun reinforcement, encouraging users to make the right decisions for their health and well-being [33]. It is essential to keep the patient motivated in physical rehabilitation. For this reason, the software engineering framework is required to use gamification elements. The papers identified the following elements: feedback, motivational factor, adaptability, challenge, levels, immersion, rewards, concentration, and avatar. Table 7 shows the gamification elements in primary studies, and Multimedia Appendix 4 shows the frequency of occurrence of each gamification element.

The gamification elements of primary studies are described below.

Table 7. Gamification elements in primary studies.

Gamification element	Study
Feedback	Pirovano et al [25], Amengual Alcover et al [26], Baranyi et al [27], Zain et al [28], Noveletto et al [29], Afyouni et al [30], Maggiorini et al [31]
Motivational factor	Baranyi et al [24], Pirovano et al [25], Amengual Alcover et al [26], Baranyi et al [27], Noveletto et al [29]
Adaptability	Baranyi et al [24], Pirovano et al [25], Zain et al [28], Afyouni et al [30], Maggiorini et al [31]
Challenge	Baranyi et al [24], Zain et al [28], Afyouni et al [30]
Levels	Baranyi et al [24], Amengual Alcover et al [26], Baranyi et al [27], Afyouni et al [30]
Immersion	Zain et al [28]
Rewards	Pirovano et al [25]
Concentration	Zain et al [28]
Avatar	Pirovano et al [25]

Feedback

In Pirovano et al [25], the feedback mechanisms were designed to show the outcome of actions to patients. For instance, whether a target is met or a movement has been successfully performed. Amengual Alcover et al [26] used "mirror feedback," which consists of projecting the user onto the screen and simulating a mirror in such a way that the users can see themselves on the screen at all times. In Baranyi et al [27], the feedback provided was either visual, aural, or haptic. In Zain et al [28], users with motor impairment received feedback on their progress, and when they lose the game, feedback is provided to continue in the right direction. Noveletto et al [29] established that serious same should reward players with feedback on progress. Afyouni et al [30] used a scoring system that was designed to keep track of the number of times the patient successfully passed through

XSL•FO
RenderX

the targets. Finally, in Maggiorini et al [31], a skeleton wireframe is drawn in red to provide immediate visual feedback, and an alarm is raised on the screen.

Motivational Factor

Baranyi et al [24] used "goals." The gameplay was based on achieving goals that should act as motivation factors. Pirovano et al [25] established that extrinsic motivational effects can be achieved through careful use of verbal praise, scoring mechanisms, and virtual reward systems. In Amengual Alcover et al [26], the development of new serious games allowed the inclusion of motivational elements to increase engagement. Baranyi et al [27] used rewards in serious games for the user. Finally, Noveletto et al [29] used the "motivational score" to improve attention during rehabilitation sessions.

Adaptability

Baranyi et al [24] proposed an adaptive system with the opportunity to adapt the game difficulty. Pirovano et al [25] established that virtual exercises should use dynamic difficulty adaptation, thus further increasing the flexibility of serious games. For Zain et al [28], an adaptive factor was important to design and develop serious games for users with motor impairment because the application, aware of the users' current cognitive load and physical limitations, can change its response, presentation, and interaction flow to improve users' experience and their task performance. In Afyouni et al [30], the framework embeds different adaptation techniques to adapt to the patients' needs. Key game-changing parameters such as patient preferences and constraints are considered. This allows the creation of personalized game features for every patient. Maggiorini et al [31] proposed that remotely controlled serious games may also provide easily tunable parameters to better adapt the game therapy to the actual patient recovery.

Challenge

Baranyi et al [24] proposed the challenge as a "key fact." They considered that the game should not be too easy nor too hard to manage. The game should be sufficiently challenging and match the player's skill level. Zain et al [28] proposed that serious games should also vary the level of difficulty and keep an appropriate pace. Afyouni et al [30] generated therapy-aware navigational movements with multiple levels of difficulty.

Levels

Baranyi et al [24] stated that the purpose of the serious game developed is to have a rehabilitation system containing different levels that were adapted and created for the individual needs of the patients and to fit their impairments. Amengual Alcover et al [26] considered that serious games must have a definition of different levels in the game. In Baranyi et al [28], when the game is started for the first time, a diagnostic routine is performed; using these data, a baseline for the exercises can be defined by the therapist to get an initial idea about how easy or complex a level might be for a patient. Afyouni et al [30] presented different levels of difficulty based on therapeutic gestures and patient performance.

Immersion

Zain et al [28] considered that immersive games draw players into the game and affect their senses through elements such as audio and narrative.

Reward

Pirovano et al [25] used a scoring system, and at the end of each exergame, a virtual reward is presented to the patients.

Concentration

Zain et al [28] considered that the more concentration a task requires in terms of attention and workload, the more absorbing it will be. The games should grab the player's attention quickly and maintain it throughout the game.

Avatar

Pirovano et al [25] used an avatar for feedback on wrong movements, changing the color of the associated avatar segments. When wrong movements persist for a long time, the game is paused, and a virtual therapist avatar pops up to advise patients.

RQ5: What Is the Targeted Disability Contemplated in the Frameworks?

This specifies whether a study focuses on a particular pathology with loss or decrease in movement. The papers established the following target pathology: 4 defined strokes [24,25,27,29], 2 defined neuromotor disorder [26,30], 1 defined users with motor impairment [28], and 1 defined rehabilitation of the elderly [31]. Stroke is mainly targeted in studies because it is the second cause of death and the third cause of disability worldwide [34]. Multimedia Appendix 5 shows the target disability percentage.

RQ6: If the Framework Includes a Case Study, Which Part of the Body Is Rehabilitated? What Is the Modality of the Serious Game? Which Interaction Technology Is Used?

As Table 8 reports, Baranyi et al [24] presented a prototype that rehabilitates patients with lower limb disabilities with balance and strength exercises using the Wii Fit Balance Board. Pirovano et al [25] developed serious games for upper limb motor rehabilitation therapy using Microsoft Kinect and lower limb with the Wii Fit Balance Board. Amengual Alcover et al [26] also rehabilitated the lower limb by allowing patients to perform repetitions in a video game controlled with Microsoft Kinect, with each repetition varied according to the participant's tolerance and the physiotherapist's recommendations. Baranyi et al [27] performed hand rehabilitation using gesture exercises, touch, and patient movement levels using mobile phone sensors. Zain et al [28] and Noveletto et al [29] did not report any case studies. Afyouni et al [30] developed a serious game for hand rehabilitation using leap motion. Game instructions can be visual (shown on the screen) or voice, depending on the perception capacity of the patient. Finally, Maggiorini et al [31] developed a prototype for rehabilitation using Microsoft Kinect. It only presents the skeleton tracking by a sensor and does not mention whether the video game implements another form of communication with the patient.

XSL•FO
RenderX

The modality is a way in which information is transmitted from the computer to the participants [35]. Baranyi et al [24,27], Pirovano et al [25], Amengual Alcover et al [26], Afyouni et al [30], and Maggiorini et al [31] used a visual modality, presenting a graphical interface for user interaction. Pirovano et al [25], Amengual Alcover et al [26], Baranyi et al [27], and Afyouni et al [30] used audio effects such as music or voice instructions. Baranyi et al [27] used haptic modality to control the video game through a touch screen. Zain et al [28] and Noveletto et al [29] did not report modalities.

Table 8. Rehabilitated limb, serious game modality, and data-acquisition device in primary studies.

Study	Rehabilitation/extremity	Modality	Interaction technology
Baranyi et al [24]	Lower limbs	Visual	Wii Fit Balance Board
Pirovano et al [25]	Lower and upper limbs	Visual, auditory	Wii Fit Balance Board and Microsoft Kinect
Amengual Alcover et al [26]	Lower limbs	Visual, auditory	Microsoft Kinect
Baranyi et al [27]	Hand	Visual, auditory, haptic	iOS platform sensors
Zain et al [28]	Not reported	Not reported	Open
Noveletto et al [29]	Not reported	Not reported	Open
Afyouni et al [30]	Hand	Visual, auditory	Leap motion
Maggiorini et al [31]	Full body	Visual	Microsoft Kinect

RQ7: What Type of Evaluation and Number of Patients Are Involved in the Clinical Trials?

The objective of this research question was to identify clinical validation of the studies and the number of patients involved. In clinical trials, participants receive specific interventions according to the research plan or protocol created by the researchers to determine the safety and efficacy of the interventions through the measurements of the outcomes [36]. Table 9 shows these data. Amengual Alcover et al [26] conducted a clinical trial and observed a significant difference between before and after scores. They used the Berg Balance Scale and their results showed a significant functional improvement (P=.002) in comparison with assessments before (mean 29.5 [SD 3.9] and after (mean 34.1 [SD 2.2]) the intervention. The Functional Reach Test revealed significant differences in functional balance before and after the intervention: right upper limb, before (mean 8.6 [SD 1.4]) and after intervention (mean 10.1 [SD 2.0]; P=.007); and left upper limb, before (mean 8.3 [SD 2.0]) and after intervention (mean 10.1 [SD 3.7]; P=.052). Finally, a significant difference between the pre- and post-assessment scores for the Tinetti Balance Test was observed at the end of the 24-week intervention period. The average score rose from 16 to 21 points on a scale of 28 points. Afyouni et al [30] reported that patients showed improved hand movement using a range of motion. They were able to document 66% of the elements in the video game. No other study reported a clinical trial.

Table 9. Type of evaluation and number of patients in the primary studies.

Study	Evaluation	Number of patients
Baranyi et al [24]	No clinical validation	N/A[a]
Pirovano et al [25]	No clinical validation	N/A
Amengual Alcover et al [26]	Clinical trial	9
Baranyi et al [27]	No clinical validation	N/A
Zain et al [28]	No clinical validation	N/A
Noveletto et al [29]	No clinical validation	N/A
Afyouni et al [30]	Clinical trial	5
Maggiorini et al [31]	No clinical validation	N/A

[a]NA: not applicable.

RQ8: Does the Framework Contemplate a Standardized Scale to Evaluate the Patient's Rehabilitation Progress?

An assessment instrument allows to objectively quantify the disability degree of the patient and measure the progress of rehabilitation [37,38]. The evaluation scales in the framework are used to quantify the improvement in rehabilitation depending on the type of exercise applied. During the analysis of primary papers, we identified 3 studies with assessment instruments: Pirovano et al [25], Amengual Alcover et al [26], and Afyouni et al [30].

RQ9: Does the Framework Contemplate Adaptability?

Adaptability is the ability to dynamically adapt difficulty in a video game according to the patient's performance [39]. Five primary studies use this characteristic. In Baranyi et al [24], the physiotherapist designed the level of difficulty of the video

XSL•FO
RenderX

game. Pirovano et al [25] established that for every exercise, quality parameters are necessary to define movement properties. This will allow one to determine the challenge degree of the exercises and adapt the difficulty to the patient's needs. Zain et al [28] mentioned that adaptability must consider the following elements: (1) user motivation, (2) experience and abilities, and (3) detection, which identifies necessary changes. Afyouni et al [30] adapted the difficulty level based on the patient's profile and performance in real time. In Maggiorini et al [31] the therapist can remotely adapt the game therapy to the patient's actual recovery. Amengual Alcover et al [26], Baranyi et al [27], and Noveletto et al [29] did not specify how adaptability is incorporated into their game. Multimedia Appendix 6 shows the percentage of frameworks contemplating adaptability.

Threats to Validity of Primary Studies Selected

Although we used search strategies and techniques to systematically find papers by using keywords in the selected databases, these words may vary within papers, so some relevant studies may have been omitted.

Discussion

Preliminary Findings

We found only few studies that used a systematic process for serious game development. Each framework analyzed in the primary papers highlighted a different feature.

Planning was the structural activity least implemented. This activity is essential because it allows goal definition, objectives, and path to follow in the software development [9,10,40,41].

Regarding applicability, most studies focused on the treatment of stroke sequelae using various modalities such as visual and auditory. The latter should also be implemented to provide feedback on patient performance. Lastly, test cases directly use playable commercial platforms such as Microsoft Kinect and Leap motion as interaction technology.

There were a few clinical trials, and the type of improvement reported varies from one study to another. Amengual Alcover et al [26] used the Berg Balance Scale and Tinetti Balance Test measurements and reported significant functional improvement from previous results. Afyouni et al [30] also reported improvements using range of motion evaluation in hand movement. No other studies used clinical trials to evaluate the framework. Clinical evaluation is essential to objectively validate the patient's rehabilitation progress [36].

Pirovano et al [25], Amengual Alcover et al [26], and Afyouni et al [30] used an evaluation scale to assess the patient's progress. It should also be used as an alternative to adaptability, which is essential for progress and motivation [42]. It is also a technique that can be used to advance game levels [5]. Game levels help engage in the game and could increase treatment compliance.

Conclusions

The objective of this study was to identify the software engineering frameworks used in the development of serious games through a literature review of 8 primary studies. The conclusions of this study are as follows:

About 75% (6/8) of the primary papers proposed a framework [25,26,28-31], whereas the rest were adaptations of the user-centered design framework (RQ1). Regarding the structural activities, 100% (8/8) of the papers applied the communication and construction activity [24-31], 63% (5/8) used modeling (known as a design in some developments) [24-27,31], 25% (2/8) considered user feedback to improve the serious games [25,27], and only 13% (1/8) included the planning phase [26] (RQ2).

Each primary study contributes in one or more aspects to the rehabilitation process. Baranyi et al [24,27] applied a user-centered design using which the physiotherapist can personalize individual needs in the serious game. Pirovano et al [25] proposed ease of play and assisted help during the rehabilitation exercise. Amengual Alcover et al [26] developed a framework for motor rehabilitation therapies using a systematized process. Zain et al [28] embraced immersion and fun in the game to maintain flow interest. Noveletto et al [29] used game scores for patient assessment. Afyouni et al [30] developed games with dynamic adaptability that were patient centered. Finally, Maggiorini et al [31] incorporated telerehabilitation and adaptability for the elderly to perform rehabilitation exercises at home (RQ3). Every study applies gamification elements that allow patients to transform rehabilitation obstacles into positive and fun reinforcements. Feedback was the gamification element most applied (7/8, 88%) [25-31]. Other elements frequently implemented were adaptability [24,25,28,30,31] and motivational factor [24-27,29] (5/8, 63%) for both; RQ4.

Stroke is the primary pathology on which serious games are focused. This pathology is the third cause of disability worldwide, and a characteristic symptom is the sudden, generally unilateral, loss of muscle strength in the arms, legs, or face (RQ5). Regarding the case studies of limb rehabilitation, 2 studies [24,26] included the lower limb, 1 [25] included lower and upper limbs, 2 [27,30] included hand, 1 [31] full body, and 2 [28,29] did not report case studies. The most used video game modality was visual (6/8, 75%) [24-27,30,31], followed by auditory (4/8, 50%) [25-27,30]. Although each case study used a different motion acquisition technology, every framework allowed a wide variety of the interaction style to obtain the patient's movement and control the serious game (RQ6).

Of the primary papers, 25% (2/8) applied a clinical evaluation to assess patient improvement when the serious game is used [26,30] (RQ7). To objectively evaluate progress and identify abilities and deficits, only 38% (3/8) of the primary studies used an assessment instrument [25,26,30] (RQ8). The assessment used standardized procedures indicating how a patient of any given age and intelligence level would perform. Adjusting the video game difficulty to the patient's rehabilitation needs is essential to avoid frustration or boredom, and 63% (5/8) of the primary studies used adaptability [24,25,28,30,31] (RQ9).

Finally, we recommend that all serious games have to be developed with a framework or methodology. If for some reason this is not possible, they should at least involve the therapist to

XSL•FO
RenderX

define requirements. It is also important to include evaluation scales to measure the patient's progress and gamification elements. Besides, the video game development must be an iterative and incremental process based on generic structural activities and the patient should be considered in the validation and feedback phases.

We propose the following recommendations for future studies:

- Carry out a study of the papers that propose a methodology for serious game development.
- Study software engineering framework proposals in serious games from other fields, such as education.
- Develop a software engineering framework applying all the structural activities and gamification elements for the creation of serious games for physical rehabilitation.

Acknowledgments

The authors thank the National Council of Science and Technology (CONACyT) for their support of this study. Additionally, this work was supported by Secretariat of Public Education (SEP) through PRODEP (Programa para el Desarrollo Profesional Docente).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Percentage of primary studies provided by each electronic database.
[PNG File , 18 KB - games_v9i4e25831_app1.png]

Multimedia Appendix 2
Percentage of serious games that used a software engineering framework.
[PNG File , 21 KB - games_v9i4e25831_app2.png]

Multimedia Appendix 3
Frequency of occurrence of structural activities in primary studies.
[PNG File , 7 KB - games_v9i4e25831_app3.png]

Multimedia Appendix 4
Frequency of occurrence of each gamification element.
[PNG File , 11 KB - games_v9i4e25831_app4.png]

Multimedia Appendix 5
Target disability contemplated in frameworks.
[PNG File , 15 KB - games_v9i4e25831_app5.png]

Multimedia Appendix 6
Frameworks contemplating adaptability.
[PNG File , 17 KB - games_v9i4e25831_app6.png]

References

1. World Report on Disability 2011. Geneva, Switzerland: World Health Organization; 2011. URL: https://apps.who.int/iris/handle/10665/44575 [accessed 2021-01-31]
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. 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. Frontiers in Human Neuroscience 2016;10:1-22 [FREE Full text] [doi: 10.3389/fnhum.2016.00442] [Medline: 27679565]
4. Michael D, Chen S. Serious Games: Games That Educate, Train, and Inform. Boston, MA: Thomson Course Technology; 2005:287.
5. 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]
6. Cortegiani A, Absalom AR. Importance of proper conduct of clinical trials. Br J Anaesth 2021 Feb;126(2):354-356. [doi: 10.1016/j.bja.2020.09.030] [Medline: 33121749]

XSL•FO
RenderX

7. Mnkandla E. About software engineering frameworks and methodologies. New York, NY: IEEE; 2009 Presented at: AFRICON 2009; September 23-25, 2009; Nairobi, Kenya. [doi: 10.1109/afrcon.2009.5308117]
8. Schmidt DC, Gokhale A, Natarajan B. Leveraging Application Frameworks. Queue 2004 Jul;2(5):66-75. [doi: 10.1145/1016998.1017005]
9. Pressman R. Software Engineering: A Practitioner's Approach. New York, NY: McGraw Hill; 2014:941.
10. Sommerville I. Software Engineering. London, UK: Pearson; 2015:816.
11. Vliet H. Software Engineering: Principles and Practice. Hoboken, NJ: Wiley; 2008:552.
12. Schmidt R. Software Engineering Architecture-Driven Software Development. St Louis, MO: Elsevier/Morgan Kaufmann; 2013:376.
13. Kumar J. Gamification at work: designing engaging business software. In: Marcus A, editor. Design, User Experience, and Usability. Health, Learning, Playing, Cultural, and Cross-Cultural User Experience. DUXU 2013. Berlin, Germany: Springer; 2013:528-537.
14. Schulz R, Martinez S, Hara T. Towards a Game-Design Framework for Evidence-Based Clinical Procedure Libraries. New York, NY: IEEE; 2019 Presented at: IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); August 5-7, 2019; Kyoto, Japan. [doi: 10.1109/SeGAH.2019.8882474]
15. 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]
16. de Castro-Cros M, Sebastian-Romagosa M, Rodríguez-Serrano J, Opisso E, Ochoa M, Ortner R, et al. Effects of Gamification in BCI Functional Rehabilitation. Front Neurosci 2020;14:882 [FREE Full text] [doi: 10.3389/fnins.2020.00882] [Medline: 32973435]
17. Steiner B, Elgert L, Saalfeld B, Wolf K. Gamification in Rehabilitation of Patients With Musculoskeletal Diseases of the Shoulder: Scoping Review. JMIR Serious Games 2020 Aug 25;8(3):1-12 [FREE Full text] [doi: 10.2196/19914] [Medline: 32840488]
18. Kitchenham B. Systematic review in software engineering: where we are and where we should be going. New York, NY: Association for Computing Machinery; 2012 Presented at: 2nd International Workshop on Evidential Assessment of Software Technologies; September 22, 2012; New York, NY. [doi: 10.1145/2372233.2372235]
19. Mubin S, Wee APM. A Review on Gamification Design Framework: How They Incorporated for Autism Children. New York, NY: IEEE; 2019 Presented at: 2019 4th International Conference and Workshops on Recent Advances and Innovations in Engineering; November 27-29, 2019; Kedah, Malaysia. [doi: 10.1109/icraie47735.2019.9037765]
20. Vargas J, García-Mundo L, Genero M, Piattini M. A systematic mapping study on serious game quality. New York, NY: ACM Press; 2014 Presented at: 18th International Conference on Evaluation and Assessment in Software Engineering; May 2014; London, UK. [doi: 10.1145/2601248.2601261]
21. Tomala-Gonzales J, Guaman-Quinche J, Guaman-Quinche E, Chamba-Zaragocin W, Mendoza-Betancourt S. Serious Games: Review of methodologies and Games engines for their development. New York, NY: IEEE; 2020 Presented at: 15th Iberian Conference on Information Systems and Technologies (CISTI); June 24-27, 2020; Sevilla, Spain. [doi: 10.23919/cisti49556.2020.9140827]
22. Brereton P, Kitchenham BA, Budgen D, Turner M, Khalil M. Lessons from applying the systematic literature review process within the software engineering domain. Journal of Systems and Software 2007 Apr;80(4):571-583. [doi: 10.1016/j.jss.2006.07.009]
23. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:1-9 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
24. 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: 2nd International Conference on Serious Games and Applications for Health (SeGAH); May 2-3, 2013; Vilamoura, Portugal. [doi: 10.1109/segah.2013.6665319]
25. 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]
26. 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):1-18 [FREE Full text] [doi: 10.1371/journal.pone.0197383] [Medline: 29768472]
27. 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. New York, NY: IEEE; 2019 Presented at: IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); August 5-7, 2019; Kyoto, Japan. [doi: 10.1109/segah.2019.8882447]
28. 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, 2012; Kuala Lumpur, Malaysia. [doi: 10.1109/iccisci.2012.6297175]

XSL•FO
RenderX

29. Noveletto F, Hounsell M, Soares A, Eichinger F, Sagawa Y, Bertemes Filho P. Stronger: A serious game framework for post-stroke rehabilitation. Annals of Physical and Rehabilitation Medicine 2018 Jul;61:e487. [doi: 10.1016/j.rehab.2018.05.1135]
30. Afyouni I, Qamar A, Hussain S, Ur RF, Sadiq B, Murad A. Motion-based serious games for hand assistive rehabilitation. New York, NY: Association for Computing Machinery; 2017 Presented at: 22nd International Conference on Intelligent User Interfaces Companion; March 13-16, 2017; New York, NY. [doi: 10.1145/3030024.3040977]
31. Maggiorini D, Ripamonti L, Zanon E. Supporting seniors rehabilitation through videogame technology: a distributed approach. New York, NY: IEEE Press; 2012 Presented at: 2012 Second International Workshop on Games and Software Engineering: Realizing User Engagement with Game Engineering Techniques (GAS); June 9, 2012; Zurich, Switzerland. [doi: 10.1109/gas.2012.6225920]
32. Csikszentmihalyi M. The flow experience and its significance for human psychology. In: Optimal Experience Psychological Studies of Flow in Consciousness. Cambridge, UK: Cambridge University Press; 1988:15-35.
33. Pereira P, Duarte E, Rebelo F, Noriega P. A review of gamification for health-related contexts. In: Marcus A, editor. Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments. DUXU 2014. Cham, Switzerland: Springer International Publishing; 2014:742-753.
34. Feigin V, Forouzanfar M, Krishnamurthi R, Mensah G, Connor M, Bennett D, et al. Global and regional burden of stroke during 1990–2010: findings from the Global Burden of Disease Study 2010. The Lancet 2014 Jan;383(9913):245-255. [doi: 10.1016/s0140-6736(13)61953-4]
35. Laamarti F, Eid M, El Saddik A. An Overview of Serious Games. International Journal of Computer Games Technology 2014;2014:1-15. [doi: 10.1155/2014/358152]
36. Wright B. Introduction to clinical trials. In: A Comprehensive and Practical Guide to Clinical Trials. Waltham, MA: Academic Press; 2017.
37. Shamley D, Wright B. A Comprehensive and Practical Guide to Clinical Trials. USA: Academic Press; 2017:210.
38. Martinez R. Clinical Trial Outcomes: What Matters to Patients. JACC Heart Fail 2019 Mar;7(3):272-273 [FREE Full text] [doi: 10.1016/j.jchf.2018.12.003] [Medline: 30819384]
39. Rego P, Moreira P, Reis L. Architecture for Serious Games in Health Rehabilitation. Berlin, Germany: Springer International Publishing; 2014 Presented at: World Conference on Information Systems and Technologies (WorldCIST); April 15-18, 2014; Funchal, Portugal p. 307-317. [doi: 10.1007/978-3-319-05948-8_30]
40. Rook P. Project planning and control. In: Software Engineer's Reference Book. Oxford, United Kingdom: Butterworth-Heinemann; 1991.
41. Rook P. Controlling software projects. Software Engineering Journal 1986;1(1):7-16. [doi: 10.1049/sej.1986.0003]
42. De Lope RP, Medina-Medina N. A Comprehensive Taxonomy for Serious Games. Journal of Educational Computing Research 2016 Dec 12;55(5):629-672. [doi: 10.1177/0735633116681301]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by N Zary; submitted 17.11.20; peer-reviewed by G Alor-Hernández, U Juárez-Martínez; comments to author 08.01.21; revised version received 05.04.21; accepted 22.07.21; published 11.11.21.

Please cite as:
Ambros-Antemate JF, Beristain-Colorado MDP, Vargas-Treviño M, Gutiérrez-Gutiérrez J, Hernández-Cruz PA, Gallegos-Velasco IB, Moreno-Rodríguez A
Software Engineering Frameworks Used for Serious Games Development in Physical Rehabilitation: Systematic Review
JMIR Serious Games 2021;9(4):e25831
URL: https://games.jmir.org/2021/4/e25831
doi:10.2196/25831
PMID:34762052

©Jorge Fernando Ambros-Antemate, María Del Pilar Beristain-Colorado, Marciano Vargas-Treviño, Jaime Gutiérrez-Gutiérrez, Pedro Antonio Hernández-Cruz, Itandehui Belem Gallegos-Velasco, Adriana Moreno-Rodríguez. Originally published in JMIR Serious Games (https://games.jmir.org), 11.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Contribution of the co.LAB Framework to the Collaborative Design of Serious Games: Mixed Methods Validation Study

Dominique Jaccard[1], MSc; Laurent Suppan[2], MD; Félicia Bielser[3], MSc

[1]Media Engineering Institute, HES-SO University of Applied Sciences and Arts Western Switzerland, Yverdon, Switzerland

[2]Division of Emergency Medicine, Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, University of Geneva Hospitals and Faculty of Medicine, Geneva, Switzerland

[3]School of Health Sciences (HESAV), HES-SO University of Applied Sciences and Arts Western Switzerland, Lausanne, Switzerland

Corresponding Author:
Dominique Jaccard, MSc
Media Engineering Institute
HES-SO University of Applied Sciences and Arts Western Switzerland
Av. Sports 20
Yverdon, 1400
Switzerland
Phone: 41 245577556
Email: dominique.jaccard@heig-vd.ch

Abstract

Background: Multidisciplinary collaboration is essential to the successful development of serious games, albeit difficult to achieve. In a previous study, the co.LAB serious game design framework was created to support collaboration within serious game multidisciplinary design teams. Its use has not yet been validated in a real usage context.

Objective: The objective of this study was to perform a first assessment of the impact of the co.LAB framework on collaboration within multidisciplinary teams during serious game design and development.

Methods: A mixed methods study was conducted, based on 2 serious game design projects in which the co.LAB framework was used. The first phase was qualitative and carried out using a general inductive approach. To this end, all members of the first serious game project team who used the co.LAB framework were invited to take part in a focus group session (n=6). In a second phase, results inferred from qualitative data were used to define a quantitative instrument (questionnaire) that was designed according to the Checklist for Reporting Results of Internet E-Surveys. Members of both project teams (n=11) were then asked to answer the questionnaire. Quantitative results were reported as median (Q1, Q3), and appropriate nonparametric tests were used to assess between-group differences. Finally, results gathered through the qualitative and quantitative phases were integrated.

Results: In both phases, the participation rate was 100% (6/6 and 11/11). Verbatim transcripts were classified into 4 high level themes: (1) influence on collaborative dimensions; (2) impact on project course, monitoring, and efficiency; (3) qualitative perceptions of the framework; and (4) influence of team composition on the use of the framework. The web-based questionnaire was then developed according to the 7 dimensions of collaboration by Burkhardt et al. In both projects, the co.LAB framework had a positive impact on most dimensions of collaboration during the multidisciplinary design and development of serious games. When all collaborative dimensions were aggregated, the overall impact of the framework was rated on a scale from –42 to 42 (very negative to very positive). The overall median score was 23 (Q1, Q3: 20, 27), with no significant difference between groups (P=.58). Most respondents also believed that all serious game design teams should include a member possessing significant expertise in serious game design to guide the development process.

Conclusions: The co.LAB framework had a positive impact on collaboration within serious game design and development teams. However, expert guidance seems necessary to maximize development efficiency. Whether such guidance can be provided by means of a collaborative web platform remains to be determined.

(JMIR Serious Games 2021;9(4):e33144) doi:10.2196/33144

KEYWORDS

serious game; educational game; education; simulation game; gaming; design; framework; methodology; mixed method; validation

XSL•FO
RenderX

Introduction

Background

Efficient multidisciplinary collaboration is essential to the successful development of serious games [1-4]. This collaboration can, however, prove difficult to achieve, as members of the development team come from different backgrounds and might have divergent expectations [4-7]. Communication and collaboration could be enhanced by using comprehensive design frameworks, the aim of which is to guide development teams during serious game design and development [4,7,8].

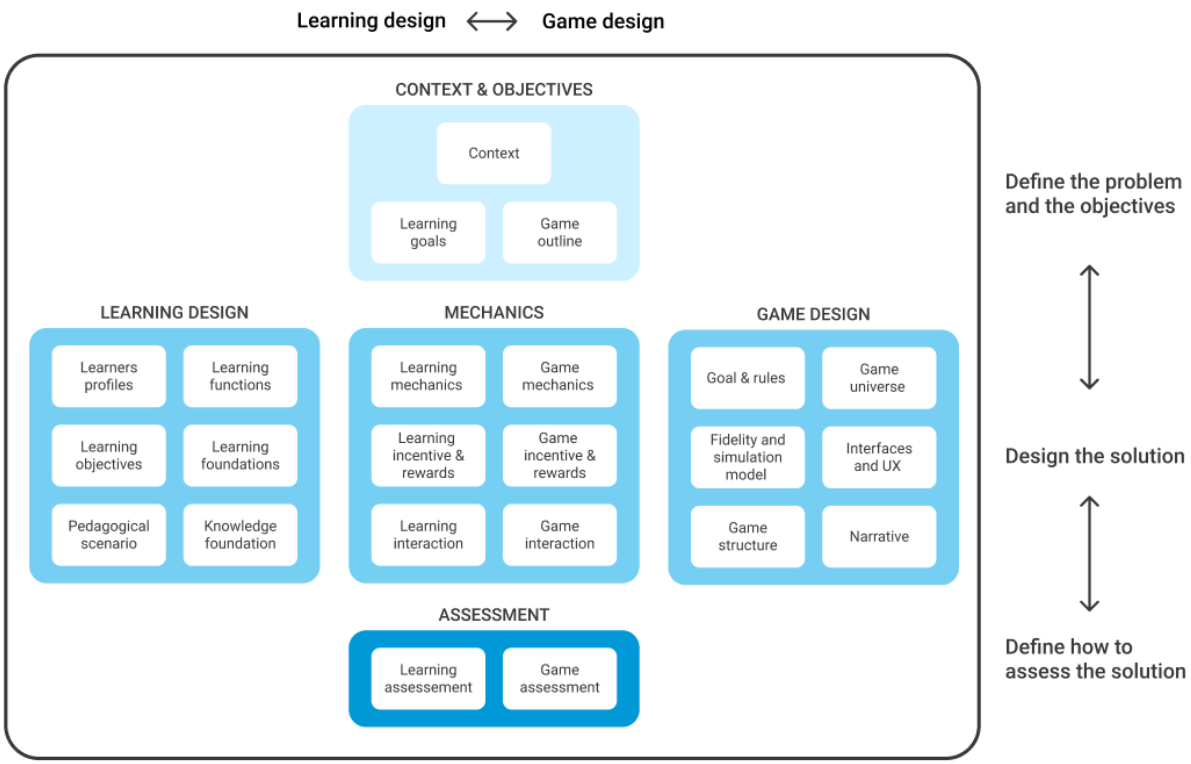
The co.LAB Methodological Framework

Funded by the Swiss National Science Foundation, the aim of the project "co.LAB - A Digital Lab for the co-Design, co-Development and co-Evaluation of Digital Learning Games" (co.LAB) is to improve efficiency and relevance in serious game design and development by supporting the collaboration between multidisciplinary development teams.

This project has already led to the creation and publication of the co.LAB generic serious game design framework (co.LAB framework) [9]. In their article, authors of the co.LAB framework stated that it was defined based on a literature review and on its authors' experiences. But the co.LAB framework has not been validated in a naturalistic context yet.

While most existing frameworks are dedicated to the design and development of specific types of educational games [4,10], the co.LAB framework was conceived as an adaptive framework that should allow the design of a wide range of educational games [9]. The co.LAB framework contains 5 main categories ("context and objectives," "game design," "mechanics," "learning design," and "assessment") and was designed to provide multidisciplinary teams with a global understanding of the design process. Different views of the framework can be used to help team members apprehend the different components of serious games design (Figure 1).

Figure 1. The co.LAB generic framework for serious game design (adapted from [9]). UX: user experience.



The horizontal view (either left-to-right or right-to-left) of the co.LAB framework emphasizes that serious game design is a blend of learning and game design, with the Mechanics category representing the link between the game and learning. In the vertical view, the co.LAB framework can be seen under a project management perspective: The upper section defines project objectives, the middle section defines the solution (the game and the associated learning concept), and the bottom section defines how the solution will be evaluated both from the game and learning perspectives [9].

To date, and to the best of our knowledge, the co.LAB framework has been used for serious game design and development by 2 multidisciplinary development teams.

Objectives

The objective of this study was to perform a first assessment of the impact of the co.LAB framework on collaboration within multidisciplinary teams during serious game design and development. The following research question was formulated: How does a multidisciplinary team perceive the contribution of the co.LAB framework to the collaborative design and development of serious games?

XSL•FO
RenderX

Methods

Overview

In order to better understand the contribution of the co.LAB framework to collaboration, mixed methods research with an exploratory design was conducted. The following phases were carried out sequentially [11]: (1) qualitative phase (focus group [FG]): qualitative data collection, data analysis, results; (2) mixing phase: development of an instrument to allow quantitative measurements; (3) quantitative phase (questionnaire): quantitative data collection, data analysis, results; (4) interpretation phase: integration of quantitative and qualitative results, generalization.

Projects and Participants

The study included team members of the first 2 projects in which the co.LAB framework was used.

The first project, Patients' Rights and Innovative Teaching Strategies (PRITS) [12], aimed at designing and developing a serious game to support health students in learning about patient rights. The project team included 2 professors of law (lawyers specializing in patients' rights), 1 health care lecturer, 1 educational researcher, 1 serious game expert, 1 graphic designer, and 2 computer scientists. The project was led by the School of Health Sciences (HESAV) and Media Engineering Institute (MEI), both at the University of Applied Sciences of Western Switzerland (HES-SO). The project began in September 2020 and ended in August 2021. It was thus almost completed at the time of this study (June 2021).

The second project, Interprofessional Major Incident Simulator (InterMIS), was a large-scale project aiming at developing a serious game to train health professionals to manage exceptional events. The project team includes 3 medical doctors, 2 paramedic instructors, 1 serious game expert, 1 graphic designer, and 2 computer scientists. This project is led by MEI at HES-SO and the Geneva University Hospitals. The project began in January 2021 and should span over 4 years. At the time of this study (June 2021), the project was at the end of the design phase, which included the design of both the learning and game aspects.

Experimental Conditions of the Uses of the co.LAB Framework

In both projects (PRITS and InterMIS), the co.LAB framework was used since inception. The framework was reproduced into an online shared folder. All team members had access to the overview of the framework and to their specific project folder. The use of the co.LAB framework for the design and development of both projects was chosen by the development leader. Therefore, participants had to assess the impact of a framework that they had not chosen and had not used before. Thus, an early adopter effect can hardly be held accountable for the results obtained in the course of this study.

Focus Group

In this first qualitative phase, a general inductive approach as described by Thomas [13] was chosen. This approach was considered appropriate in this context as it allows for the simple inference of findings from data derived from focused questions and discussions [13]. Moreover, FGs are particularly suitable for exploring people's knowledge and daily experiences [14,15] and for exploring ideas and opinions operating through communication [16].

Procedures

The FG was conducted according to the Consolidated Criteria for Reporting Qualitative Research (COREQ) guidelines [17] (Multimedia Appendix 1). An FG interview guide was established (Multimedia Appendix 2). The questions stated in this FG interview guide are only intended to give an overall idea of the questions actually asked during the FG. Indeed, as acknowledged by Krueger [18], questions are often asked in a somewhat different way than reported in such general guides.

An FG, which lasted 1.5 hours, was held at HESAV and recorded using an audio device. Participants (n=6) were asked to discuss their use of the co.LAB framework during their serious game design and development activities (individual and collective) and to explore its impact on collaboration.

Participants provided their consent to the recording and use of FG data for research purposes. The FG was held by an educational scientist with previous experience in conducting qualitative research and moderating such groups. One of the authors of the co.LAB framework introduced the FG and attended the FG as an external observer.

Participants

As one of the team members was the main author of the co.LAB framework, he did not act as a participant during the FG session but was nevertheless present as an observer. A second team member participated in the definition of the assessment of the contribution of the co.LAB framework to collaboration and acted as a facilitator rather than as a participant. Of the 6 members of the PRITS project who attended the FG, 3 were from the learning design and development side of the co.LAB framework (1 project owner and 2 professors of law), and 3 were from the game design and development side of the co.LAB framework (2 computer scientists and 1 graphic designer).

Data Analysis

A thematic analysis was performed on the data to identify themes. The thematic analysis was based on the process described by Braun and Clarke [19], with the steps outlined in Textbox 1.

XSL•FO
RenderX

Textbox 1. Steps followed during the thematic analysis to identify themes.

1. Transcription:
   - A "clean read" transcript was created based on the audio recording.
   - Based on the transcript, a summary of the discussion between participants was generated in connection with the questions.
2. Coding
   - Verbatim transcripts were coded with the qualitative analysis software MAXQDA (VERBI Software GmbH, Berlin, Germany).
3. Searching for themes
   - Themes were generated from codes.
4. Reviewing and defining themes
   - Themes were grouped into high-level themes.
5. Producing the report
   - A write-up of each theme content was performed, and representative quotes were selected.
   - As the focus group took place in French, we did an automatic English translation of selected quotes with DeepL Translate (DeepL GmbH, Köln, Germany). We chose to do an automatic translation as it guarantees the reproducibility and reduces the risks of interpretation by authors while translating. This automatic translation was proofread by an author and validated by a second author.

In addition, the team dynamic during the FG was analyzed, and agreements and disagreements between participants were identified and coded.

Each phase was initially carried out by one author, then confirmed by another. Along the process, authors in charge of the initial proposition were alternated. Any disagreement was resolved by reaching consensus.

Web-Based Questionnaire

Based on the FG results, a 3-tiered, web-based questionnaire was developed following Eysenbach's Checklist for Reporting Results of Internet E-Surveys (CHERRIES) [20] (Multimedia Appendix 3).

Questionnaire Design

The first part was designed to gather sociodemographic data, including prior serious game design experiences and roles held in these projects.

The second part focused on the contribution of the co.LAB framework to the collaboration. This part was constructed based on the 7 dimensions of collaboration proposed by Burkhardt et al [21]. The authors' original version was intended to assess the quality of the collaboration within a team. Since our goal was to evaluate the contribution of a methodology as a support to collaboration, we developed questions designed to specifically assess the contribution of the co.LAB framework. For each of the dimensions, we developed 3 questions, with one of them being formulated as a reversed item. The questions were proposed by a first author, completed by a second author, and validated by all authors.

In the third part, we added questions related to elements of collaboration and usability of the co.LAB framework that emerged from the FG analysis.

A translated version of this questionnaire is available (Multimedia Appendix 4).

Web-Based Platform

The questionnaire was hosted on a web platform created under the Joomla! 3.9 content management system (Open Source Matters, New York, NY) [22]. It was created using the Community Surveys 5.6 component (Corejoomla, Hyberabad, India) and administered to a convenience sample of 11 participants: 6 from the PRITS project and 5 from the InterMIS project. The web platform was protected by 2 different software firewalls: RSFirewall 3 (RSJoomla, Constanta, Romania) and Admin Tools 6 (Akeeba Ltd, Nicosia, Cyprus). To avoid potential double entries, participants were required to fill in a short registration form before they could access the survey. All data were automatically recorded and securely stored in an encrypted MySQL-compatible database (MariaDB 5.5.5; MariaDB Foundation, Wakefield, MA) hosted on a Swiss server.

Answers to multiple answer and multiple-choice questions were mandatory, and all questions had to be completed before participants were allowed to proceed to the next page. Answers could be modified until the questionnaire was marked as completed (after clicking the "Finish" button). Questionnaires could be resumed if participants were disconnected or chose to log out temporarily.

The platform, registration form, questionnaire, and data extraction mechanism were thoroughly tested by all co-authors prior to the quantitative phase. Only then were participants invited, by email, to complete the questionnaire. No incentive was provided to encourage participants to complete the questionnaire.

Data Extraction and Statistical Analysis

Data were extracted from the MariaDB database to a CSV file. It was then imported and curated under Stata 16.1 (Statacorp

XSL•FO
RenderX

LLC, College Station, TX). Given the limited sample size (n=11), data are presented as median (Q1, Q3) rather than as mean (SD). Questionnaires containing inconsistent answers were excluded. Inconsistent answers were detected when either the maximum (5) or minimum value (1) was given to all 3 questions assessing a specific dimension of collaboration as each dimension contained inversely phrased questions.

Answers based on 5-point Likert scales were then ascribed numerical values. Neutral answers (eg, "neither agree nor disagree") were given a value of 0, answers backing the use of the co.LAB framework were given positive values (either 1 or 2), and answers opposing it were given negative values (either –1 or –2).

Each question regarding the 7 dimensions of collaboration was first analyzed separately. Then, the 3 questions belonging to each specific dimension were grouped, thereby generating a score ranging from –6 to 6. Finally, all 21 questions assessing the 7 dimensions by Burkhardt et al [21] were pooled to give an overall representation of the framework's impact on a score ranging from –42 to 42. All questions were evenly weighted.

The answers to Likert-based questions assessing other elements of collaboration were assigned points ranging from 1 (for "Very small/none") to 5 (for "Very high").

Nonparametric tests were used to assess differences between the InterMIS and PRITS groups. Individual questions were assessed using Fisher exact tests, and the Mann-Whitney U test was used to evaluate pooled results (overall score and scores by dimension). A P value <.05 was considered significant.

Consent and Institutional Review Board

All participants were provided with project information and gave their consent to participate and to the use of data for research purposes.

According to the Swiss Human Research Act [23], institutional review board approval was not necessary for this study, as participants did not belong to a vulnerable population and there was no medical intervention whatsoever.

Results

Focus Group

The classification of the verbatim transcripts enabled us to cluster the discussions about the contribution of the co.LAB framework into 4 high-level themes. These high-level themes, along with their subthemes, can be seen in Table 1.

Table 1. Themes and subthemes inferred from the focus group discussions.

High-level themes	Subthemes
Influence on collaborative dimensions	Fluidity of collaboration, information exchange for problem solving, argumentation and reaching consensus, sustaining mutual understanding, individual task orientation, task and time management
Impact on project course, monitoring, and efficiency	Use at all stages of the project, the project's common thread, role clarification, monitoring, common objectives, structuring, outcomes related to the framework
Qualitative perceptions of the framework	Positive perceptions, negative perceptions
Influence of team composition on the use of the framework	Team size, team experiences, team quality and diversity, team autonomy in using the framework

The next sections provide a summary of each theme, with a few quotes selected for their representativeness. The final section presents the results of the team dynamics analysis.

Influence on Collaborative Dimensions

The co.LAB framework positively influenced 6 of the 7 dimensions of collaboration proposed by Burkhardt et al[21]. Participants found that the fluidity of collaboration was supported by this framework, as each member knew both what to do and what the others were to do. Consensus was more easily reached, as each member was able to react to what had been written or proposed by others (this was achieved by using shared files embedding collaborative features). Information sharing was promoted by granting access to centralized and structured information. Personal motivation was bolstered by making it possible to monitor the progress of the project. Mutual understanding was strengthened by a clear definition of the roles and using a common terminology. Finally, management of time and activities was facilitated by allowing an almost permanent and real-time overview of the project progress and of the tasks that still had to be performed.

Regarding fluidity of collaboration, one participant emphasized what was previously said by others, as follows:

> *The method helps with the fluidity of collaboration because you know who is doing what; it makes things clearer for everyone. The fluidity of collaboration is what impressed me the most for the reasons already mentioned.* [Designer]

The co.LAB framework was implemented in a shared online environment where team members could find the information. A participant who joined the development team while the project was already underway expressed how the framework had helped her assimilate information previously shared by and between team members:

> *I came in during the course of the project, so it allowed me to get into the flow of the previous information. When you come into a project, it's always extremely difficult to know what has been discussed. I appreciated having a place where everything was brought together in a clear way.* [Developer 2]

Another participant acknowledged that "we are forced to reach consensus" (Developer 1). This participant considered that the

XSL•FO
RenderX

framework reproduced in shared files helped team members to find agreement because it allowed them to see what others had written and gave them the possibility to react accordingly.

A participant pointed out that a common understanding was promoted by the common terminology established by virtue of the framework:

> *For mutual understanding, this method allows us to have common terms, to know what we are talking about and where we stand.* [Professor of law 1]

Mutual understanding was also strengthened by clarifying the respective roles while setting common objectives:

> *It's allowed for common goals, without getting lost, to guide our work. Especially when there are a lot of us, coming from different fields.* [Professor of law 1]

The framework also allowed people from different disciplines to give their opinion and to take notice of those from others:

> *In a project with multiple disciplines, it helped to know when we had a voice. It's not forbidden to speak, but we knew when we had more legitimacy, on which points of co.LAB we could contribute more, and when we should stay in the background, leave the place to the person who has the competence, rather than having everyone throw in their two cents, not necessarily at the best time.* [Developer 1]

The framework was seen as having the potential to positively influence motivation:

> *For motivation, I think it helps indirectly. It wasn't the case, but if I had had a drop in motivation, I think that going back to the framework and seeing the progress, what's going on in green, would have helped me get motivated again.* [Designer]

The methodology also positively influenced fluid and spontaneous task allocation between project members:

> *As the main applicant, I was concerned about making sure we got to the end of the project, that all the tasks were done. But I didn't have to impose anything. People assigned themselves the tasks naturally, and my role was comfortable.* [Project Owner]

Impact on Project Course, Monitoring, and Efficiency

Using the co.LAB framework was perceived as having a positive impact on the course of the project. It was used from inception and all along the project. Throughout the project, it was perceived as a guide, a common thread. In that way, the co.LAB framework was considered as a supporting structure for the design and development process and for project monitoring. The use of the co.LAB framework was also perceived as having a positive impact on efficiency. Participants shared their thoughts on using this methodology in other contexts, for example without an expert or in another team configuration, but there were no responses showing a negative impact of the framework on collaborative work.

All participants answered that the framework was used as soon as they joined the project and then throughout it. A participant said that she had even already used it for a project proposal:

> *We used it from the very beginning and already in the preparation of the project proposal.* [Professor of law 2]

Several participants agreed about the use of co.LAB framework throughout the project and its role as a guideline and project monitoring tool:

> *If I remember correctly, each meeting included the milestones of the methodology. It gave us a roadmap to follow. It really made an impact because it structured our meetings, allowing us to know what we had to pay attention to for the next sessions. Beyond the tools, the organizational aspect of the methodology was very comfortable. It gave a vision of the next steps.* [Project Owner]

The co.LAB framework was seen as enabling people to better understand what to do, in a consistent way across disciplines:

> *On the development side, it helped to know what mechanics to put in place. I wasn't present at every meeting, and going to co.LAB was a way to know if we needed to implement a button or something else. It was a way to get direction, not to do something that was not discussed, not to do something that wasn't planned: not to do too much, but to do the right thing, to refocus.* [Developer 1]

Team members frequently used the word "structure," which illustrates their perception of the framework as assembling parts of a whole. The project was considered to have been carried out effectively by virtue of the structuring effect and of the centralization of information provided by the co.LAB framework:

> *It was a structure that made it possible to work quickly and well.* [Developer 1]

> *The structure is nice. With the centralized information, I could go and look for what had to do with IT development. Nothing bothered me; on the contrary, I could focus on the points that were important to me. Sometimes there were meetings where I felt less concerned, but in the end it's not such a bad thing; it helps to understand the project. Without the help of the co.LAB method, we might not have had this efficiency.* [Developer 1]

The project owner raised a question about the context of use of the framework in larger teams:

> *It's more of a question: We only had 4 disciplines and a small group (8 people). Can we transpose this methodology to a larger group with more disciplines?* [Project Owner]

Qualitative Perceptions of the Framework

Team members were asked to discuss positive and negative characteristics of the co.LAB framework. The discussion showed a clear predominance of positive aspects, with terms such as comfortable, clear, reassuring, liked, user-friendly, easy, understandable, and pleasant. These positive aspects are linked to most of the themes already presented: information sharing, project monitoring, structure, or guidelines. The only negative

XSL•FO
RenderX

aspect was the fact that the content was mainly introduced and synthesized by one person, who guided the project.

Usability and simplicity were seen as factors favoring the adoption of the framework:

> *I liked it because it's a user-friendly tool. It's simple to use, logical to follow. This ease of use allows the method to be adhered to.* [Project Owner]

It was perceived as a convenient source of information, both for internal and external uses:

> *It's quite comfortable to bring out information among others for our funders. It has been an essential source of information for me.* [Project Owner]

One team member reported positive aspects of the framework, as follows:

> *The method is very clear. And there was availability, a nice collaboration, and I think the method helps in that.* [Professor of law 1]

On the negative side, a participant said that she was involved in creating the content but had limited participation at other levels:

> *On the negative aspects, I did not complete the content. We were involved in the content but not in proposing the steps.* [Professor of law 2]

Influence of Team Composition on the Use of the Framework

All team members described themselves as having no previous experience in serious game design nor had they heard about the co.LAB framework before starting the project. As such, they pointed out the value of having a team member (who did not participate in the FG) with expertise in serious game design and who had extensive knowledge of the co.LAB framework. The presence and role of this expert were seen as a success factor for the use of the framework. It raised a question about the importance of having an expert and whether the framework could be used without guidance.

Regarding the benefits of having an expert of the framework among the team, a professor of law said:

> *It made it easier to have a person responsible [the expert] for filling in the items. Someone who has the logic in mind, who knows what to write in which area, and then to discuss it with the people involved in the project. I'm not sure I'm able to fill in the right items.* [Professor of law 1]

Although the methodology was seen as more easily usable if an expert was present, some team members thought that it could nevertheless serve as a guideline for teams with no prior experience in serious game development:

> *I was involved, but a lot of it was handled by [the expert]. I don't know if I would have been able to do this. But it would have been a good base, a checklist.* [Professor of law 2]

> *I especially have questions: Without the coaching we received, would we have been able to follow the method with the same efficiency? I think that it is a very interesting tool, but there is still the need to have someone who translates the issues behind the method.* [Project Owner]

Some team members said that, in future projects, while gaining experience, they would be interested in using the co.LAB framework more independently:

> *In future projects, I would be really interested in doing it individually. But I'd also be interested in having someone behind me telling me if I did it right or wrong. I don't know if I would be able to, but when I see the methodology, it looks simple; it structures a project. I'd be interested in trying to use it independently.* [Developer 1]

Team Dynamics During the Focus Group

While speaking about the use of the co.LAB framework for teamwork and collaboration, team members mostly agreed with each other. The agreement with what was already said was sometimes made explicit with the use of expressions like “I agree with…” and “I’m joining what was said.” Even though it was not explicit at other moments, content analysis still showed convergent opinions. Content similarity was stronger among participants who had the same functions in the team, while the project owner expressed several specific points of view.

There was no explicit disagreement between team members during the discussion. Team members never confronted each other directly but different points of view were expressed, particularly regarding the personal use of the framework and how they felt about using it in autonomy for further projects.

Participants raised some questions about the use of the framework in other team configurations and wondered whether its use would also be appropriate for larger development teams.

Questionnaires

We sent 11 emails to members of development teams belonging to either of the 2 projects, all of whom completed the questionnaire (100%). They completed the questionnaire between June 29, 2021 and July 7, 2021. No questionnaire was excluded as our exclusion criterion was not met. Participants' characteristics are detailed in Table 2.

XSL•FO
RenderX

Table 2. Characteristics of the study participants.

Characteristics	PRITS[a] project (n=6)	InterMIS[b] project (n=5)
Age (years), median (Q1, Q3)	33 (33, 38)	38 (36, 38)
Gender, n		
Female	4	1
Male	2	4
Other	0	0
Role(s) in the project[c], n		
Computer scientist	2	1
Game designer	2	0
Graphic designer	1	0
Pedagogue	0	1
Serious game expert	0	1
Subject matter expert	1	4
Teacher	3	0
Level of experience in serious game development prior to project inception, n		
None	5	4
Limited	1	0
Intermediate	0	0
High	0	1

[a]PRITS: Patients' Rights and Innovative Teaching Strategies.

[b]InterMIS: Interprofessional Major Incident Simulator.

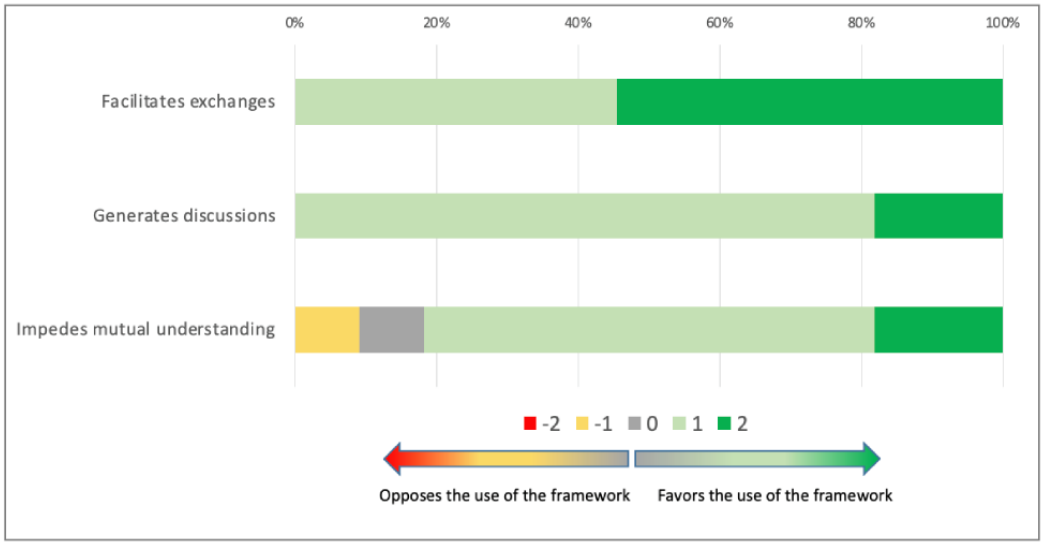
[c]Participants could have more than one role.

The following sections present the results of the questionnaire regarding the influence of the co.LAB Framework on the collaborative dimensions identified by Burkhardt et al [21] and then on the other specific items inferred from the FG analysis.

Fluidity of Collaboration

According to the participants, the co.LAB framework facilitates exchange between specialists from different disciplines (median 2; Q1, Q3: 1, 2), helps generate discussion between these specialists (median 1; Q1, Q3: 1, 1), and does not impede mutual understanding (median 1; Q1, Q3: 1, 1). A graphical representation of these results is displayed in Figure 2.

Figure 2. Assessment of fluidity of collaboration with the co.LAB framework.



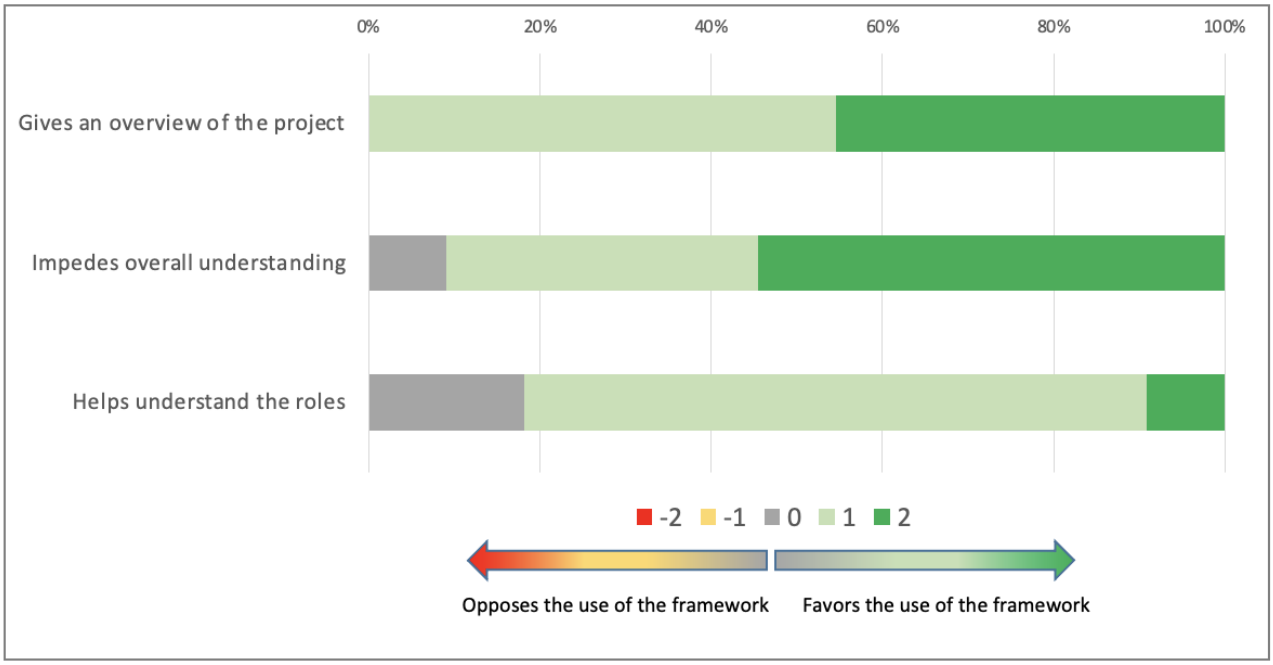
XSL•FO
RenderX

The overall assessment of this dimension was in favor of the co.LAB framework (median 4; Q1, Q3: 3, 5), with no significant difference between groups (*P*=.06).

Sustaining Mutual Understanding

The co.LAB framework was considered to provide an overall view of the project (median 1; Q1, Q3: 1, 2). It also helped understand the roles of the different team members (median 1; Q1, Q3: 1, 1) and did not make overall understanding difficult (median 2; Q1, Q3: 1, 2). These results are shown in Figure 3.

Figure 3. Sustaining mutual understanding with the co.LAB framework.

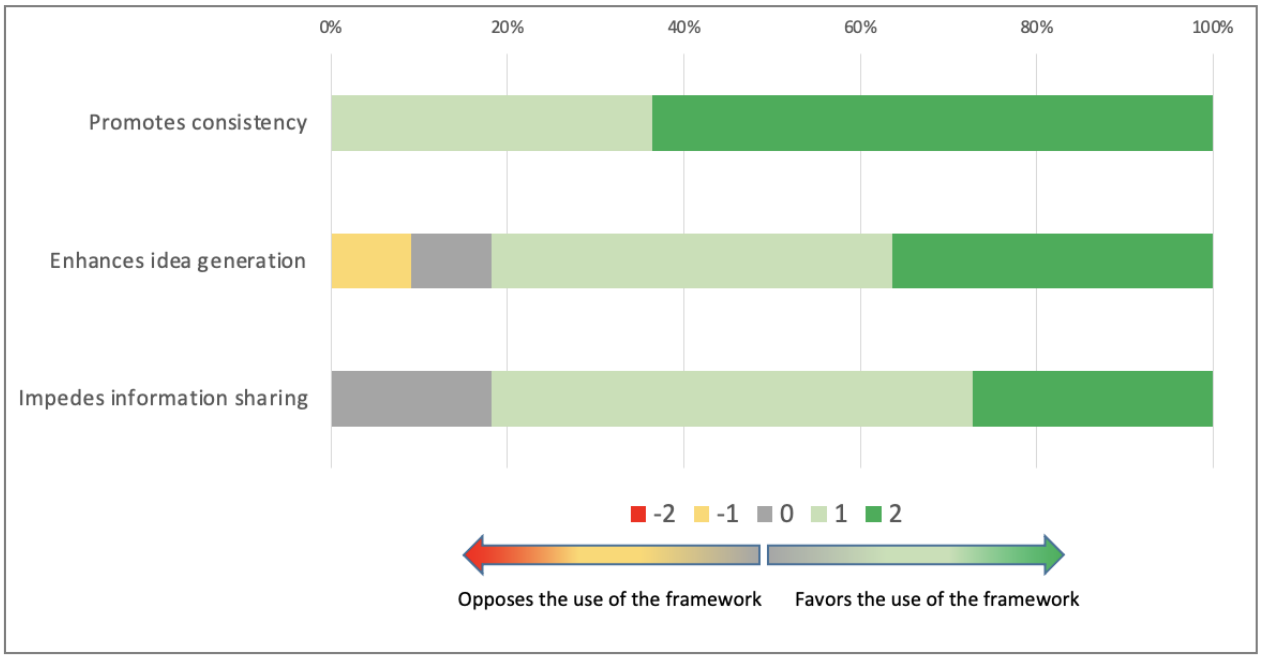


Overall, this dimension was clearly supported by the use of the co.LAB framework (median 4; Q1, Q3: 3, 5). There was no statistically significant difference between the PRITS and InterMIS groups (*P*=.64).

Information Exchange for Problem Solving

The participants answered that using the co.LAB framework promoted consistency in the collaborative search for solutions (median 2; Q1, Q3: 1, 2) and enhanced idea generation (median 1; Q1, Q3: 1, 2). It was not considered to make information sharing more difficult (median 1; Q1, Q3: 1, 2; Figure 4).

Figure 4. Information exchange for problem solving using the co.LAB framework.



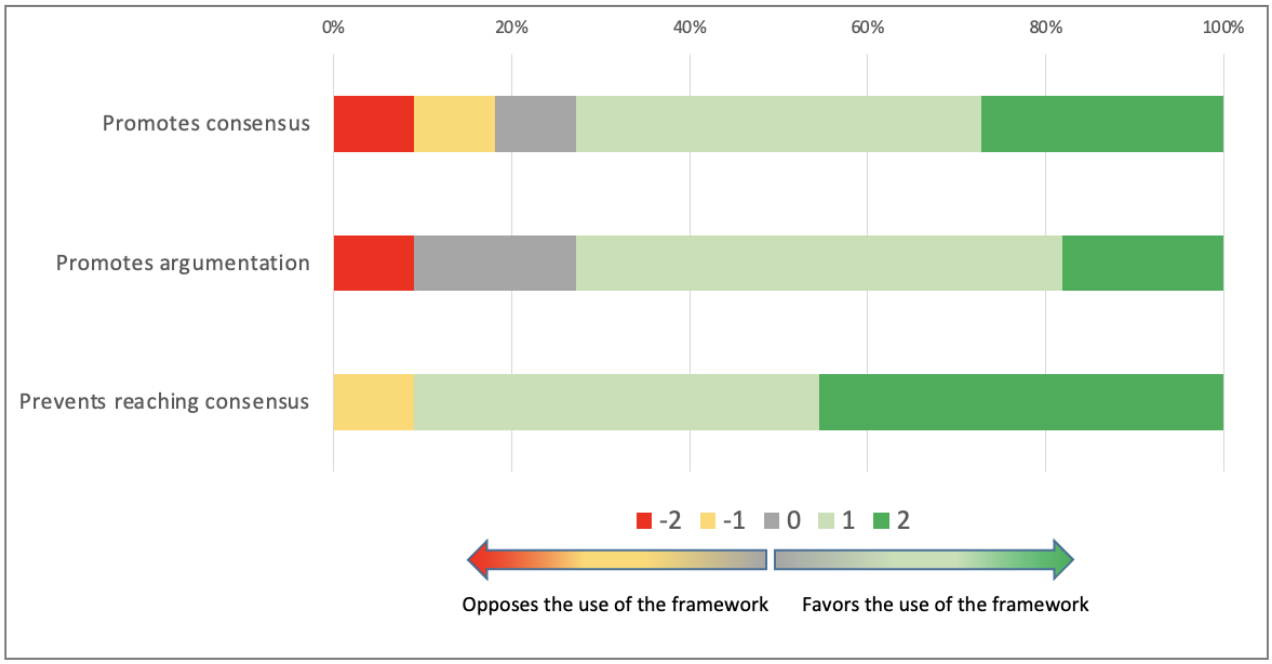
XSL•FO
RenderX

Overall information exchange was therefore facilitated by the use of the co.LAB framework (median 4; Q1, Q3: 2, 5). PRITS and InterMIS participants provided similar answers (*P*=.52).

Argumentation and Reaching Consensus

Consensus building was promoted by the use of the co.LAB framework (median 1; Q1, Q3: 0, 2). This framework was also seen as promoting argumentation on alternative solutions (median 1; Q1, Q3: 0, 1) and was not considered as preventing reaching consensus (median 1; Q1, Q3: 1, 2; Figure 5).

Figure 5. Argumentation and reaching consensus using the co.LAB framework.



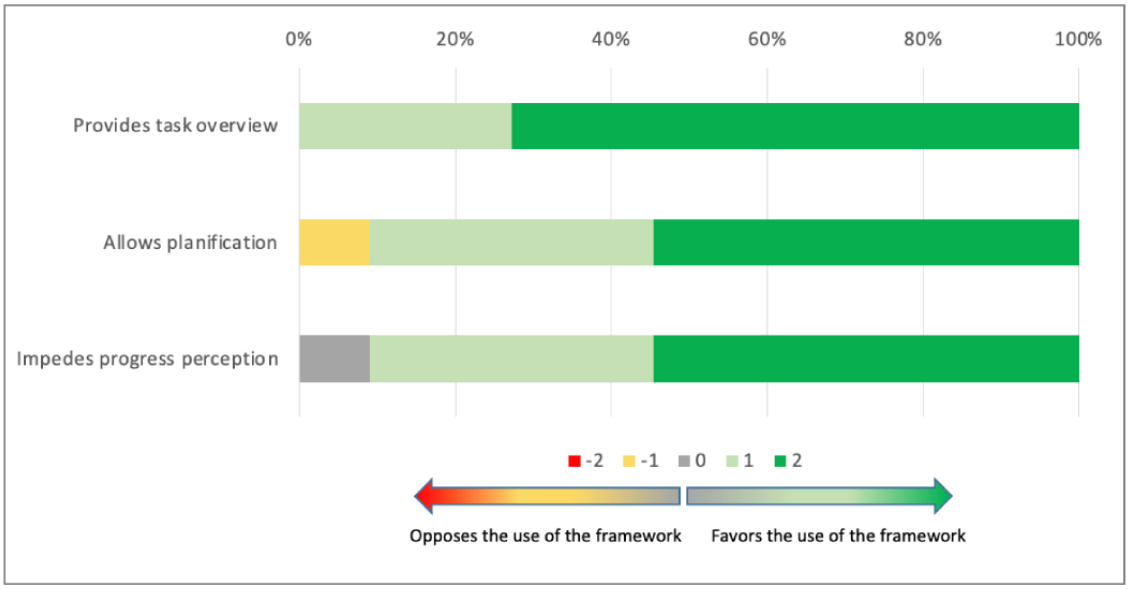
Overall, this dimension was also supported by the use of the co.LAB framework (median 4; Q1, Q3: 2, 4) with no difference between groups (*P*=.70).

Task and Time Management

Participants generally agreed that the co.LAB framework was effective at providing an overview of the work to be achieved (median 2; Q1, Q3: 1, 2) and allowed planning of tasks (median 2; Q1, Q3: 1, 2). They did not think that its use made it difficult to understand the progress of the project (median 2; Q1, Q3: 1, 2; Figure 6).

Overall, this dimension received high ratings (median 5; Q1, Q3: 3, 6), with no difference between groups (*P*=.40).

Figure 6. Task and time management using the co.LAB framework.

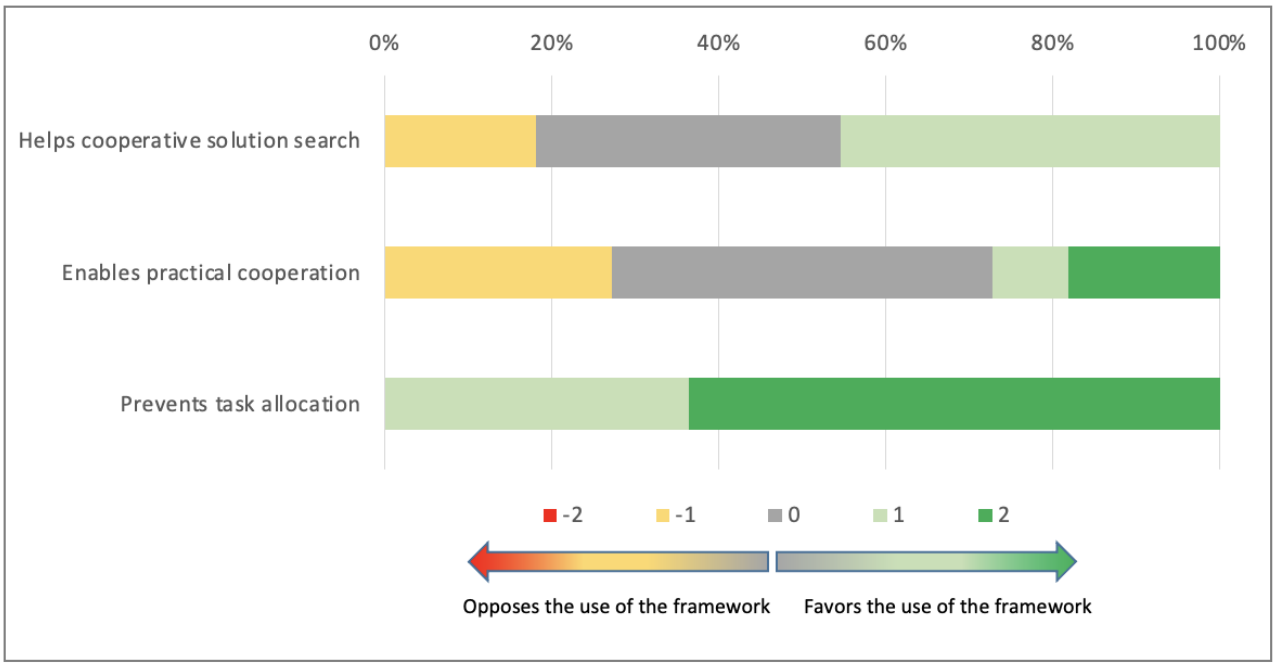


XSL•FO
RenderX

Cooperative Orientation

Participants were less convinced by the usefulness of the co.LAB framework regarding the promotion of equal contributions in the search for a solution (median 0; Q1, Q3: 0, 1) or in achieving solutions (median 0; Q1, Q3: –1, 1). They were, however, convinced that using this framework did not interfere with task distribution (median 2; Q1, Q3: 1, 2; Figure 7).

Figure 7. Cooperative orientation using the co.LAB framework.

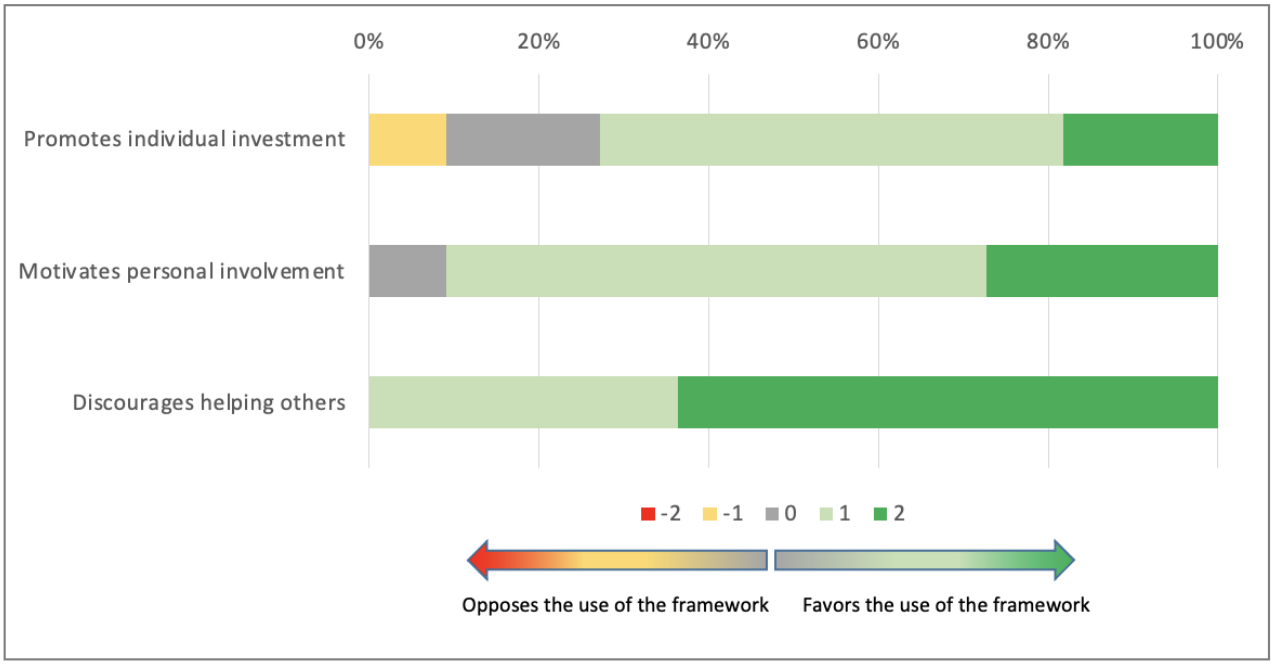


The overall analysis of this dimension nevertheless favored the use of the co.LAB framework (median 2; Q1, Q3: 1, 4), with no statistically significant difference between groups (P=.09).

Individual Task Orientation

The co.LAB framework was considered as efficient in promoting individual investment throughout the project (median 1; Q1, Q3: 0, 1) and in motivating personal involvement (median 1; Q1, Q3: 1, 2). Participants who used the framework were not less prone to help others (median 2; Q1, Q3: 1, 2; Figure 8).

Figure 8. Individual task orientation using the co.LAB framework.



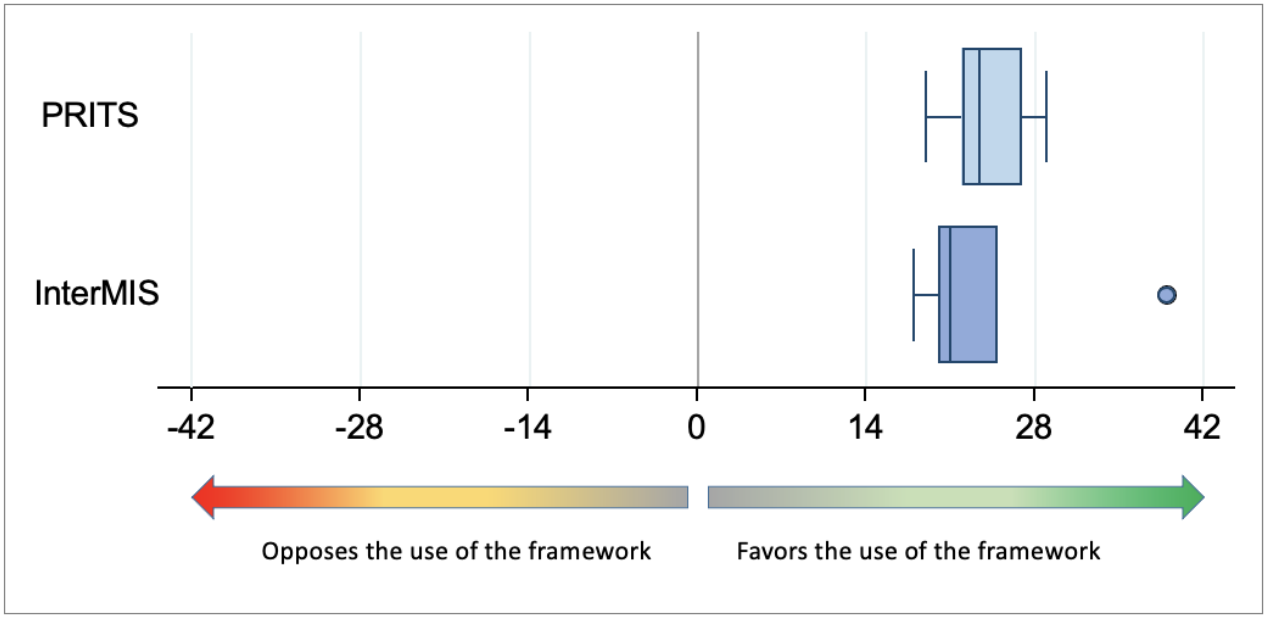
XSL•FO
RenderX

The overall analysis of this last dimension also supported the use of the co.LAB framework (median 3; Q1, Q3: 2, 5). There was no difference between groups (*P*=.08).

Overall Assessment

Figure 9 shows the results obtained after pooling the 21 questions used to assess the 7 collaborative dimensions by Burkhardt et al [21].

Figure 9. Overall assessment of the 7 dimensions of collaboration. InterMIS: Interprofessional Major Incident Simulator; PRITS: Patients' Rights and Innovative Teaching Strategies.



No statistically significant difference was detected (*P*=.58). Figure 9 shows that there was a clear outlier in the InterMIS group, who rated the co.LAB framework much higher than all other participants. This outlier was a subject matter expert who had no prior experience in serious game development.

Collaboration Items Specific to the co.LAB Framework

All participants agreed that the co.LAB framework represents a guide for the design and development of serious games (median 2; Q1, Q3: 1, 2), with no difference between groups (*P*=.99). One participant commented that any comparison would be difficult without knowing any other framework or method. Another highlighted the fact that the framework "does not give the solution but structures the different steps necessary to the development [of the serious game]." The 2 participants who had prior experience in serious game design and development said that they had not used a specific framework or method during their previous venture. Both rated their experience with the co.LAB framework as either positive or very positive in comparison with their prior development.

Most participants thought that, for a team without prior experience in serious game design, a serious game expert would be mandatory to use the framework (median 2; Q1, Q3: 1, 2). A participant commented that "it [was] mandatory" while another was of the opinion that "it might not be mandatory, but it certainly has an added value." In line with this latter comment, participants were not convinced that the co.LAB framework could serve as a standalone guide for a team without prior experience and without an expert (median 0; Q1, Q3: –1, 1). However, all participants were convinced that having a person responsible for summarizing the information regarding the project and its progress was necessary regardless of the experience of the development team (median 2; Q1, Q3: 1, 2).

The effect of the co.LAB framework on the quality of the collaboration during the design and development of the serious game was highly rated (median 4; Q1, Q3: 4, 4), with no difference between groups (*P*=.99). Participants rated even higher the impact of personal characteristics (personalities, past experiences) on the quality of the collaboration during the design and development phase (median 5; Q1, Q3: 4, 5). Both groups gave similar ratings to this item (*P*=.99).

Finally, all participants were convinced that the co-LAB framework should be implemented in a collaborative web platform (median 4; Q1, Q3: 4, 5), with no significant difference between groups (*P*=.24).

Discussion

Principal Findings

Essential Findings

The co.LAB framework was conceived to facilitate the design and development of serious games through a collaborative, multidisciplinary, adaptive, systemic approach [9]. The main objective of the present study was to gain insight regarding the contribution of the co.LAB framework to the collaboration within multidisciplinary teams during serious game design and development. Overall results (FG and questionnaires) show that the co.LAB framework had a positive impact on collaboration within multidisciplinary teams during serious game design and development.

XSL•FO
RenderX

Impact on Collaboration

The qualitative analysis of the FG session shows that the co.LAB framework was perceived as having a positive influence on several dimensions of collaboration. Participants spontaneously noted the positive impact of the co.LAB framework on the fluidity of communication, mutual understanding, information sharing, argumentation, motivation, management of project activities, and monitoring. This positive influence was similarly noted by participants from different professional disciplines. One dimension was not addressed: the fair balance of verbal contributions and activities carried out by the actors. While participants shared the impression that the co.LAB framework enhanced multidisciplinary collaboration, they also felt that it helped them manage their own specific tasks. These were related either to the discipline or to their role (eg, the game developer using information for the design of interfaces or the project owner seeking information about the project progresses for the funders). Thus, we can assume that the co.LAB framework could simultaneously support and build a shared understanding, while providing individual team members with specific information necessary to carry out their activities. These findings reveal that the use of this framework was positively perceived by team members. Moreover, it allows them to overcome some difficulties such as mutual understanding and working with individuals from different fields, which have been described as simultaneously a necessary and challenging process [4].

The quantitative analysis yielded results consistent with those obtained through qualitative analysis. The use of the co.LAB framework was perceived to have a positive impact on all dimensions of the collaboration. The dimension that was the least impacted by the use of the framework was facilitation of equal contributions, both in the search for solutions and in the development of those solutions. This is consistent with the FG results, as this dimension was the only one not mentioned by the participants during the session. We hypothesize that the rather low impact of the co.LAB framework on the equality of contribution could be explained by the fact that the framework seeks to give an overview of the solution to be developed but does not give any guidelines on the distribution of contributions. Moreover, we believe that the design and development of a serious game by a multidisciplinary team does not necessarily imply a proportional distribution of contributions. For example, a computer scientist may have to work several months on design and development, but a professional expert may spend only 1 day on the validation of knowledge foundations.

Collaboration has previously been described as a challenging but necessary process for serious game conception [4]. Overall qualitative and quantitative results of this study support the hypothesis that the use of a visual design framework, such as the co.LAB framework, provides an overview that enables both mutual understanding and the collaborative development of an integrated and coherent solution. In addition, a positive impact on time and task management was also reported. Thus, overall results suggest that the co.LAB framework helps overcome some of the inherent challenges linked with collaboration during serious game design and development.

Need for a Serious Game Expert

This study also revealed some additional points of interest regarding contingency factors influencing the use of the co.LAB framework. During the FG, an unexpected element was mentioned by participants: All of them agreed on the need to have a serious game development expert who understands the methodology and can guide the team. The results of the web-based questionnaire, which contained 3 questions designed to specifically assess this element, are consistent with this finding. There was no clear agreement on the possible use of the methodology without resorting to such an expert.

We hypothesize that this can be explained in the following way: The co.LAB framework defines “building blocks” (eg, learning objectives, pedagogical scenarios, game mechanics) that need to be designed and developed. Disciplinary skills are necessary for each of those domains (for example, a user experience [UX]/user interface [UI] specialist is needed to develop user interfaces). In the same way, the co.LAB framework provides guidelines for overall serious game design as these elements are not independent but interconnected [9]. Therefore, disciplinary skills in serious game design are needed to understand the interrelation between design elements and encourage coherence between them. This can also explain why the majority of team members, experts in their field but novices in serious game design, felt comfortable using the co.LAB framework in their field of expertise but not in all areas of serious game design. However, we believe that the co.LAB framework could allow novices to progressively develop expertise in serious game design and a better understanding of the other dimensions at stake. From a time perspective, some members expressed interest in developing serious game design expertise.

Therefore, we argue that a serious game design team should include disciplinary skills needed to cover each element of the serious game design and address the need for competencies related to the overall design and development process.

Limitations

The first limitation of our study is related to its limited sample size, despite a very high participation rate. This could hardly be helped as, to date, only 2 projects have been developed with the methodology supported by the co.LAB framework. In addition, at the time of this study, both projects were going well in terms of both deadlines and quality of the results obtained. This may have had an impact on the perception of the contribution of the co.LAB framework but could also be seen as an effect of the use of the framework.

Furthermore, there are methodological limitations related to the use of FGs in qualitative research. Some of the recognized limitations for this type of method are the fact that some people speak more while others remain in the background (although this should have been counterbalanced by the moderation), and alignment of ideas, status, and roles can influence the opinions of the participants. To mitigate these limitations, some decisions were made. As one of the team members was the main author of the co.LAB framework, he did not participate as a participant in the FG but as an observer. As a second team member

XSL•FO
RenderX

contributed to this study (ie, participated in the definition of the assessment of the contribution of the co.LAB framework to collaboration), she did not act as a participant but as a facilitator. Since she had a good understanding of the subject at hand, she was able to conduct the FG with valuable prior knowledge. Her experience as a researcher in qualitative research was another advantage in facilitating the FG.

Another limitation is related to the context of use of the co.LAB framework. In both projects, an expert of the co.LAB framework was part of the serious game development team. Therefore, this study does not give information regarding the possible use of the co.LAB framework without a methodological expert but rather shows that it can be useful for novices provided they receive appropriate guidance. It is not possible to state whether the importance of the presence of the expert was dependent on the framework itself or on team configuration, as in both projects, the same expert was part of the team. Further investigation is needed to measure the impact of this factor on team collaboration.

Finally, even though the questionnaire administered to participants was developed on the basis of the questionnaire by Burkhardt et al [21], it has not been validated. Given the small sample size, we refrained from performing a reliability analysis, and future studies including larger samples should be used for this purpose.

Future Work

The co.LAB framework is currently being implemented in a web platform in which it is planned to embed questionnaires to evaluate its contribution to collaboration. This will allow for larger-scale validation.

Beyond assessing the impact of the framework on collaboration, this study revealed the perceived importance to the team of having a serious game expert. The analysis of the expert's contribution versus the framework's contribution, as well as the influence of the expert's personality, are topics for future research.

Conclusion

Qualitative and quantitative results obtained through this study support the use of the co.LAB framework to facilitate collaboration within multidisciplinary serious game design and development teams.

In both projects included in this study, the co.LAB framework enhanced several dimensions of collaboration within multidisciplinary teams. However, expert guidance seems necessary to maximize development efficiency. Whether such guidance can be provided by means of a collaborative web platform remains to be determined.

Acknowledgments

The co.LAB project is supported by the Swiss National Science Foundation (NRP 77), under Grant number 407740_187275.

The InterMIS project is co-funded by the Geneva University Hospital and InnoSuisse.

The PRITS project was supported by the digital skills center of the HES-SO under grant PGBDS8-104294.

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ: Consolidated Criteria for Reporting Qualitative Research.
[XLSX File (Microsoft Excel File), 15 KB - games_v9i4e33144_app1.xlsx]

Multimedia Appendix 2
Focus group interview guide translated.
[PDF File (Adobe PDF File), 56 KB - games_v9i4e33144_app2.pdf]

Multimedia Appendix 3
Cherries checklist.
[PDF File (Adobe PDF File), 120 KB - games_v9i4e33144_app3.pdf]

Multimedia Appendix 4
Web-based questionnaire translated.
[PDF File (Adobe PDF File), 30 KB - games_v9i4e33144_app4.pdf]

References

1. Bourgonjon J, De Grove F, De Smet C, Van Looy J, Soetaert R, Valcke M. Acceptance of game-based learning by secondary school teachers. Computers & Education 2013 Sep;67:21-35. [doi: 10.1016/j.compedu.2013.02.010]

XSL•FO
RenderX

2. Chaudy Y, Connolly T, Hainey T. EngAGe: A Link between Educational Games Developers and Educators. 2015 Presented at: 6th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES); September 9-12, 2014; Valletta, Malta URL: https://ieeexplore.ieee.org/document/7012156 [doi: 10.1109/vs-games.2014.7012156]
3. Djaouti D. Serious Games pour l'éducation : utiliser, créer, faire créer ? trema 2016 May 01(44):51-64. [doi: 10.4000/trema.3386]
4. Verschueren S, Buffel C, Vander Stichele G. Developing theory-driven, evidence-based serious games for health: Framework based on research community insights. JMIR Serious Games 2019 May 02;7(2):e11565. [doi: 10.2196/11565]
5. De Troyer O. Towards effective serious games. 2017 Presented at: 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games); September 6-8, 2017; Athens, Greece URL: https://ieeexplore.ieee.org/document/8056615 [doi: 10.1109/vs-games.2017.8056615]
6. De Troyer O, Janssens E. Supporting the requirement analysis phase for the development of serious games for children. International Journal of Child-Computer Interaction 2014 May;2(2):76-84. [doi: 10.1016/j.ijcci.2014.05.001]
7. Rooney P. A theoretical framework for serious game design: Exploring pedagogy, play and fidelity and their implications for the design process. International Journal of Game-Based Learning 2012;2(4):41-60. [doi: 10.4018/ijgbl.2012100103]
8. Aleven V, Myers E, Easterday M, Ogan A. Toward a Framework for the Analysis and Design of Educational Games. 2010 Presented at: Third IEEE International Conference on Digital Game and Intelligent Toy Enhanced Learning; April 12-16, 2010; Kaohsiung, Taiwan. [doi: 10.1109/digitel.2010.55]
9. Jaccard D, Suppan L, Sanchez E, Huguenin A, Laurent M. The co.LAB generic framework for collaborative design of serious games: development study. JMIR Serious Games 2021 Jul 02;9(3):e28674 [FREE Full text] [doi: 10.2196/28674] [Medline: 34255655]
10. Plass JL, Mayer RE, Homer BD. Handbook of Game-Based Learning. Cambridge, MA: MIT Press; 2020.
11. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015 Nov;13(6):554-561 [FREE Full text] [doi: 10.1370/afm.1865] [Medline: 26553895]
12. De Oliveria DC, Bielser F, Bonnard D, Songuel Y, Jaccard D. A Serious Game for Patient' Rights Education. 2021 Presented at: IEEE 9th International Conference on Serious Games and Applications for Health (SeGAH); August 4-6, 2021; Dubai, United Arab Emirates URL: https://ieeexplore.ieee.org/document/9551898 [doi: 10.1109/segah52098.2021.9551898]
13. Thomas D. A general inductive approach for analyzing qualitative evaluation data. American Journal of Evaluation 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
14. Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995 Jul 29;311(7000):299-302 [FREE Full text] [doi: 10.1136/bmj.311.7000.299] [Medline: 7633241]
15. Kitzinger J, Markova I, Kalampalikis N. Qu'est-ce que les focus groups? Bulletin de psychologie, Groupe d'étude de psychologie 2004;57(3):237-243 [FREE Full text]
16. Moscovici S. The Phenomenon of Social Representations. In: Farr RM, Moscovici S, editors. Social Representations. Cambridge, MA: Cambridge University; 1984.
17. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
18. Krueger RA. Developing questions for focus groups. London, England: Sage Publications Inc; 1997.
19. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
20. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
21. Burkhardt JM, Détienne F, Hébert AM, Perron L, Safin S, Leclercq P. An approach to assess the quality of collaboration in technology-mediated design situations. 2009 Presented at: ECCE '09: European Conference on Cognitive Ergonomics: Designing beyond the Product; September 30-October 2, 2009; Helsinki, Finland. [doi: 10.1007/978-3-642-03658-3_20]
22. co.LAB survey. URL: https://colab.online-studies.net/ [accessed 2021-11-13]
23. 810:30 Federal Act on Research involving Human Beings. Fedlex. 2011 Sep 30. URL: https://www.fedlex.admin.ch/eli/cc/2013/617/en [accessed 2021-11-13]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
co.LAB: A Digital Lab for the co-Design, co-Development and co-Evaluation of Digital Learning Games
COREQ: Consolidated Criteria for Reporting Qualitative Research
FG: focus group
HESAV: School of Health Sciences
HES-SO: University of Applied Sciences of Western Switzerland
InterMIS: Interprofessional Major Incident Simulator

XSL•FO
RenderX

MEI: Media Engineering Institute
PRITS: Patients' Rights and Innovative Teaching Strategies
UI: user interface
UX: user experience

Edited by N Zary; submitted 25.08.21; peer-reviewed by A White, M Szeto, V Verma, M Aksoy; comments to author 16.09.21; revised version received 23.09.21; accepted 12.10.21; published 24.11.21.

Please cite as:
Jaccard D, Suppan L, Bielser F
Contribution of the co.LAB Framework to the Collaborative Design of Serious Games: Mixed Methods Validation Study
JMIR Serious Games 2021;9(4):e33144
URL: https://games.jmir.org/2021/4/e33144
doi:10.2196/33144
PMID:34822336

©Dominique Jaccard, Laurent Suppan, Félicia Bielser. Originally published in JMIR Serious Games (https://games.jmir.org), 24.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Team Building Through Team Video Games: Randomized Controlled Trial

Mark J Keith[1*], MSc, PhD; Douglas L Dean[1*], MSc, PhD; James Gaskin[1*], MSc, PhD; Greg Anderson[1*], MSc, PhD

Information Systems Department, Marriott School of Business, Brigham Young University, Provo, UT, United States

[*]all authors contributed equally

Corresponding Author:
Mark J Keith, MSc, PhD
Information Systems Department
Marriott School of Business
Brigham Young University
Campus Dr
Provo, UT, 84602
United States
Phone: 1 801 674 5159
Email: mark.keith@gmail.com

Abstract

Background: Organizations of all types require the use of teams. Poor team member engagement costs billions of US dollars annually.

Objective: This study aimed to explain how team building can be accomplished with team video gaming based on a team cohesion model enhanced by team flow theory.

Methods: In this controlled experiment, teams were randomly assigned to a team video gaming treatment or a control treatment. Team productivity was measured during both pretreatment and posttreatment team tasks. After the pretest, teams who were involved in the team video gaming treatment competed against other teams by playing the Halo or Rock Band video game for 45 minutes. After the pretest, teams in the control treatment worked alone for 45 minutes. Then, all teams completed the posttest team activity. This same experimental protocol was conducted on 2 different team tasks.

Results: For both tasks, teams in the team video gaming treatment increased their productivity significantly more (F_1=8.760, P=.004) on the posttest task than teams in the control treatment. Our flow-based theoretical model explained team performance improvement more than twice as well (R^2=40.6%) than prior related research (R^2=18.5%).

Conclusions: The focused immersion caused by team video gaming increased team performance while the enjoyment component of flow decreased team performance on the posttest. Both flow and team cohesion contributed to team performance, with flow contributing more than cohesion. Team video gaming did not increase team cohesion, so team video gaming effects are independent of cohesion. Team video gaming is a valid practical method for developing and improving newly formed teams.

(JMIR Serious Games 2021;9(4):e28896) doi:10.2196/28896

KEYWORDS

team video gaming; team building; flow; team cohesion; video games; gamification; team; teamwork; cohesion; theory; framework; performance

Introduction

Important organizational work is often performed in teams. Therefore, effective teamwork has long been a critical area for research [1,2]. Yet, evidence continues to emerge that shows teamwork is often ineffective. For example, one practitioner survey found that 97% of employees and executives believe that team alignment is critical to their effectiveness, yet 86% blame a lack of effective team collaboration for workplace failures [3]. As a result, although a strong body of research exists on effective teamwork, there is continual need for research on how to build productive teams.

A common thread through prior research is that effective teamwork requires interpersonal skills [4], the development of effective ways to share work, and team members' willingness to exert effort toward the team task [5]. To help teams develop

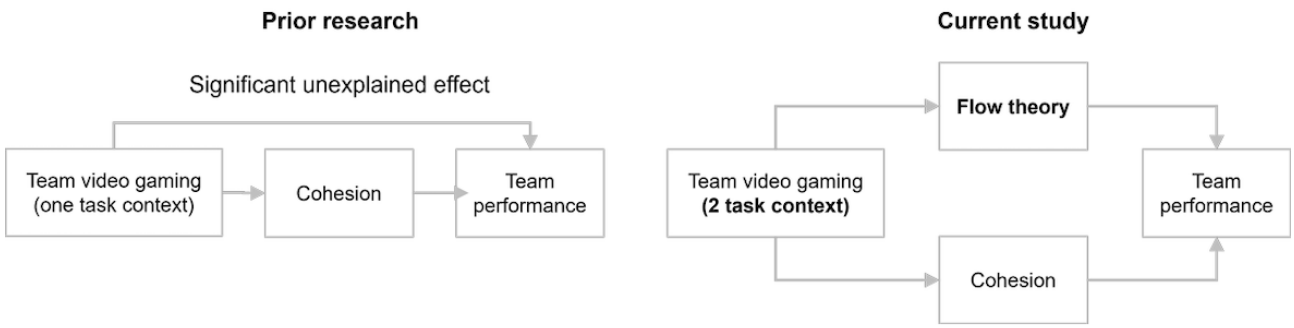
XSL•FO
RenderX

these critical skills and norms, some organizations employ team-building activities. Meta-analyses show that team-building activities significantly increase team performance and productivity [6]. And given that employee disengagement costs the US economy as much as US $350 billion every year [7], some organizations invest extensively in team-building activities. But team-building activities used by organizations can be both time-consuming and expensive. Ropes courses and other challenge retreats are common team-building interventions that have been moderately effective in increasing team performance [8]. But these interventions take employees away from work for a day or longer. The time and expense required by many team-building interventions have led some to argue that team building is a waste of resources [9]. This raises an important question and opportunity for research. Can effective team-building interventions be created that increase team performance but take less time? This study examines team video gaming (TVG) as a possible solution.

For many people, video games have become a pervasive part of life. In the year 2020, 2.6 billion people worldwide played video games [10]. Research on online games shows that players are drawn to games for opportunities for achievement, entertainment, immersive experiences, and interactions with other players [11]. Not all video games are designed exclusively for fun, entertainment, and escape [12]. Video games can use entertainment and enjoyment as a means to help accomplish a variety of worthwhile objectives, such as education [13], training [14], gaining experience in complex situations [15], and facilitating social networks [16].

TVG gives players practice at forming and working effectively with other players in teams [17]. Players work together against an opponent to achieve a common goal. Video games such as Halo, a competitive and cooperative first-person shooter game, and Rock Band, a cooperative music game, encourage players to collaborate and engage as a team to successfully complete shared objectives. Much like competitive sports, teams playing TVG exhibit player engagement, team communication, and strategy formation to achieve a common goal [18]. Another reason TVG is a candidate for team building is that some research suggests that TVG improves team cohesion [19]. Cohesive teams developed social relationships and trust, attraction to the team and to the team tasks, and a knowledge of how to work together [20,21]. This improves team performance on subsequent tasks [22-24]. Some organizations provide video game lounges containing gaming equipment because employees enjoy playing together [25]. However, only one academic study has examined the use of TVG for team building and measured subsequent team performance. Keith et al [26] compared the effects of team building with and without TVG on subsequent team performance in a single task context. They found that teams using TVG increased team cohesion and increased subsequent team performance. However, there was a direct positive effect of TVG on team performance that was not explained by team cohesion. Our study found that TVG also promotes team flow, which accounts for the previously unexplained increase in team productivity. Thus, it provides a better explanation of how team building with TVG improves team performance. This study also found that TVG improved team performance in 2 different task contexts. Figure 1 reflects how the current study compares with that of the study by Keith et al [26] (labeled as "Prior research") and the unique contributions of our study in bold.

Figure 1. Summary of study contributions.



We draw from gamification theory to help explain how TVG can help teams become more productive. *Gamification* is the practice of using game-design elements (ie, rules, goals, interactions, rewards) in nongaming contexts [27,28] or making processes more "game-like" [29]. An important goal of gamification is to promote player engagement [30] so the player is fully involved or immersed in a physical or mental activity [31]. This gamification perspective motivated our selection of *flow* theory to enhance the team cohesion model by Keith et al [26] of TVG-driven team performance. *Flow* theory explains the psychological state of being fully immersed in, and focused on, a task [12,30,32,33]. Research has shown that engagement on work tasks can be manipulated and that a team's collective engagement on one task can recur on subsequent tasks [34-36]. Therefore, we expect teams that experience flow during TVG will also experience flow on subsequent team tasks. To promote this state of flow, we had new teams play commercially available team video games with which they were already familiar, which benefits the team-building process. There are some unique aspects of team cohesion and flow theory that are relevant to the TVG context, thus necessitating both theories. Our research questions are (RQ1) "Can gamifying the team building process via TVG be an effective strategy to improve team performance?" and (RQ2)**"**Can team flow explain the effects of TVG on team performance above and beyond that of team cohesion alone?"

XSL•FO
RenderX

To test our theoretical model, we implemented TVG as a treatment in a laboratory experiment in 2 different contexts or tasks. For the first task, we used the same geocaching task for the pretest and posttest used by Keith et al [26]. For the second task, we used a completely different tower-building task for the pretest and posttest. Using a second task that is distinctly different from the first study allowed us to demonstrate that the effects of the TVG were not limited to 1 task. Our research focuses on helping new teams become productive. Such teams are different from long-term, existing teams because the team members have few, if any, prior relationships among team members. New teams may persist for long or short durations. Helping newly formed, short-term teams become productive is important because many teams in organizations are formed for short durations [37] or may include temporary employees [38].

Theory: How Team Video Gaming Improves Team Performance

Team Building With Video Gaming

Games have long been practiced as a strategy for team building. Games and challenges are used in many traditional team building interventions (eg, ropes courses, competitive team games). The video games we used were developed for entertainment purposes and not for team building but can be used to effectively accomplish the goals of team building. Research on team building has revealed 4 fundamental components of team-building interventions: goal setting, interpersonal relations, problem solving, and role clarification [39-42]. These components are not necessarily activated equally in every team-building intervention. A team-building intervention might include one, multiple, or all these components in varying degrees. Although specific team video games differ in the degree of support for all 4 team-building interventions, some TVG can activate all 4. While playing team video games, individuals increase their skill and share knowledge with teammates [43]. Individuals learn to depend on and be dependable for other team members, which results in building relationships and strengthening teamwork skills [44]. Moreover, the team attributes developed during this process can and do carry over into their cooperative work in high-tech, cross-functional, team-centered workplaces [6,26,45,46]. Our first hypothesis will test this theory: (H1) Teams using TVG will perform better on subsequent team tasks than teams who do not use TVG.

While this finding alone is useful, it is even more important to understand why TVG has this effect. In the following sections, we review 2 theories that may explain the effect of TVG.

Team Cohesion

Team cohesion theory is the dominant historical explanation of how team-building interventions increase team performance (eg, [22,47-49]). At a high level, team cohesion theory states that teams should develop the following attributes to be able to work effectively together: (1) social relationships and trust, (2) attraction to the team, (3) general attraction to team tasks, and (4) knowledge of how to work together [20,21]. Traditional team-building interventions that do not use team video games create team cohesion that subsequently carries over into other tasks, thus improving team performance [22-24].

Team flow is often the "go-to" theory used to explain team performance. However, prior research shows that team cohesion theory has only limited capability to explain the team performance effect that comes from video gaming [26]. Therefore, we turn to flow theory, which is more appropriate for video game designs.

Flow

Flow state refers to the psychological condition of being totally immersed in, and focused on, a task [12,30,32]. The concept of flow has been used to explain the deep engagement of individual persons in activities involving high levels of intrinsic motivation [50] and the optimal state of engagement [12,30]. Flow is typically measured using 5 subconstructs: focused immersion, heightened enjoyment, control, curiosity, and time dissociation.

When individuals enter a state of flow, they feel like they are in control, their curiosity is engaged, and they enjoy themselves [51]. They focus intently on the task at hand and tune out outside stimuli. When experienced by a team (eg, [52,53]), these aspects of engagement may promote team performance—because the team is curiously engaged, feels in control, focuses intently on their task, and tunes out distractions, all while having a good time.

Video games are known to induce flow states easily because they meet 3 necessary preconditions [50]. First, to experience flow, there must be a clear goal to be achieved. Playing video games typically has a superordinate goal of winning the game or performing well. Users must perform tasks that support achieving the superordinate goal. Second, there must be feedback that reflects the degree of performance toward the goal [54]. The games used in this research provide clear, real-time feedback to teams through dashboards that show players how well they are performing in real time. As the games progress, teams can see where they stand and can focus their efforts accordingly. Finally, and perhaps most critically, there must be a balance between skill and the level of "appropriate challenge" in the task of interest [50,54-58]. Flow can be achieved only if the challenge is appropriately matched with skill. In other words, the task cannot be either too easy or too difficult. In the TVG context, playing against a much more skilled team would result in frustration, whereas playing against a much less skilled team would be boring. Thus, video games do not drive flow; rather, a state of flow occurs when using video games provides an appropriate challenge. (H2) TVG will increase the degree of perceived appropriate challenge in a subsequent task. (H3) Appropriate challenge has a positive effect on flow.

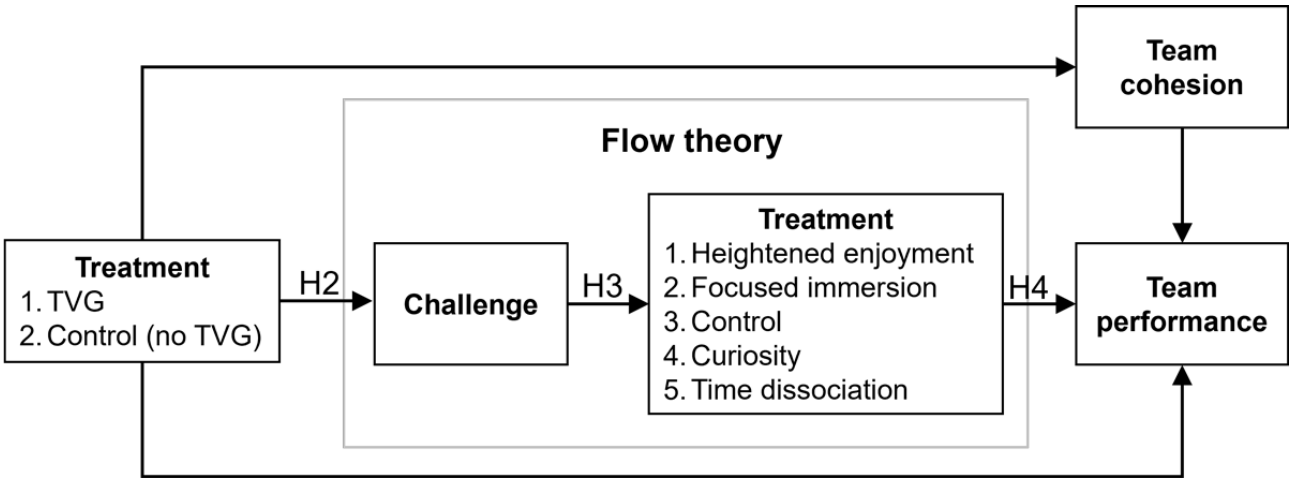
Flow and Performance

The positive effects of flow have been observed in prior research [59-61]. As flow increases, team performance should increase [52,53] because flow represents deep engagement and focus on the task at hand [62]. The strongest outcome of flow in the model by Agarwal and Karahanna [62] was the effect on perceived ease of use. Such a finding suggests that when we experience flow, we perceive the task at hand to be easier to accomplish. Furthermore, Rutkowski et al [63] found that *focused immersion* (a subdimension of flow) led to greater

XSL•FO
RenderX

performance in virtual teams. When teams are immersed in a task, they ignore external (distracting) stimuli that may divert attention away from the task [12]. Thus, an increase in flow should increase performance (Figure 2). (H4) Flow, including (1) focused immersion, (2) heightened enjoyment, (3) control, (4) curiosity, and (5) time dissociation will increase team performance.

Figure 2. Theoretical model. TVG: team video gaming.



Methods

Participants

We used a laboratory experiment design to test our theoretical model. Participants in this study were undergraduate students at a private western university in the United States. They were randomly assigned to teams of 4 participants, and teams were randomly assigned to treatments. However, we questioned the participants prior to making their final team assignment to detect whether they had any preexisting relationships with potential teammates. If they did, the teams were rearranged to minimize the impact of prior relationships. Participants assigned to TVG treatments could select 1 of 2 popular games based on their familiarity with the game. Subjects chose either Rock Band 3 or Halo 4. The popularity of these games makes it likely that students have played them before. These games have also been used in past research [16,64]. This allowed team members to play the game for which they had the most experience and that they found most interesting and engaging. This should have maximized the likelihood of the video game effect and how a team would select a video game in real life. To control for video game ability (and thereby balance appropriate challenge and skill), we asked participants about their level of experience with the game and then balanced the level of video game experience across teams so that teams were competing against other teams of roughly equivalent skill. A total of 586 individuals were divided across 155 teams. However, only 444 participants completed all survey responses measuring latent constructs used to estimate the theoretical model. A few teams had only 3 participants due to no-shows at the lab. Comparative tests showed few differences between 3-person and 4-person teams as noted in the Results section. Of those who chose to report, 141 (141/583, 24.2%) were female, 469 (469/583, 80.4%) were Caucasian, 59 (59/583, 10.1%) were Asian, and 41 (41/583, 7.0%) were Hispanic. The average age of participants was 22.9 years.

Study Design

The study involved a (1) pretest of team task performance to establish a baseline, (2) treatment, and (3) posttest of team task performance, to measure performance improvement.

Task

The tasks in the experiment were designed to replicate the context of a newly formed work team under time pressure. Therefore, the task met the following criteria: (1) it was time sensitive—there was a limited amount of time available to complete the task; (2) it had objective performance measures that were readily calculable—this allowed teams to evaluate their own performance and compare their performance to other teams; (3) the teams selected their own strategies and division of labor—this allowed team members to benefit from their own creativity and ingenuity; and (4) the task required team members to coordinate and collaborate to achieve the best results. We implemented 2 distinct tasks that allowed us to collect objective measures of team performance. Using 2 different tasks allowed us to measure the influence of the tasks on the results.

Task 1 was based on a mobile application called “Findamine” (pronounced “find a mine”) that was created for research purposes and has been successfully used in prior field experiments in information system (IS) research [26,65], but with modifications for our context. Findamine is a geocaching mobile application that generates clues for finding specific landmarks. Rather than giving GPS-based latitude/longitude coordinates, the application gives players short, text-based clues (eg, “This statue depicts the founder of this university.”) that help participants identity the specific location. The destination locations were distributed across the large campus. Teams earned points by successfully deciphering the clue, travelling to the location, and taking a picture of themselves at the location. The pictures of subjects in front of the landmark were automatically uploaded through the mobile application.

Participants could identify and visit more locations by dividing into pairs. So, division of labor, communication between the

XSL•FO
RenderX

team members, and collaboration were rewarded. However, the application tracks the total time elapsed from opening a clue until the correct picture (verified by the GPS coordinates embedded in the photo) is submitted through the application. The natural log of the minutes elapsed was deducted from the possible clue points to reward teams for the speed of their performance in addition to accuracy.

At the conclusion of the task, teams returned to the start location where their performance was displayed in a "leaderboard" so that they could compare their results with those of other teams. In summary, this task enabled all 3 preconditions for flow, namely (1) a clear goal, (2) performance feedback, and most importantly, (3) a challenge that is commensurate with their skills [50,55]. However, given that this geocaching task is somewhat "game-like" itself and may be correlated with the characteristics of TVG, we created a second task (teams only completed 1 of the 2 tasks).

Task 2 was drawn from prior laboratory research on team tasks and performance [66]. It included a timed task of building a tower out of dry uncooked spaghetti noodles and marshmallows. However, there was no leaderboard or real-time feedback about how they were performing relative to other teams, thus minimizing the competitive element. Like the geocaching task, participants were divided into teams of 4 (while minimizing the likelihood of prior relationships among team members). Teams were given 7 minutes to build the tallest tower possible that would remain standing. Performance was measured as the height in inches of the tower. Table 1 shows the number of teams in each treatment assigned to each task.

Table 1. Sample size: number of teams for each treatment and task.

Treatment	Task 1 (geocaching app)			Task 2 (tower building)		
	Number of participants	Number of teams	Number of participants per team, mean	Number of participants	Number of teams	Number of participants per team, mean
Control	191	51	3.75	147	38	3.87
Team video games	112	30	3.73	136	36	3.78

After all data were collected, z scores were calculated for both Task 1 and 2 performances in order to make the results comparable. We also included *task type* (Task 1 coded as 1 and Task 2 as 0) as a covariate in our hypothesis testing.

After being assigned to teams, participants performed either Task 1 or Task 2 (depending on the date of the study) as a baseline measurement of team performance. Similar to prior research [26], those assigned to Task 1 (geocaching) downloaded the app on only 2 of the phones possessed by team members. The app had 6 clues ready with a 25-minute time limit. The team's total score was the combined total of the points on both phones. The phones of the other team members could still be used for communication. Teams were given 5 minutes to plan a strategy. Immediately at the 5-minute mark, the 25-minute timer began during which they could find the clues. As an incentive, teams were notified that the highest-scoring team from each day's participants would earn US $20 Visa cash cards for each member. As locations were found, the results were loaded into a website leaderboard in real time. Upon returning, each team was shown their standing on the leaderboard, and their performance was recorded.

Teams assigned to Task 2 (tower) were placed in an isolated room that was not visible to any other participant teams. They were given a standard number of marshmallows and dry spaghetti noodles. They were also given a brief review of the rules: (1) build free-standing towers not adhered to any furniture or walls, (2) no use of smartphones for ideas or tips, and (3) total score is the combined height of the 2 tallest towers (to give teams the opportunity to determine how to divide roles). Finally, a 7-minute timer was started and left in the room with the team. Similar to Task 1, participants were notified beforehand that the highest-scoring team of that day would earn US $20 Visa gift cards.

Team Intervention: Treatments

Upon completing the pretest, teams were randomly assigned to 1 of 2 treatments: (1) TVG or (2) control (no TVG). Those assigned to the *control* treatment were asked to spend the next 45 minutes by themselves. Team members were instructed to not speak with each other until the posttest began. This was intended to replicate the practical context where no team building occurs. However, they were left at liberty to work individually on homework or any other pursuit unrelated to the experiment.

In the *TVG* treatment, participants played Rock Band 3 or Halo 4. These 2 games were selected primarily because of the interdependent nature of the team tasks. In Rock Band, the players must coordinate their activities to perform the songs correctly. In Halo, the players must coordinate their attacks and defensive strategies to beat the other team. Teams in the Rock Band condition were tasked to earn the highest possible score across any 4 songs of their choosing. The team that earned the highest score earned large candy bars for each member. Those in the Halo 4 condition played 3 rounds of the team-based "capture the flag" subgame against the other team in their cohort. The team winning at least 2 out of 3 matches earned large candy bars. Both treatments lasted 45 minutes. Figures 3 and 4 visualize the gameplay of both games.

XSL•FO
RenderX

Figure 3. Screenshot of Halo.

Figure 4. Screenshot of Rock Band.

Importantly, those in the TVG treatment were not simply left to themselves. Rather, they were playing in a cooperative-competitive environment in which the team they were competing against was in a nearby, but separate, room. A facilitator was assigned to handle technical problems and ensure that the teams played the games according to instructions and fully participated.

Posttask: Measuring Change in Team Performance

After the treatment, teams were again assigned to complete another round of the same pretest task so that team performance could be measured as the relative percent improvement from the pretest. For Task 1 (geocaching), the study administrators downloaded 7 new clues for new locations on campus to the 2 phones for each team running the application. The increase from 6 to 7 clues in Task 2 was because pilot tests revealed that participants gained experience and skill from Task 1 that translated into faster task completion times. Therefore, to even

XSL•FO
RenderX

out the total time required by Task 1 and Task 2, the number of clues was increased for Task 2. Once again, the teams had 5 minutes to strategize and then 25 minutes to find and photograph themselves at as many of the locations as possible. Upon finishing this task, the teams viewed their standing on the leaderboard and completed another survey measuring flow and the other key variables. For Task 2 (tower), participants were placed back into the same room with a fresh set of spaghetti and marshmallows and a clean workspace. They were given another 7-minute timer and set to work.

After the task, each team member took a survey measuring flow and several other covariates. It is important to note that the survey measures referred to the participants' flow state during the geocaching or tower-building task—not the TVG treatment. This is significant because an important assumption of our research is that achieving a flow state during a team intervention (ie, TVG) can increase the likelihood of entering a state of flow on subsequent team tasks as with the geocaching task.

Measures

The attitudinal variables in this study were measured using latent construct items drawn from prior research and adapted for this study. Team flow was adapted from the construct of cognitive absorption [62]. It is an aggregate, second-order construct [67] that includes the subdimensions of *focused immersion* (FI), *heightened enjoyment* (HE), *control* (CO), *curiosity* (CU), and *time dissociation* (TD) [62]. Based on conceptualizations from flow theory [50], 3 items measuring *appropriate challenge* (CH)—or how well the task was equal to (not beneath or above) the skill level of the team—were created new for this study. They were first pilot-tested and met all criteria for validity and reliability (as demonstrated in the following sections). As aforementioned, team performance was measured as the team composite percent increase over the score earned in the pretest task.

Team cohesion was measured based on the conceptual model by Carron et al [68]. Although there are 4 primary subcomponents of team cohesion, the 2 most relevant to the short time duration of our TVG treatment are *group integration-task* (GI-T) and the *individual attractions to the group-task* (ATG-T). The variables Group Integration-Social (GI-S) and Individual Attractions to the Group-Social (ATG-S) were eliminated from the study because we found that the items demonstrated very poor validity and reliability despite being drawn directly from prior research. We believe this indicates the procedures did not provide enough time for the social constructs measured by these instruments to develop to the point where they could play a role in our results. Team flow and team cohesion were modeled as first-order reflective and second-order formative constructs based on prior research [62,68]. *Task interdependence* was included as a control variable because it is a known potential confound in team cohesion research [69]. Items were adapted from Sharma and Yetton [70].

In addition, we included 2 controls: *task type* (geocaching or tower-building) to control for task-related performance differences and their score on the pretest baseline task (standardized to a z score). The latter control is important because teams in both tasks may have been able to achieve a "ceiling" effect where, if they performed extremely well on their pretest task, they would have less room for improvement on the posttest task.

Measurement Model Testing

Pre-analysis was performed to test the convergent and discriminant validity of the reflective subdimension measures, test for multicollinearity, and ensure reliabilities. The constructs, question wording, and outer loadings are summarized in Multimedia Appendix 1 (Table A1). As indicated in Tables A2 and A3 in Multimedia Appendix 1, after removing HE4, CO2, ATG-T3, and ATG-T5, all validity criteria were met. In particular, all average variances extracted (AVEs) were above the 0.50 recommended cutoff and greater than the squared correlation between the focal construct and the subdimensions [71]. Composite reliability was greater than 0.70 for every subdimension. Concerning discriminant validity, the cross-loading matrix in Multimedia Appendix 1 (Table A3) reveals that all item loadings were greater than their cross-loadings. However, the validity scores (Cronbach alpha) for ATGT and CO were slightly below the 0.70 threshold (see Multimedia Appendix 1, Table A2) [72]. Because ATGT is a well-validated instrument, we retained the items as indicated. CO likely had minor problems because only 2 items remained to measure it [73]. However, as all other reliability criteria were met, we continued with the analysis.

We also tested for reliability and validity at the second-order factor level for flow and team cohesion (see Multimedia Appendix 1, Table A4). To test these constructs, latent factor scores were first generated for the subconstructs of both flow (HE, CO, CU, FI, TD) and team cohesion (ATG-T and GIT). These latent factor scores were then used as reflective indicators for the second-order factors. In summary, after testing the measurement model at the highest-level of each construct, all criteria for convergent validity, discriminant validity, and reliability were met. In addition, multicollinearity was not a significant concern as all variance inflation factor scores were below the recommended maximal cutoff of 10 [73]. Therefore, we opted to keep the remaining items and proceed with the hypothesis testing. Overall, the results indicated acceptable factorial validity and minimal multicollinearity or common method bias based on the standards for IS research [74].

Team-Level Constructs

In this study, our treatments were administered to teams, our hypotheses are at the team level, and we predict team-level performance. Thus, the team is our unit of analysis. These perceptions are what Klein and Kozlowski [75] refer to as "emergent unit properties" because they "originate in experiences, attitudes, perceptions, cognitions, or behaviors that are held in common by team members." The team-level attitudinal measures reported in this research were aggregated as means of individual scores of team members. We used the *direct consensus* composition model for the constructs of challenge, components of team flow, and group interaction. We used the *referent shift* composition model for attraction to the group [76]. Multimedia Appendix 2 describes the approach we used to justify the aggregation of individual-level responses into team-level measures.

XSL•FO
RenderX

Comparison of Productivity Increase Across Treatments

To examine changes in team productivity between the pretest and posttest in the control treatment versus the video game treatment, we calculated the percent productivity change for each team and used a one-sided *t* test to compare average changes in productivity across the treatments.

Validation of Theoretical Model

We validated our flow-based theoretical model using the survey responses analyzed in 2 path models using the partial least squares (PLS) structural equation modeling technique in SmartPLS 3.2.6 [77]. Use of this analytical approach was appropriate because PLS does not depend on normal distributions or interval scales [78] making it ideal for our objective measures of task performance. The *t* statistics were generated from running 10,000 bootstrap procedures. We estimated 2 models. In the first model, we combined all aspects of flow into one flow measure so we could measure the effects of challenge on flow taken as a whole and how flow impacts team performance. In the second model, we estimated the effects of challenge on the subcomponents of flow and the effects of the subcomponents of flow on performance to examine the differential role of components of flow as suggested in prior research [12]. Both models were based on team-level scores.

Results

Comparison of Productivity Increase Across Treatments

Table 2 depicts the percentage increase from pretest to posttest for the work tasks. For both tasks, teams in the TVG treatments increased their posttest performance over their pretest performance more than the teams in the control treatment. The TVG treatment in Task 1, geocaching, resulted in a 49.6% average improvement compared with 20.3% for the control teams, for a difference of 29.3% (*t* test, *P*<.001). The tower-building task (Task 2) produced greater variance in term performance than Task 1 and a clear "ceiling" effect—meaning that teams who built a very tall tower during the pretest were not able to improve their score by as great a percentage as the geocaching task. Despite this, the TVG treatment resulted in a 72.1% average improvement compared with 49.5% improvement for the control teams, for a difference of 22.6%.

Table 2. Percent change in team productivity from the pretest to posttest for Tasks 1 and 2.

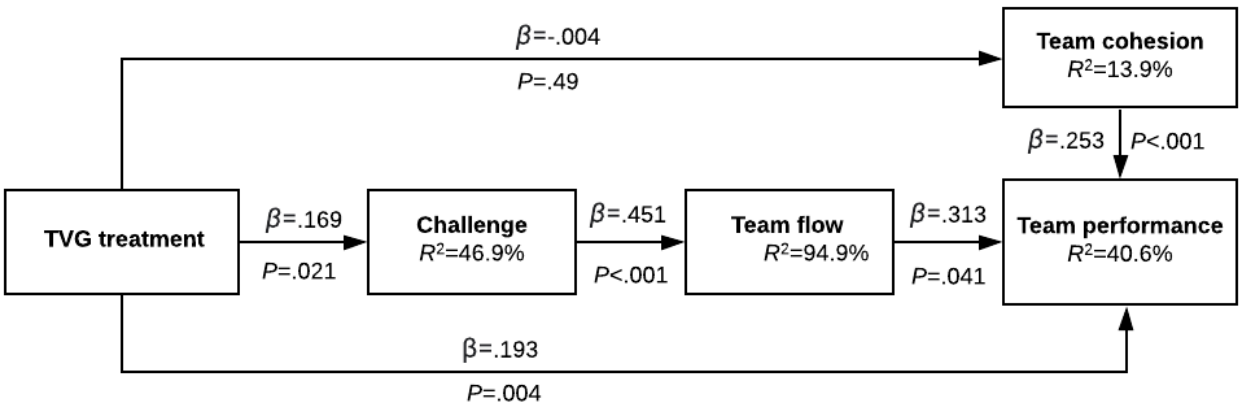
Task	Control treatment mean (SD)	Team video games treatment mean (SD)	Percent difference, (%)
Task 1 (geocaching app)	20.3 (.328)	49.6 (.399)	29.3
Task 2 (tower building)	49.5 (.754)	72.1 (.801)	22.6

An analysis of variance controlling for game (Halo or Rock Band), gender (percent female), age, team size, and task showed that the TVG treatment had a very strong and significant effect on improving team performance (F_1=8.760, *P*=.004). Thus, H1, which predicted that teams using TVG would perform better on subsequent team tasks than teams who did not use TVG, was supported for both tasks. Additionally, we examined differences in mean scores by video game, task, gender, and team size; these are described in Multimedia Appendix 3.

Validation of Theoretical Model

With H1 supported, our next task was to validate our theoretical model explaining why TVG has a strong positive effect on team performance. Figure 5 shows the path coefficients (β) above each relationship and *P* values in parentheses based on a team-level analysis. R^2 values are shown for the endogenous variables. Although not depicted in Figure 5, task type, gender (percent female), and team size were included as covariates explaining each endogenous variable.

Figure 5. Hypothesis testing. TVG: team video gaming.



Supporting H2, the TVG treatment increased the teams' perceptions of the challenging nature of the subsequent tasks (β=.169, *P*=.021). As predicted by H3, challenge increased the perception of team flow (β=.451, *P*<.001). The exceptionally high R^2 for team flow (94.9%) is not unexpected since, as noted earlier, challenge is an essential prerequisite for flow. Perhaps

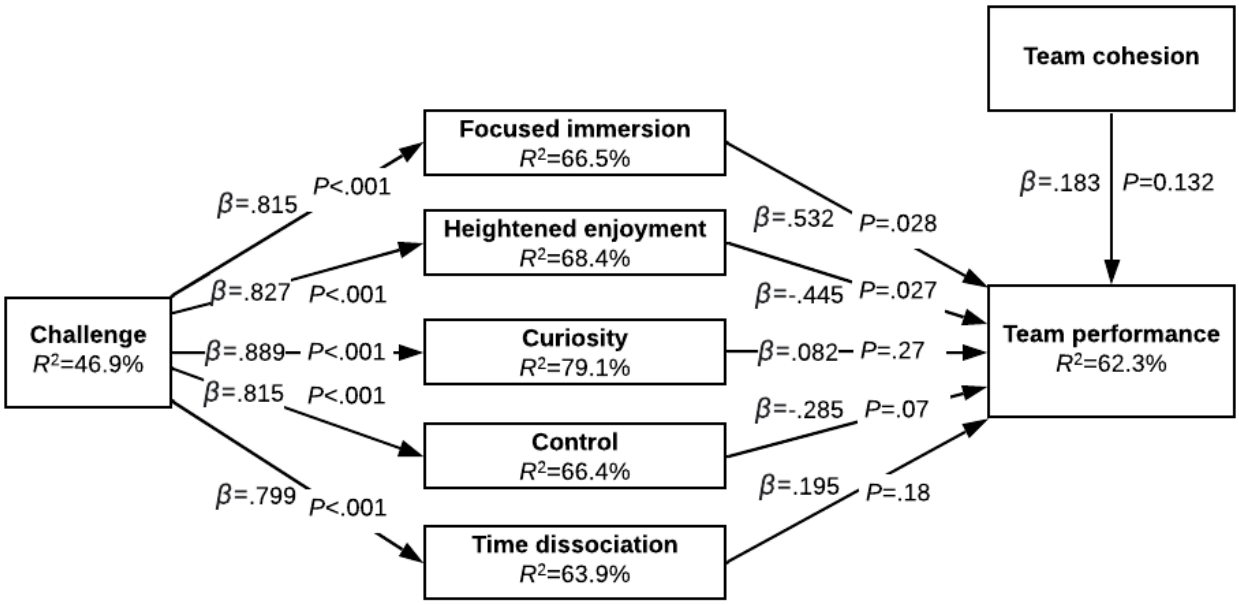
XSL•FO
RenderX

most central to our study, H4 was supported, in that team flow significantly increased performance (β=.313, *P*=.041). The control variables gender (β=–.056, *P*=.21), team size (β=.017, *P*=.40), and task type (geocaching versus tower building; β=.265, *P*=.20) had no significant impact on team performance.

Although not specifically hypothesized, we re-examined the relationships tested in prior research on team cohesion. TVG did not significantly increase team cohesion after controlling for task type (β=–.004, *P*=.49). Team cohesion did increase team performance (β=.253, *P*<.001).

With hypothesis testing completed in Figure 5, Figure 6 provides a deeper understanding of team flow by depicting the results of the model when the subcomponents of flow are estimated as opposed to the second-order flow factor. This allows us to better compare the effects of flow versus team cohesion. The treatment was removed from this model for simplicity. Challenge contributed positively and significantly to all components of flow (*P*<.001). However, the effects of the components of flow on performance were much more differentiated. Focused immersion positively influenced team performance (β=.532, *P*=.028). Surprisingly, heightened enjoyment had a significant negative effect on team performance (β=–.445, *P*=.027). Control had a moderately significant negative effect (β=–.285, *P*=.07). Curiosity (β=.082, *P*=.27) and time dissociation (β=.195, *P*=.18) had no significant effects on performance. Breaking flow into its subconstructs produced a model that better explains the effects of a TVG treatment on performance. The R2 value for team performance improved from 40.6% in Figure 5 to 62.3% in Figure 6.

Figure 6. First-order team flow.



Although not depicted in Table 2 and Figures 5 and 6, there were significant effects of the type of task performed (1=geocaching versus 0=tower-building): After controlling for all other relevant paths, those who participated in the geocaching task experienced a greater sense of appropriate challenge (β=.739, *P*<.001) and thus, greater flow (β=.613, *P*<.001) while developing less team cohesion (β=–.375, *P*<.001). In addition, as expected, higher performances in the pretest led to lower relative increases in posttest performance (β=–.539, *P*<.001).

Finally, the control variables gender (β=–.082, *P*=.12), team size (β=.004, *P*=.48), and study (geocaching versus tower building; β=–.181, *P*=.36) had no significant impact on team performance. Table 3 summarizes the hypotheses and main results of this study.

XSL•FO
RenderX

Table 3. Summary of hypotheses and findings.

Hypothesis number	Support	Finding
H1	Yes	Team video gaming increased team performance on subsequent team tasks.
H2	Yes	Team video gaming increased perceived appropriate challenge in subsequent team tasks.
H3	Yes	Appropriate challenge increased flow.
H4	Yes	Flow, comprised of all flow components, increased team performance.
H4.1	Yes	*Focused immersion* increased team performance.
H4.2	No	*Heightened enjoyment* decreased team performance.
H4.3	No	*Curiosity* did not increase team performance.
H4.4	No	*Control* did not increase team performance.
H4.5	No	*Time dissociation* did not increase team performance.

Discussion

Principal Findings

In response to our 2 research questions, the primary findings of this research are first that the team-building process can be implemented with TVG. Unlike traditional team-building exercises, which focus the attention of team members on specific team-building objectives, these objectives are met naturally as the teams focus on cooperating to play the game. Existing features in the games used in this research supported cooperation and competition through enjoyable and challenging game scenarios, thereby effectively accomplishing the team-building process. Second, teams that played TVG experienced greater flow and exhibited greater performance on subsequent team tasks. This effect was found with 2 different video games and with 2 different work tasks; thus, the positive effects of team building with TVG are not limited to 1 video game or 1 work task. TVG provides some benefits traditionally attributed to team cohesion: A limited form of team attraction, attraction to the task, and roles and scenarios embodied in the game serve to facilitate effective division of labor and cooperation within the team. But TVG does not manipulate team cohesion. The effects of team cohesion on team performance are independent of TVG. Team cohesion has a small, positive impact on team performance but is not affected by TVG. Instead, TVG improves performance by increasing appropriate challenge, which increases flow. The positive effect of flow on performance is stronger than the positive effect of team cohesion on performance. Focused immersion is the component on flow that increases team performance. Heightened enjoyment is a component of flow that decreases team performance. These findings constitute a better explanation of how TVG increases team performance than prior research. We found a much stronger effect size for team performance. In particular, our model explained 62.3% of the variance (R^2) in team performance across 2 distinct types of team tasks compared with the 18.5% explained by Keith et al [26].

Implications

This study contributes to the team-building literature in a unique way. We demonstrated that TVG is an effective team-building strategy. TVG does this by creating an immersive environment that motivates team members to rise to a challenge and achieve a flow state. Most importantly, much like the concept of work engagement state [79], this research demonstrated that flow state can be carried over to improve performance on subsequent team tasks.

One important implication of this research is that the clarity and explanatory power of our model improved substantially when we analyzed the components of flow at the first-order subconstruct level. Given the adequate measurement model metrics we found for flow as a second-order construct, one could reasonably expect that each subconstruct of flow would have an important and positive effect on team performance. However, that was clearly not the case in our data. The positive flow effect came entirely from focused immersion. The focused immersion aspect of flow experienced during TVG carried over into subsequent tasks. That is not the case with the other subconstructs. Enjoyment had a negative effect on the subsequent task performance, whereas curiosity, control, and time dissociation produced no significant effects on team performance.

These noneffects can likely be attributed to the specific attributes of our experimental tasks. Curiosity requires time for contemplation, but time was limited during both the geocaching and tower-building tasks. Further, teams were competing against other teams and had limited time and could not see the scores of the other teams until after they completed the task. This likely explains why they felt little control. Lastly, because our tasks were short and time was controlled and limited, the teams had to keep track of the time they had left while completing the pretest and posttest tasks. Thus, they could not experience time dissociation. These aspects of flow might be experienced in team tasks that are less time constrained.

However, the time limitations of our tasks do not account for the significant *negative* effect of heightened enjoyment on team performance improvements (β=–.445, P=.027). Any heightened enjoyment developed during TVG may have led to greater heightened enjoyment of subsequent tasks, which led to lower performance on those tasks. This effect is somewhat harder to explain, particularly considering that prior research has found that heightened enjoyment is positively correlated with an employee's absorption in work tasks and motivation [46] and even work teams' collective efficacy beliefs [36]. The contradiction of our findings with prior research may be

XSL•FO
RenderX

explained by the source of enjoyment that comes from video games versus work fulfillment. For example, the enjoyment that comes from work accomplishment may be different from the enjoyment from leisure experiences [80]. Therefore, while TVG may help build team focus, it may also create a "let down" effect where the type of enjoyment built during TVG is incongruent with the type of enjoyment that comes from accomplishing work tasks. Future research should explore this possible negative effect of enjoyment on team tasks in the workplace. If this negative effect persists, future research could also examine whether there are types of team video games that produce a form of heightened enjoyment that is more in line with work fulfillment and work enjoyment.

There are other reasons future research should measure and analyze flow at the subconstruct level. For example, the influence of challenge on flow was somewhat suppressed when flow was measured as a second-order construct (β=.451, P<.001) but was substantially stronger, ranging from .799 to .889, for the individual subconstructs. Another consequence of assessing flow at the aggregate was to understate the impact of flow on performance and to overstate the positive influence of team cohesion on performance. Our results also suggest that flow can be measured and assessed at the team level by aggregating responses of individuals when members of teams receive the same treatments and there is a high degree of reliability in terms of how team members rate the constructs [52,53,81]. These criteria were met in this research.

There are many additional research applications in which the effects of team flow could be examined. For example, our results demonstrate an effect of team flow in a controlled, laboratory environment; however, team flow could also be adapted to information technology (IT) project teams and subteams to see how time-critical IT projects can be enhanced with video gaming at a project kickoff meeting.

An important implication of these findings is that video games manipulated flow but not team cohesion and that flow had a stronger impact on performance than cohesion. Team cohesion has received decades of research support in a variety of settings, and we do not claim flow will be more important to performance than cohesion for all settings. Flow is relevant in the TVG context where the task requires focused immersion, heightened enjoyment, control, curiosity, and time dissociation. While some work tasks meet these criteria, some do not. It remains to be seen whether team building with TVG will benefit these other tasks.

Furthermore, both flow and team cohesion matter. Both theories help explain team performance, but each explains different aspects of team building. Team cohesion explains the social integration of the team and their attraction to the task at hand [20] whereas flow explains the state the team is in, the *manner* in which a task is performed. We believe the type of task will likely determine which factor provides the dominant explanation. Flow exerted a stronger influence in our context—newly formed work teams completing time-sensitive, short-term tasks. Cohesion will likely exert greater influence on performance when there is time to develop relationships with team members and an attraction to the task. Future research might extend our results by discovering the boundaries between task types that explain where one theory might be dominant over the other—although both are likely to be relevant to some degree in every task.

The overall effect of TVG to increase team performance was significant for both tasks used in this study: 29.2% for Task 1 and 22.5% for Task 2. Thus, they exceeded the increase of 20% found by Keith et al [26]. This means if tasks performed in our study are representative of other tasks, team building for 45 minutes with TVG for newly formed work teams may increase performance for subsequent tasks requiring several hours or more. TVG requires much less cost and time than traditional team-building activities like retreats and ropes courses. However, it should be noted that the teams in our study were highly engaged in the TVG task.

The finding that the geocaching task produced less team cohesion may be the result of the teams splitting up to find different landmarks, whereas all team members worked together on the tower-building task. Plus, geocaching teams had to determine what each landmark was, determine where it was located, and go to that landmark to take a picture there, so the geocaching task was more challenging and produced more flow than the other task. The descriptive summary in Multimedia Appendix 1 (Table A7) also has some interesting implications. For example, Halo appeared to induce greater improvements in team performance than Rock Band but did not produce greater flow. Therefore, there may be alternative explanations to explain the difference between game features that lead to performance differences. And although Halo led to greater team performance, women were more likely to select Rock Band. It is also interesting that women found more heightened enjoyment in the posttest in our study. Since women were the minority participants, this effect may have come from being part of teams that were more diverse demographically. Future research should explore how team homophily moderates the TVG effect.

Lastly, the strong and consistent impact of appropriate challenge on each subdimension of flow provides strong empirical evidence of the validity of flow theory (ie, that flow is enabled by appropriate challenge). This also implies that studies that neglect to measure or manipulate appropriate challenge may be missing an important antecedent to flow. TVG does not result in flow unless challenge is balanced with skill. Thus, employees who dominate, or get dominated, in the TVG arena may not feel appropriately challenged, which would reduce the likelihood of attaining a state of flow. Future users of TVG for team building should be aware of this.

Limitations

A few limitations of this research are worth noting. First, this was a laboratory experiment. Although laboratory experiments are a necessary and useful first step in establishing a phenomenon, future research is needed to ensure the results of this study are generalizable to practical workplace settings. One limitation arising from the experimental setting is the artificial time pressure to which we attributed mixed effects of some aspects of flow on team performance.

XSL•FO
RenderX

Managers should not assume all TVG will be beneficial. Each team had a facilitator that encouraged them to participate fully, and the participants may have been motivated by knowing that their cooperation would help the researchers. Therefore, the positive effects of TVG may not be replicated in actual settings if the teams do not engage in the video games.

Our geocaching and tower-building tasks were designed to be enjoyable. The geocaching task was somewhat game-like in that it included a leaderboard and real-time feedback about competitive performance. This may have contributed to the carryover effects of challenge and flow from the TVG treatment to the subsequent tasks. However, we included it to keep our results comparable to prior research on TVG [26]. In contrast, the tower-building task was selected to avoid this bias. There was no leaderboard or competitive feedback, making it more distinct from the TVG treatment. Nevertheless, neither of the tasks were particularly representative of common business work tasks. Therefore, future research should replicate our findings using more generalizable business-oriented tasks.

Another limitation concerns our use of 2 video games. While the video games used represented very different genres, these games have feature sets that represent only some video game characteristics that could be useful as team-building interventions. Future research could map the characteristics of other game features to the traditional team building treatments (interpersonal relations, goal setting, role clarification, problem solving) to see which game types are most effective for each team-building treatment.

Our results were found specifically with participants who were previously unfamiliar with each other, yet interdependent in terms of accomplishing a team task. The TVG treatment may not have the same effect on preexisting teams who have already established norms, biases, and opinions about other team members. In these settings, competitive video gaming may reinforce existing negative biases in relationships that already exist.

Additionally, our participants were college-age students who were generally familiar with video games. The advantage of using student subjects is that (1) it allowed us to replicate the context of new teams and (2) students are typically younger and may be more interested in video gaming than older employees [82]. So, students do not represent all types of employees, and some employees may have negative attitudes toward TVG. Moreover, these students had experience with the games studied in this research. Future research could explore how TVG might work with those who are not familiar with video games. Similarly, if a workplace is not characterized by time-critical, objectively measured tasks, then TVG may become more of a distraction than a team-building activity. Future research should examine how flow works in settings without time pressure.

Finally, although our operationalization of the control group treatment represented a practical example of companies that do not facilitate any sort of team-building activity, it also allowed the opportunity for many possible explanations for the difference with the TVG treatment. For example, participants in the TVG treatment not only got to play video games together but also had the opportunity to develop communication skills and get rewards, whereas the control treatment participants were not allowed to interact in any way. Therefore, we cannot definitively conclude that it was the TVG treatment that caused the performance improvements and not simply the communication or candy rewards. Future research should include additional control treatments that allow basic socializing and communication in between tasks as well as break apart the elements of (1) video gaming, (2) cooperation, and (3) competition to understand the degree to which each element contributes to flow and positive performance effects of TVG.

Conclusion

Overall, our research contributes to work on TVG outcomes in 3 ways. First, it confirms prior research [26] that found TVG has a positive overall effect on subsequent team performance. Second, flow theory explains significantly more variance in team performance than prior research. TVG creates a norm of finding an appropriate level of challenge and engagement with team tasks. Finally, it demonstrates that flow is a distinct and complementary construct to the traditional team cohesion theory used to explain team performance. We hope our results will inspire additional research into challenge, flow, and the potential benefits of TVG for small work teams.

Conflicts of Interest

None declared.

Editorial Notice

This randomized study was not registered, explained by authors as follows: “This was a study with no physical risks of injury of any kind. We are still waiting for the PRS account approval”. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials, given that the subjects were undergraduate students. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness.

Multimedia Appendix 1
Measurement model tables and summaries.
[DOCX File , 72 KB - games_v9i4e28896_app1.docx]

Multimedia Appendix 2
Aggregation of individual-level responses into team-level measures.

XSL•FO
RenderX

[DOCX File , 33 KB - games_v9i4e28896_app2.docx]

Multimedia Appendix 3
Differences in mean scores by video game, task, gender, and team size.
[DOCX File , 20 KB - games_v9i4e28896_app3.docx]

Multimedia Appendix 4
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 730 KB - games_v9i4e28896_app4.pdf]

References

1. Salas E, Fiore SM, editors. Team cognition: Understanding the factors that drive process and performance. Washington, DC: American Psychological Association; 2004.
2. Stewart GL, Manz CC, Sims HP. Team Work and Group Dynamics. New York, NY: Wiley; 1999.
3. Gates M. Is poor collaboration killing your company? Confessionsoftheprofession.com. URL: https://confessionsoftheprofessions.com/poor-collaboration-killing-company-infographic/ [accessed 2021-11-05]
4. Kawamoto E, Ito-Masui A, Esumi R, Ito M, Mizutani N, Hayashi T, et al. Social Network Analysis of Intensive Care Unit Health Care Professionals Measured by Wearable Sociometric Badges: Longitudinal Observational Study. J Med Internet Res 2020 Dec 31;22(12):e23184 [FREE Full text] [doi: 10.2196/23184] [Medline: 33258785]
5. Salas E, Priest HA, DeRouin RE. Team building. In: Stanton N, Hedge A, Brookhuis K, Salas E, Hendrick H, editors. Handbook of human factors and ergonomics methods. London, England: CRC Press LLC; 2005:48-41.
6. Klein C, DiazGranados D, Salas E, Le H, Burke CS, Lyons R, et al. Does Team Building Work? Small Group Research 2009 Jan 16;40(2):181-222. [doi: 10.1177/1046496408328821]
7. McDade S. Calculating the Cost of Employee Disengagement. PeopleMetrics. URL: https://www.peoplemetrics.com/ex-blog/calculating-the-cost-of-employee-disengagement [accessed 2021-11-02]
8. Gillis LH, Speelman E. Are Challenge (Ropes) Courses an Effective Tool? A Meta-Analysis. Journal of Experiential Education 2010 Apr 23;31(2):111-135. [doi: 10.5193/jee.31.2.111]
9. Valdes-Dapena C. Stop wasting money on team building. Harvard Business Review. 2018 Sep 11. URL: https://hbr.org/2018/09/stop-wasting-money-on-team-building [accessed 2021-11-02]
10. Clement J. Number of active video gamers worldwide from 2015 to 2023. Statista. 2021 Jun 01. URL: https://www.statista.com/statistics/748044/number-video-gamers-world/ [accessed 2021-11-02]
11. Ryan RM, Rigby CS, Przybylski A. The Motivational Pull of Video Games: A Self-Determination Theory Approach. Motiv Emot 2006 Nov 29;30(4):344-360. [doi: 10.1007/s11031-006-9051-8]
12. Lowry P, Gaskin J, Twyman N, Hammer B, Roberts T. Taking “Fun and Games” Seriously: Proposing the Hedonic-Motivation System Adoption Model (HMSAM). JAIS 2013 Nov;14(11):617-671. [doi: 10.17705/1jais.00347]
13. Gaskin J, Berente N. Video Game Design in the MBA Curriculum: An Experiential Learning Approach for Teaching Design Thinking. CAIS 2011;29:103-122. [doi: 10.17705/1cais.02906]
14. Eichenbaum A, Bavelier D, Green CS. Video games: Play that can do serious good. American Journal of Play 2014;7(1):50-72 [FREE Full text] [doi: 10.4135/9781412972000.n330]
15. Huntemann NB, Payne MT, editors. Joystick Soldiers: The Politics of Play in Military Video Games. Milton Park, Abingdon-on-Thames, Oxfordshire, England, UK: Routledge; 2009.
16. McGonigal J. Reality Is Broken: Why Games Make Us Better and How They Can Change the World. London, England: Penguin Books; 2011.
17. Liaw SY, Ooi SW, Rusli KDB, Lau TC, Tam WWS, Chua WL. Nurse-Physician Communication Team Training in Virtual Reality Versus Live Simulations: Randomized Controlled Trial on Team Communication and Teamwork Attitudes. J Med Internet Res 2020 Apr 08;22(4):e17279 [FREE Full text] [doi: 10.2196/17279] [Medline: 32267235]
18. Cottrell C, McMillen N, Harris BS. Sport psychology in a virtual world: Considerations for practitioners working in eSports. Journal of Sport Psychology in Action 2018 Dec 18;10(2):73-81. [doi: 10.1080/21520704.2018.1518280]
19. Bozanta A, Kutlu B, Nowlan N, Shirmohammadi S. Effects of serious games on perceived team cohesiveness in a multi-user virtual environment. Computers in Human Behavior 2016 Jun;59:380-388. [doi: 10.1016/j.chb.2016.02.042]
20. Carron AV, Brawley LR. Cohesion. Small Group Research 2012 Nov 19;43(6):726-743. [doi: 10.1177/1046496412468072]
21. Dyer WG, Dyer JH, Dyer WG. Team building: Proven strategies for improving team performance. San Francisco, CA: Jossey-Bass; 2013.
22. Senécal J, Loughead T, Bloom G. A season-long team-building intervention: examining the effect of team goal setting on cohesion. J Sport Exerc Psychol 2008 Apr;30(2):186-199. [doi: 10.1123/jsep.30.2.186] [Medline: 18490790]
23. Bloom G, Stevens D, Wickwire T. Expert Coaches' Perceptions of Team Building. J. of Appl. Sport Psych 2003;15(2):129-143. [doi: 10.1080/10413200305397]

XSL•FO
RenderX

24. DiMeglio K, Padula C, Piatek C, Korber S, Barrett A, Ducharme M, et al. Group cohesion and nurse satisfaction: examination of a team-building approach. J Nurs Adm 2005 Mar;35(3):110-120. [doi: 10.1097/00005110-200503000-00003] [Medline: 15761307]
25. Lucasimo. Video games in the workplace. Section G2: Data, Information, and Culture Internet. 2014 Nov 3. URL: http://www.ericrettberg.com/datacultg2/longer-essays/video-games-in-the-workplace/ [accessed 2021-11-02]
26. Keith MJ, Anderson G, Gaskin JE, Dean DL. Team Gaming for Team-building: Effects on Team Performance. THCI 2018:205-231. [doi: 10.17705/1thci.00110]
27. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification. using game-design elements in non-gaming contexts. 2011 Presented at: CHI'11 Extended Abstracts on Human Factors in Computing Systems; May 7-12, 2011; Vancouver BC Canada p. 2425-2428. [doi: 10.1145/1979742.1979575]
28. Treiblmaier H, Putz LM, Lowry PB. Research Commentary: Setting a Definition, Context, and Theory-Based Research Agenda for the Gamification of Non-Gaming Applications. AIS Transactions on Human-Computer Interaction 2018;10(3):129-163 [FREE Full text] [doi: 10.17705/1thci.00107]
29. Werbach K. (Re)Defining Gamification: A Process Approach. In: Spagnolli A, Chittaro L, Gamberini L, editors. Persuasive Technology. PERSUASIVE 2014. Lecture Notes in Computer Science, vol 8462. Cham, Switzerland: Springer; 2014:266-272.
30. Sweetser P, Wyeth P. GameFlow. Comput. Entertain 2005 Jul;3(3):3-3. [doi: 10.1145/1077246.1077253]
31. Lee J, Liu M. Design of Fantasy and Their Effect on Learning and Engagement in a Serious Game. Hershey, PA: IGI Global; 2017.
32. Admiraal W, Huizenga J, Akkerman S, Dam GT. The concept of flow in collaborative game-based learning. Computers in Human Behavior 2011 May;27(3):1185-1194. [doi: 10.1016/j.chb.2010.12.013]
33. Csikszentmihalyi M. Finding Flow: The Psychology of Engagement with Everyday Life. New York, NY: Basic Books; 1997.
34. Schaufeli WB, Bakker AB, Salanova M. The Measurement of Work Engagement With a Short Questionnaire. Educational and Psychological Measurement 2016 Jul 02;66(4):701-716. [doi: 10.1177/0013164405282471]
35. Rodríguez-Sánchez AM, Devloo T, Rico R, Salanova M, Anseel F. What Makes Creative Teams Tick? Cohesion, Engagement, and Performance Across Creativity Tasks: A Three-Wave Study. Group & Organization Management 2016 Mar 21;42(4):521-547. [doi: 10.1177/1059601116636476]
36. Salanova M, Rodríguez-Sánchez AM, Schaufeli WB, Cifre E. Flowing together: a longitudinal study of collective efficacy and collective flow among workgroups. J Psychol 2014 Jan 21;148(4):435-455. [doi: 10.1080/00223980.2013.806290] [Medline: 24946388]
37. Valentine MA, Retelny D, To A, Rahmati N, Doshi T, Bernstein MS. Flash Organizations: Crowdsourcing Complex Work by Structuring Crowds As Organizations. 2017 Presented at: CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, CO. [doi: 10.1145/3025453.3025811]
38. Wilkin CL, de Jong JP, Rubino C. Teaming up with temps: the impact of temporary workers on team social networks and effectiveness. European Journal of Work and Organizational Psychology 2017 Dec 21;27(2):204-218. [doi: 10.1080/1359432x.2017.1418329]
39. Beer M. The technology of organization development. In: Dunnette MD, editor. Handbook of Industrial and Organizational Psychology. Chicago, IL: Rand McNally & Company; 1976.
40. Beer M, Nohria N. Cracking the code of change. Harvard Business Review. 2000. URL: https://hbr.org/2000/05/cracking-the-code-of-change [accessed 2021-11-02]
41. Dyer WG. Team Building: Issues and Alternatives. Boston, MA: Addison-Wesley; 1987.
42. Buller PF, Bell CH. Effects of team building and goal setting on productivity: a field experiment. Academy of Management Journal 1986 Jun 01;29(2):305-328. [doi: 10.2307/256190]
43. Gee JP. Good video games and good learning: Collected essays on video games, learning, and literacy. Academic Advanced Distributed Learning Co-Laboratory. 2007. URL: https://academiccolab.org/resources/documents/Good_Learning.pdf [accessed 2021-11-02]
44. Cole H, Griffiths MD. Social interactions in massively multiplayer online role-playing gamers. Cyberpsychol Behav 2007 Aug;10(4):575-583. [doi: 10.1089/cpb.2007.9988] [Medline: 17711367]
45. Gee JP. What video games have to teach us about learning and literacy. Comput. Entertain 2003 Oct;1(1):20-20. [doi: 10.1145/950566.950595]
46. Salanova M, Bakker AB, Llorens S. Flow at Work: Evidence for an Upward Spiral of Personal and Organizational Resources*. J Happiness Stud 2006 Mar;7(1):1-22. [doi: 10.1007/s10902-005-8854-8]
47. Klein HJ, Mulvey PW. Two Investigations of the Relationships among Group Goals, Goal Commitment, Cohesion, and Performance. Organizational Behavior and Human Decision Processes 1995 Jan;61(1):44-53. [doi: 10.1006/obhd.1995.1004]
48. Ensley MD, Pearson AW. An Exploratory Comparison of the Behavioral Dynamics of Top Management Teams in Family and Nonfamily New Ventures: Cohesion, Conflict, Potency, and Consensus. Entrepreneurship Theory and Practice 2017 Dec 04;29(3):267-284. [doi: 10.1111/j.1540-6520.2005.00082.x]
49. Carron AV, Colman MM, Wheeler J, Stevens D. Cohesion and performance in sport: A meta analysis. Journal of Sport & Exercise Psychology 2002;24(2):168-188. [doi: 10.1123/jsep.24.2.168]

XSL•FO
RenderX

50. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience. New York, NY: Harper and Row; 1990.
51. Kuipers DA, Terlouw G, Wartena BO, van 't Veer JT, Prins JT, Pierie JPE. The Role of Transfer in Designing Games and Simulations for Health: Systematic Review. JMIR Serious Games 2017 Nov 24;5(4):e23 [FREE Full text] [doi: 10.2196/games.7880] [Medline: 29175812]
52. van den Hout JJ, Gevers JM, Davis OC, Weggeman MC. Developing and Testing the Team Flow Monitor (TFM). Cogent Psychology 2019 Aug 09;6(1):1643962. [doi: 10.1080/23311908.2019.1643962]
53. van den Hout JJ, Davis OC, Weggeman MC. The Conceptualization of Team Flow. J Psychol 2018 Aug 18;152(6):388-423. [doi: 10.1080/00223980.2018.1449729] [Medline: 30089084]
54. Csikszentmihályi M, Abuhamdeh S, Nakamura J. Flow. In: Elliot AJ, Dweck CS, Yeager DS, editors. Handbook of Competence and Motivation. New York, NY: The Guilford Press; 2005.
55. Hoffman DL, Novak TP. Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations. Journal of Marketing 2018 Nov 28;60(3):50-68. [doi: 10.1177/002224299606000304]
56. Silic M, Lowry PB. Using Design-Science Based Gamification to Improve Organizational Security Training and Compliance. Journal of Management Information Systems 2020 Mar 01;37(1):129-161. [doi: 10.1080/07421222.2019.1705512]
57. Santhanam R, Liu D, Shen WM. Research Note—Gamification of Technology-Mediated Training: Not All Competitions Are the Same. Information Systems Research 2016 Jun;27(2):453-465. [doi: 10.1287/isre.2016.0630]
58. Suh A, Cheung CM, Ahuja M, Wagner C. Gamification in the Workplace: The Central Role of the Aesthetic Experience. Journal of Management Information Systems 2017 Apr 20;34(1):268-305. [doi: 10.1080/07421222.2017.1297642]
59. Bakker AB, Oerlemans W, Demerouti E, Slot BB, Ali DK. Flow and performance: A study among talented Dutch soccer players. Psychology of Sport and Exercise 2011 Jul;12(4):442-450. [doi: 10.1016/j.psychsport.2011.02.003]
60. Chu LC, Lee CL, Wu CJ, Hsu KF. Team Members' Flow Experiences and their Influence on Project Performance. The Journal of Human Resource and Adult Learning 2012;8(1):46-55 [FREE Full text]
61. Heyne K, Pavlas D, Salas E. An Investigation on the Effects of Flow State on Team Process and Outcomes. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2011 Sep 01;55(1):475-479. [doi: 10.1177/1071181311551098]
62. Agarwal R, Karahanna E. Time Flies When You're Having Fun: Cognitive Absorption and Beliefs about Information Technology Usage. MIS Quarterly 2000 Dec;24(4):665. [doi: 10.2307/3250951]
63. Rutkowski A, Saunders C, Vogel D, van Genuchten M. “Is It Already 4 a.m. in Your Time Zone?”. Small Group Research 2016 Nov 15;38(1):98-129. [doi: 10.1177/1046496406297042]
64. Peacocke M, Teather RJ, Carette J, MacKenzie IS. Evaluating the effectiveness of HUDs and diegetic ammo displays in first-person shooter games. 2015 Presented at: IEEE Games Entertainment Media Conference (GEM); October 14-16, 2015; Toronto, ON, Canada. [doi: 10.1109/gem.2015.7377211]
65. Steinbart PJ, Keith MJ, Babb J. Examining the Continuance of Secure Behavior: A Longitudinal Field Study of Mobile Device Authentication. Information Systems Research 2016 Jun;27(2):219-239. [doi: 10.1287/isre.2016.0634]
66. Cook LS, Olson JR. The Sky’s the Limit: An Activity for Teaching Project Management. Journal of Management Education 2016 Jul 26;30(3):404-420. [doi: 10.1177/1052562905279220]
67. Polites GL, Roberts N, Thatcher J. Conceptualizing models using multidimensional constructs: a review and guidelines for their use. European Journal of Information Systems 2017 Dec 19;21(1):22-48. [doi: 10.1057/ejis.2011.10]
68. Carron AV, Widmeyer WN, Brawley LR. The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire. Journal of Sport and Exercise Psychology 1985;7(3):244-266. [doi: 10.1123/jsp.7.3.244]
69. Gully SM, Devine DJ, Whitney DJ. A Meta-Analysis of Cohesion and Performance. Small Group Research 2016 Aug 19;26(4):497-520. [doi: 10.1177/1046496495264003]
70. Sharma R, Yetton P. The Contingent Effects of Training, Technical Complexity, and Task Interdependence on Successful Information Systems Implementation. MIS Quarterly 2007;31(2):219. [doi: 10.2307/25148789]
71. Fornell C, Larcker DF. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research 2018 Nov 28;18(1):39-50. [doi: 10.1177/002224378101800104]
72. Santos JRA. Cronbach's alpha: A tool for assessing the reliability of scales. Journal of Extension 1999;37(2):2TOT3.
73. Hair JF, Anderson RE, Tatham R, Black WC. Multivariate data analysis. London, England: Pearson; 1998.
74. Pavlou PA, Liang HG, Xue YJ. Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective. MIS Quarterly 2007;31(1):105. [doi: 10.2307/25148783]
75. Klein KJ, Kozlowski SWJ. From Micro to Meso: Critical Steps in Conceptualizing and Conducting Multilevel Research. Organizational Research Methods 2016 Jun 29;3(3):211-236 P 215. [doi: 10.1177/109442810033001]
76. Chan D. Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. Journal of Applied Psychology 1998;83(2):234-246. [doi: 10.1037/0021-9010.83.2.234]
77. SmartPLS 3. URL: https://www.smartpls.com/ [accessed 2021-11-02]
78. Chin WW, Marcolin BL, Newsted PR. A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption Study. Information Systems Research 2003 Jun;14(2):189-217. [doi: 10.1287/isre.14.2.189.16018]

XSL•FO
RenderX

79. Knight C, Patterson M, Dawson J. Building work engagement: A systematic review and meta-analysis investigating the effectiveness of work engagement interventions. J Organ Behav 2017 Jul 13;38(6):792-812 [FREE Full text] [doi: 10.1002/job.2167] [Medline: 28781428]
80. Tinsley HE, Hinson JA, Tinsley DJ, Holt MS. Attributes of leisure and work experiences. Journal of Counseling Psychology 1993;40(4):447-455. [doi: 10.1037/0022-0167.40.4.447]
81. van Mierlo H, Vermunt JK, Rutte CG. Composing Group-Level Constructs From Individual-Level Survey Data. Organizational Research Methods 2008 Apr 08;12(2):368-392. [doi: 10.1177/1094428107309322]
82. Lenhart A, Kahne J, Middaugh E, Macgill AR, Evans C, Vitak J. Teens, Video Games, and Civics: Teens' gaming experiences are diverse and include significant social interaction and civic engagement. Pew Internet & American Life Project. 2008 Sep 16. URL: https://www.pewresearch.org/internet/2008/09/16/teens-video-games-and-civics/ [accessed 2021-11-02]

Abbreviations

ATG-S: Individual Attractions to the Group-Social
ATG-T: individual attractions to the group-task
AVE: average variance extracted
CH: appropriate challenge
CO: control
CU: curiosity
FI: focused immersion
GI-S: Group Integration-Social
GI-T: group integration-task
HE: heightened enjoyment
IS: information system
IT: information technology
PLS: partial least squares
TD: time dissociation
TVG: team video gaming

Edited by N Zary; submitted 17.03.21; peer-reviewed by T Ye; comments to author 09.04.21; revised version received 02.08.21; accepted 24.09.21; published 14.12.21.

Please cite as:
Keith MJ, Dean DL, Gaskin J, Anderson G
Team Building Through Team Video Games: Randomized Controlled Trial
JMIR Serious Games 2021;9(4):e28896
URL: https://games.jmir.org/2021/4/e28896
doi:10.2196/28896
PMID:34904954

©Mark J Keith, Douglas L Dean, James Gaskin, Greg Anderson. Originally published in JMIR Serious Games (https://games.jmir.org), 14.12.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Changes in Self-Reported Web-Based Gambling Activity During the COVID-19 Pandemic: Cross-sectional Study

Emma Claesdotter-Knutsson[1], MD, PhD; Anders Håkansson[2], MD, PhD

[1]Child and Adolescent Psychiatry, Department of Clinical Sciences, Lund University, Lund, Sweden

[2]Psychiatry, Department of Clinical Sciences, Faculty of Medicine, Lund University, Lund, Sweden

Corresponding Author:
Emma Claesdotter-Knutsson, MD, PhD
Child and Adolescent Psychiatry
Department of Clinical Sciences
Lund University
Barav 1
Lund, 221 85
Sweden
Phone: 46 768871765
Email: emma.claesdotter-knutsson@med.lu.se

Abstract

Background: The COVID-19 pandemic has affected not only somatic health with over 3.7 million deaths worldwide, but also has had a huge impact on psychological health, creating what amounts to a mental health crisis. The negative effect of the pandemic on traditional addictions is well described and concerning, and the same has been seen for gambling.

Objective: This study explores self-reported web-based gambling behavior during the COVID-19 pandemic in Sweden. We investigated overall changes, but also changes in specific web-based gambling types, and whether they are associated with certain risk factors or lifestyle changes.

Methods: Our study is based on an anonymous web-based survey of web panel participants in Sweden (N=1501) designed to study a range of behavioral changes during the COVID-19 pandemic. Increases in gambling were analyzed using logistic regression models against sociodemographic data and psychological distress.

Results: The majority of the respondents who gambled reported no changes in their gambling habits during the COVID-19 pandemic. We found significant associations with the problem gambling severity index (PGSI), the Kessler score (indicating psychological distress), employment status, changes in alcohol habits, and self-exclusion when looking at overall changes in gambling activity in the pandemic. In the subgroup that reported an increase in gambling activity, we found an association with both the PGSI and Kessler scores. The PGSI score was also an independent predictor for all specific web-based gambling (horses, sports, poker, and casino) whereas the Kessler score only had a significant impact on changes in casino gambling. In addition, male gender was an independent predictor for gambling on sports and casino gambling.

Conclusions: The majority of respondents who gambled reported no changes in their gambling activity during the COVID-19 pandemic. The group that reported an increase in overall gambling activity during the COVID-19 pandemic represent a group with gambling problems and psychological distress. The group that reported increased sports and casino gambling were often male, and this group seemed to experience more psychological distress.

(JMIR Serious Games 2021;9(4):e30747) doi:10.2196/30747

KEYWORDS

COVID-19; pandemic; web-based gambling; psychological distress; gender

Introduction

The first case of COVID-19 was detected in December 2019 in Wuhan, China [1] and then the virus spread rapidly, reaching Europe in early 2020 and Sweden in January 2020 [2]. The COVID-19 pandemic has had an enormous effect on mental health [3-5]. Research has confirmed increasing rates of stress, depression, addiction, anxiety, and other psychiatric disorders during the pandemic [6,7].

XSL•FO
RenderX

Virus protection strategies in Sweden have differed from those in other countries and not included lockdown or stay-at-home orders, but rather recommendations. Most work places encouraged their workers to work from home and regulations in opening hours have been in effect for restaurants and shopping malls, leading people in general to spend a lot more time indoors with possible access to screens, and several studies have reported increased screen time during the pandemic [8-11]. Excessive screen time use has been shown to be associated with a range of negative mental health outcomes such as anxiety and depression [12-14].

Historically in economic crises, when people experienced stress due to, for example, isolation, gambling activity per se increased, and so did gambling problems [15-17], but recent studies on potential changes in gambling activity during the COVID-19 pandemic have reported different changes in behavior [18]. One possible explanation might be the restrictions in place in the field of study, along with differences in study populations. Auer et al [19] and Lindner et al [20] found a substantial decrease in overall gambling activity, especially in gambling, where there were far fewer betting opportunities because of cancelled or postponed sports events such as football leagues.

This was also reported in an Italian study conducted during the lockdown, when gamblers reported not only a decrease in gambling but also the relief they felt because the restrictions meant they could not gamble as before [21]. Auer et al [19] confirmed this situation in a population of Swedish, German, Finnish, and Norwegian citizens, who reported no increase in gambling activity. Our research group, focusing on COVID-19–related changes in gambling in a Swedish setting, observed increased gambling activity in a subgroup of high-risk gamblers [22,23]. Lindner et al [20] observed the same pattern among high-risk gamblers and Frisone et al [24] observed this among moderate-risk gamblers.

Early during the pandemic, the Swedish government voiced its concerns about an anticipated increase in gambling owing to rising anxiety, boredom, and depression prevalence expected as the pandemic accelerated [20]. Social distancing to prevent virus transmission was thought likely to encourage gambling [20]. To prevent this, the Swedish government introduced a temporary legislation (Förordning 2020:495) to govern the Swedish gambling market, including a deposit limit at web-based casinos, of Swedish kronor (SEK) 5000 (US $580.48) a week (per gambling provider) and obligatory time limits because of concerns in the early stages of the outbreak that social distancing would increase web-based activities such as gambling and, by extension, the problems associated with gambling. Other countries followed suit [20,23].

Gambling is strongly associated with comorbid disorders, including alcohol use disorder, depression, substance use disorders, nicotine dependence, anxiety disorders, and antisocial personality disorder [25,26]. Gambling is a potentially addictive behavior; gambling disorder is an addiction diagnosis given the same status as alcohol and drug addiction [27,28]. Web-based gambling, known for its autonomy, accessibility, and speed of play, is considered a higher-risk form of gambling, which easily leads to addiction [29,30].

In our study, we explored self-reported web-based gambling behavior during the COVID-19 pandemic in Sweden. We considered overall changes, along with changes for specific types of gambling, and whether potential changes in gambling may be associated with specific risk factors such as level of education, employment status, disposable income, and psychological distress.

Methods

Setting

The first wave of the COVID-19 pandemic in Sweden was in spring 2020. After a decline in virus transmission in the summer, there was a second wave in the autumn of 2020. In the first few months of 2021, there was a third wave of virus transmission and a further increase in the disease burden of hospitalizations (Swedish Public Health Agency 2021). This cross-sectional study is based on a self-report web-based survey carried out in Sweden in March 2021, during the third wave of the COVID-19 pandemic in the country. At the time of this study, upper-secondary schools in Sweden had resumed in-person teaching, but the nationwide COVID-19 restrictions on leisure activities and gatherings of more than 8 people were still in effect.

Participants

We consulted the market survey company Ipsos and used its internet-based web survey panel. We have previously used the Ipsos web panel for internet-based surveys for research purposes [30]. Respondents from the general population, aged 16 years and above, from the Ipsos web panel were invited with the information that the survey would address “computer gaming, gambling for money, and other behavioural patterns in Sweden during Covid-19—Association of mental health, social situation, and attitudes to the pandemic.” The language used in the survey was Swedish. Respondents could only access the survey once they had provided electronic informed consent. Most Ipsos surveys are worth 1 point to respondents, and each point is worth about €1 (US $1.16) when redeemed as goods or services. In this study, invitations were sent until 1500 complete answers were obtained. In addition, in this study, the final distribution of age groups, gender, and geographical location (region) was compared by Ipsos to the general population, such that the data set was weighted in accordance with a summarized weighting score derived from these 3 variables. The survey was halted when the final sample consisted of 1501 individuals. The study was carried out between March 19 and 29, 2021. The study was reviewed and approved by the Swedish Ethical Review Board (File: 2021/00369).

Measures

The basic sociodemographic variables included gender (female or male), age (divided into 2 age groups of 16-24 years and ≥25 years), monthly income (divided into 3 groups of SEK 10,000-20,000 [US $1164.40-$2328.81], SEK 20,000-40,000 [US $2328.81-$4657.62], or >SEK 40,000 [>US $4657.62]), level of education (university, upper-secondary school [ages 16-19 years], compulsory school [ages 6-16 years], or other), and employment status (studying, employed, unemployed,

XSL•FO
RenderX

retired, or other). The questionnaire continued with questions about changes in personal behavior during the COVID-19 pandemic ("since these changes in Sweden started"): the time they spent at home (much more, slightly more, unchanged, or less time at home) and how much alcohol they consumed (more, less, unchanged, or "I don't drink, now nor before"). The questionnaire continued with a general question about their web-based gambling on horse racing, sports, poker, and casino games (more, less, unchanged, or "I don't gamble, now or before"). This was followed by questions for each type of web-based gambling in turn (horses, sports, poker, and casinos). A question was asked about self-exclusion from gambling (gambling breaks) using spelpaus.se, a national self-exclusion scheme that covers all forms of licensed gambling, which has been available since January 1, 2019 (yes, no, or prefer not to answer) [30].

Psychological distress was measured using the Kessler-6 scale to describe symptoms of depressed mood and anxiety [31]. It is a 6-item scale of symptoms perceived in the preceding 6 months, with 6 questions about depressive and anxiety-related symptoms with options ranging from "not at all" to "all the time," including a "cannot answer/prefer not to say" response. For the Kessler-6 scale, the scores of 0-4 for each question were summed, and a total score of ≥5 was classified as at least moderate psychological distress.

The level of potential gambling problems was measured with the 9-item problem gambling severity index (PGSI) [32], where each statement addresses the preceding 12 months, with options including "never," "sometimes," "most of the time," and "almost always." For the PGSI scale, the scores of 0-3 for each question were summed: a total score of 0 indicated no problem with gambling; 1-2, a certain risk of gambling problems; 3-7, a moderate risk of gambling problems; and ≥8 indicated gambling problems.

Statistical Analysis

The reporting of prevalence measures and group-wise comparisons related to the weighted data and statistical tests were applied using the chi-square test, whereas binary logistic regression analyses were carried out using the unweighted data. Binary logistic regressions were carried out with each increased change in gambling (yes/no) as the dependent variable to study potential–independent variables associated with the outcome. In a subsequent step, greater changes in specific forms of gambling (horses, sports, poker, and casino) were established as outcomes. For all models, odds ratios with 95% CIs are presented. SPSS (version 25.0) was used for all statistical analyses.

Availability of Data and Materials

The data sets used and analyzed in the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

All procedures performed in this study involving human participants were in accordance with the ethical standards of the national research committee and with the tenets of the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Patient Consent for Publication

Informed consent was obtained from all individual participants included in the study.

Informed Consent

The manuscript does not contain any individual person's data in any form.

Results

Descriptive Data on Changes in Web-Based Gambling Activity and Type of Gambling

The descriptive data of the sample are shown in Table 1 (weighted data). Of the 1501 respondents, 29% (n=437) responded they had never gambled and were therefore excluded from further analysis. Of the remaining 1064 who were further assessed, 56% (n=596) were male and 44% (n=468) were female. The distribution regarding the different types of web-based gambling activity is described in Table 1.

Table 1. Study population by total and self-reported changes in web-based gambling activity (weighted) (N=1501).

	Changes in gambling activity				
	Total, n (%)	Horses, n (%)	Casino, n (%)	Poker, n (%)	Sports, n (%)
Never gambled	437 (29.1)	1073 (71.5)	1122 (74.8)	1264 (84.2)	1007 (67.1)
Study population	1064 (70.9)	428 (28.5)	379 (25.2)	237 (18.8)	494 (32.9)
Gender					
Female	468 (44.0)	150 (35.1)	123 (32.3)	56 (23.5)	140 (28.4)
Male	596 (56.0)	278 (64.9)	256 (67.7)	181 (76.5)	354 (71.6)

Based on the findings, approximately 10% (n=111) of the participants reported an increase in overall gambling, 82% (n=876) reported unchanged gambling activity, and 7% (n=77) reported a decrease in gambling (Table 2). The figures for increased activity by gambling type were similar at approximately 12% (11.0%-12.9%). The same was seen for the self-reported decrease in activity by type of gambling on horses, sports, and casinos at approximately 13% (13.2%-14.2%), but self-reported decrease in poker was 21% (Table 2).

XSL•FO
RenderX

Table 2. Changes in web-based gambling activity overall and by gambling type (weighted).

Changes	Change in gambling, n (%)				
	Overall (n=1064)	On horses (n=428)	On casino (n=380)	On poker (n=237)	On sports (n=494)
Increased	111 (10.4)	47 (11.0)	49 (12.9)	28 (11.8)	56 (11.3)
Unchanged	876 (82.3)	322 (75.2)	277 (72.9)	160 (67.5)	373 (75.5)
Decreased	77 (7.2)	59 (13.8)	54 (14.2)	49 (20.7)	65 (13.2)

Changes in Web-Based Gambling Activity by Demographic and Socioeconomic Factors

Table 3 presents the weighted data for changes in gambling by demographic and socioeconomic factors. We found significant differences in employment status ($P<.001$), PGSI ($P<.001$), Kessler score ($P<.001$), gambling paus ($P<.001$), as well as changes in alcohol habits ($P<.001$) (Table 3).

XSL•FO
RenderX

Table 3. Changes in gambling activity by demographic and socioeconomic characteristics (weighted).

	Increased (n=111), n (%)	Unchanged (n=876), n (%)	Decreased (n=77), n (%)	*P* value	Total (N=1064), n (%)
Gender				.10	
Female	39 (34.9)	398 (45.4)	32 (41.4)		468 (44.0)
Male	72 (65.1)	478 (54.6)	45 (58.6)		596 (56.0)
Age groups (years)				.07	
16-24	19 (17.3)	96 (11.0)	13 (16.4)		128 (12.0)
≥25	92 (82.7)	780 (89.0)	64 (83.6)		936 (88.0)
Level of education				.51	
Primary school	55 (49.2)	402 (46.0)	41 (52.9)		498 (46.8)
Secondary school	45 (40.4)	364 (41.5)	28 (36.5)		437 (41.0)
University	9 (7.7)	82 (9.3)	3 (4.5)		94 (8.8)
Other	3 (2.7)	28 (3.2)	5 (6.1)		36 (3.3)
Employment status				<.001[a]	
Studying	21 (18.9)	86 (9.9)	12 (16.1)		120 (11.3)
Employed	63 (56.3)	499 (57.0)	50 (65.0)		612 (57.5)
Unemployed	16 (14.2)	44 (5.1)	4 (4.6)		63 (6.0)
Retired	10 (8.9)	225 (25.7)	11 (14.3)		246 (23.1)
Other	2 (1.9)	21 (2.4)	0 (0)		23 (2.2)
Disposable income				.25	
≥SEK 40,000 (US $4657.62)	35 (31.8)	241 (27.5)	29 (38.1)		305 (28.7)
SEK 20,000-40,000 (US $2328.81-$4657.62)	60 (54.1)	474 (54.2)	37 (48.2)		572 (53.7)
≤SEK 10,000-20,000 (US $1164.40-2328.81)	16 (14.2)	161 (18.3)	11 (13.7)		187 (17.6)
Problem Gambling Severity Index				<.001[a]	
No problem with gambling	25 (24.5)	600 (80.0)	41 (61.2)		666 (62.6)
Certain risk of gambling problems	17 (17.2)	86 (11.4)	14 (21.6)		117 (11.0)
Increased risk of gambling problems	34 (33.9)	55 (7.3)	7 (10.7)		96 (9.1)
Gambling problems	25 (24.4)	10 (1.4)	4 (6.5)		39 (3.7)
Kessler score				<.001[a]	
Score 0-4; no psychological distress	19 (16.7)	443 (51.1)	25 (34.7)		487 (45.8)
Score 5-24; psychological distress	92 (83.2)	423 (48.9)	48 (65.3)		563 (52.9)
Gambling pause				<.001[a]	
Yes	21 (18.2)	24 (2.8)	12 (14.9)		51 (5.2)
No	91 (81.7)	846 (9.7)	65 (85.1)		1000 (94.8)
Time at home				.11	
Much more	76 (67.8)	496 (5.7)	44 (57.3)		614 (57.8)
Slightly more	29 (25.8)	258 (29.5)	20 (25.8)		307 (28.8)

XSL•FO
RenderX

	Increased (n=111), n (%)	Unchanged (n=876), n (%)	Decreased (n=77), n (%)	*P* value	Total (N=1064), n (%)
Unchanged	7 (5.6)	119 (13.6)	12 (15.4)		137 (12.9)
Less time	1 (0.9)	4 (0.4)	1 (1.6)		6 (0.5)
Change in alcohol habits				.001[a]	
Yes more	26 (23.2)	74 (8.5)	8 (9.6)		105 (10.1)
Unchanged	38 (34.0)	530 (60.5)	24 (31.2)		592 (55.6)
Less alcohol	30 (27.3)	186 (21.2)	33 (43.2)		249 (23.4)
Does not drink	17 (15.5)	86 (9.9)	12 (16.0)		116 (10.9)

[a]Statistically significant.

Comparison of Increased Gambling in Different Outcomes

Multivariable analysis using binary logistic regression models of the potential predictors of increased gambling (yes vs no) is presented in Table 4 (unweighted data). Increased gambling was associated with all 3 categories of PGSI score, whether “Certain risk of gambling problems” (OR 4.40, 95% CI 2.17-8.93), “Increased risk of gambling problems” (OR 12.53, CI 6.54-24.04), or “Gambling problems” (OR 32.42, 95% CI 13.77-76.35). The PGSI for “Gambling problems” was also correlated with increased gambling in all specific forms of gambling: horses (OR 12.64, 95% CI 3.94-40.68), sports (OR 38.81, 95% CI 11.65-129.28), poker (OR 10.53, 95% CI 2.24-49.43), and casinos (OR 26.17, 95% CI 7.09-96.62). An increased change in gambling overall was associated with the >5 Kessler score (OR 2.62, 95% CI 1.39-4.91) and with increased casino gambling (OR 4.47, 95% CI 1.15-17.3). In addition, male gender is correlated with increased sports betting (OR 2.58, 95% CI 1.09-6.09) and casino gambling (OR 2.73, 95% CI 1.06-7.07). The correlation between “Does not drink” and increased sports betting was also significant (OR 3.95, 95% CI 1.02-15.33).

XSL•FO
RenderX

Table 4. Multivariable logistic regressions model with increased gambling as the outcome of interest (yes vs no).

	Increased gambling, odds ratio (95% CI)	Increased gambling on horses, odds ratio (95% CI)	Increased gambling on casino, odds ratio (95% CI)	Increased gambling on poker, odds ratio, (95% CI)	Increased gambling on sports, odds ratio (95% CI)
Gender					
Female	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Male	1.46 (0.86-2.48)	0.84 (0.39-1.81)	2.58[a] (1.09-6.09)	1.51 (0.50-4.59)	2.73[a] (1.06-7.07)
Age groups (years)					
16-24	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
≥25	1.09 (0.48-2.50)	2.09 (0.35-12.52)	0.43 (0.14-1.34)	0.91 (0.18-4.54)	0.67 (0.24-1.88)
Level of education					
Primary school	1.00 (reference)	1.00 (reference)	1.00 (reference)	—[b]	1.00 (reference)
Secondary school	0.73 (0.27-1.99)	0.80 (0.19-3.47)	1.62 (0.26-9.93)	—	2.39 (0.46-12.54)
University	1.18 (0.44-3.19)	0.85 (0.20-3.69)	1.69 (0.27-10.43)	—	2.82 (0.52-15.23)
Other	1.31 (0.27-6.36)	0.72 (0.05-9.64)	2.62 (0.15-45.97)	—	5.83 (0.38-89.42)
Employment status					
Studying	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Employed	0.61 (0.22-1.69)	0.38 (0.07-2.03)	0.41 (0.09-1.96)	0.49 (0.05-5.00)	0.50 (0.13-1.92)
Unemployed	1.79 (0.62-5.19)	1.59 (0.28-9.11)	2.84 (0.63-12.85)	0.89 (0.10-8.18)	0.86 (0.20-3.62)
Retired	0.64 (0.21-1.98)	0.73 (0.12-4.53)	0.51 (0.04-5.82)	—	0.79 (0.17-3.81)
Other	2.11 (0.37-12.06)	7.64 (0.70-83.40)	—	—	—
Disposable income					
≥SEK 40,000 (US $4657.62)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
SEK 20,000-40,000 (US $2328.81-$4657.62)	0.90 (0.42-1.95)	0.51 (0.19-1.36)	0.59 (0.15-2.33)	1.23 (0.27-5.58)	0.90 (0.28-2.89)
≤SEK 10,000-20,000 (US $1164.40-2328.81)	0.73 (0.26-2.04)	0.30 (0.08-1.21)	0.41 (0.06-2.65)	0.44 (0.04-5.01)	0.84 (0.17-4.09)
Problem Gambling Severity Index					
No problem with gambling	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Certain risk of gambling problems	4.40 (2.17-8.93)	4.01[a] (1.26-12.81)	4.54[a] (1.13-18.30)	7.49[a] (1.52-36.78)	6.14[a] (2.00-18.92)
Increased risk of gambling problems	12.5[a] (6.54-24.04)	10.93[a] (3.88-30.82)	6.02[a] (1.68-21.56)	3.84 (0.80-18.51)	10.91[a] (3.71-32.07)
Gambling problems	32.4[a] (13.77-76.35)	12.64[a] (3.93-40.68)	26.17[a] (7.09-96.62)	10.53[a] (2.24-49.43)	38.81[a] (11.65-129.28)
Kessler score					
Score 0-4; no psychological distress	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Score 5-24; psychological distress	2.62[a] (1.39-4.91)	2.06 (0.84-5.07)	4.47[a] (1.15-17.30)	2.80 (0.69-11.74)	1.09 (0.46-2.57)
Gambling pause					
Yes	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
No	1.14 (0.43-3.02)	0.77 (0.22-2.72)	1.66 (0.47-5.94)	0.92 (0.18-4.82)	2.48 (0.69-8.94)
Time at home					
Much more	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Slightly more	0.56 (0.30-1.04)	0.86 (0.36-2.05)	0.81 (0.29-2.20)	0.27 (0.07-1.12)	0.42 (0.16-1.07)
Unchanged	0.25 (0.08-0.77)	0.13 (0.01-1.18)	0.28 (0.05-1.76)	—	0.34 (0.08-1.49)

XSL•FO
RenderX

	Increased gambling, odds ratio (95% CI)	Increased gambling on horses, odds ratio (95% CI)	Increased gambling on casino, odds ratio (95% CI)	Increased gambling on poker, odds ratio, (95% CI)	Increased gambling on sports, odds ratio (95% CI)
Less time	0.78 (0.04-13.93)	—	—	—	—
Change in alcohol habits					
Yes more	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Unchanged	0.66 (0.30-1.44)	0.28 (0.09-0.86)	0.72 (0.18-2.81)	0.18[a] (0.03-0.98)	0.76 (0.24-2.40)
Less alcohol	0.77 (0.34-1.76)	0.53 (0.17-1.67)	1.45 (0.36-5.72)	0.24 (0.04-1.48)	0.62 (0.18-2.09)
Does not drink	1.31 (0.51-3.34)	1.22 (0.33-4.37)	1.32 (0.25-6.95)	2.07 (0.34-12.69)	3.95[a] (1.02-15.33)

[a]Significant correlations observed.

[b]—: not determined.

Discussion

Principal Findings

This cross-sectional study considers self-reported changes in web-based gambling in a general population aged 16 years and over during the third wave of the COVID-19 pandemic in Sweden. We set out to study possible changes in gambling behavior and relate them to risk factors such as gender, level of education, employment status, disposable income, alcohol consumption, and psychological distress.

The majority of the respondents who gambled reported no changes in their gambling activity during the COVID-19 pandemic. Regarding overall changes in gambling, we found significant associations with the PGSI (indicating the severity of the problem), the Kessler score (indicating psychological distress), employment status, changes in alcohol habits, and self-exclusion (gambling breaks). In the subgroup that reported an increased gambling activity, we found an association with both the PGSI score and the Kessler score. The PGSI score was also an independent predictor for all web-based gambling activities (horses, sports, poker, and casinos) whereas the Kessler score only had a statistically significant impact on changes in casino gambling. In addition, male gender was an independent predictor for sports betting and casino gambling.

Changes in Web-Based Gambling Activity During the Pandemic

In a previous study conducted during the first wave of the COVID-19 pandemic in 2020, we found a 4% self-reported increase in gambling [23], whereas in another study 7 months later, we found a 6% self-reported increase in gambling [22]. Surprisingly, Lindner et al [20], who also studied changes in gambling activity in Sweden during the first wave, found an overall decrease in gambling [20]. In this study conducted in March 2021 the majority (82%) of the respondents who reported that they gambled had not changed their gambling type during the pandemic, but we found a self-reported increase in gambling activity of 10%. Even though the studies are not fully compared because they use different subjects and this study is limited to web-based gambling only, one can nevertheless discern a consequent rise in gambling during the pandemic. This might be explained by people being more likely to change their behavior in the initial stages of the pandemic, similar to other times of crisis, but with time people tended to return to a stable, “normal” level, whether in gambling or other activities.

In the early stages of the pandemic, all sports betting events such as the top football leagues were cancelled [19]. Gradually, sporting events restarted, making sports betting easier. Another explanation for the increase in betting during the pandemic could be changes in society, such as restrictions on social events and restaurant opening hours, which implies that people were more likely to stay at home, where their screens—and gambling opportunities—were more readily available, thus affecting web-based gambling. The seemingly steady rise in self-reported gambling is something to follow as the pandemic continues.

Characteristics Among Those Reporting an Increase in Specific Types of Gambling

The PGSI, Kessler score (psychological distress), employment status, gambling breaks, and changes in alcohol habits were all individually associated with changes in gambling—something that was not true for gender, age group, level of education, disposable income, and changes in time spent at home. The only factors that remained significantly associated with an overall increase in gambling were self-reported gambling problems and psychological distress.

Our findings are in line with those of previous findings that the minority who reported increased gambling during the COVID-19 pandemic also reported increased gambling problems and greater psychological distress [18,20,23,33]. Problem gambling is known to be associated with mental health problems, adding to the impression that individuals who gambled more during the COVID-19 pandemic were a vulnerable subgroup of the population. Increased casino gambling was the only form of gambling associated with psychological distress. Casino gambling is considered potentially highly addictive and is associated with problem gambling and personal debt [30,34,35]. Notably, at the onset of the COVID-19 pandemic, both Lindner et al [20] and Håkansson [23] found that respondents who gambled using web-based casinos, rather than gambling less (as was the case for other types of gambling), increased their gambling activity.

Among respondents reporting an increase in sports betting and casino gambling, we found the traditionally male gender correlation [36,37]. This had not been true of sports betting in our previous study in April 2020, early in the COVID-19

XSL•FO
RenderX

pandemic [23]. At that stage the sports betting market had shrunk—a plausible explanation for not seeing any gender differences in self-reported increased sports betting.

As mentioned earlier, gambling is strongly associated with comorbid disorders and is also considered a potentially addictive behavior that may result in related harms [26,27,38]. Problem gambling is expensive for not only gamblers and their personal networks but also for society in terms of health care and legal costs, lost productivity owing to unemployment, and poorer quality of life [39].

Limitations

This study has some limitations. First, the data are self-reported, with the inevitable risk of recall bias. Respondents who gamble might be more likely to answer questions about it. Findings based on self-reporting would resist generalization were it not for the representative sociodemographic and geographic distribution. The advantage of our study setting is thus our use of a weighted sample: we continued to include respondents until we had a sample that was both geographically and sociodemographically representative. Another study limitation is that Sweden differs to most other countries in regard to its COVID-19 regulations [40]. There were no lockdowns and no shopping bans because policies to prevent transmission were centered on recommendations rather than regulations, making the results difficult to transfer to other settings [40].

Conclusions

Compared to our previous studies from earlier during the pandemic, we found that a higher proportion of the general population had increased their gambling activity during the COVID-19 pandemic. The group that reported increased overall gambling was small but characterized by gambling problems and psychological distress. The group that reported an increase in sports betting and casino gambling were predominantly male. The group that reported an increase in casino gambling were independently related to psychological distress.

It is concerning that those who reported increased gambling activity during the pandemic are a growing group. Moreover, they are a vulnerable group that needs to be addressed by caregivers, the gambling industry, and policy makers.

We have yet to see the full effects of the COVID-19 pandemic on gambling. More studies are needed to chart gambling patterns in possible subsequent waves of the pandemic and later in its aftermath.

Acknowledgments

The authors acknowledge Region Skåne, Yngre ALF (a Swedish government research grant), the Crafoord Foundation, the Fanny Ekdahl Foundation for Paediatric Research, and Svenska Spel's Research Council.

Authors' Contributions

All authors have contributed significantly and agree with the content of the paper.

Conflicts of Interest

AH is employed as a researcher in addiction medicine at Lund University sponsored by AB Svenska Spel (the Swedish state-owned gambling operator) as part of its responsible gambling policies. He also has research funding from Svenska Spel's Research Council, and from the research council of Systembolaget AB (the Swedish alcohol retailing monopoly) and from the Swedish Sports Confederation. ECK has research funding from Svenska Spel's Research Council. None of these organizations have been involved in or had any influence on any part of the present work.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, China Novel Coronavirus Investigating and Research Team. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med 2020 Feb 20;382(8):727-733 [FREE Full text] [doi: 10.1056/NEJMoa2001017] [Medline: 31978945]
2. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet 2020 Feb 29;395(10225):689-697 [FREE Full text] [doi: 10.1016/S0140-6736(20)30260-9] [Medline: 32014114]
3. Rajkumar RP. COVID-19 and mental health: A review of the existing literature. Asian J Psychiatr 2020 Aug;52:102066 [FREE Full text] [doi: 10.1016/j.ajp.2020.102066] [Medline: 32302935]
4. Holmes EA, O'Connor RC, Perry VH, Tracey I, Wessely S, Arseneault L, et al. Multidisciplinary research priorities for the COVID-19 pandemic: a call for action for mental health science. Lancet Psychiatry 2020 Jun;7(6):547-560 [FREE Full text] [doi: 10.1016/S2215-0366(20)30168-1] [Medline: 32304649]
5. Mansfield KE, Mathur R, Tazare J, Henderson AD, Mulick AR, Carreira H, et al. Indirect acute effects of the COVID-19 pandemic on physical and mental health in the UK: a population-based study. Lancet Digit Health 2021 Apr;3(4):e217-e230 [FREE Full text] [doi: 10.1016/S2589-7500(21)00017-0] [Medline: 33612430]
6. Öngür D, Perlis R, Goff D. Psychiatry and COVID-19. JAMA 2020 Sep 22;324(12):1149-1150. [doi: 10.1001/jama.2020.14294] [Medline: 32960247]

XSL•FO
RenderX

7. Columb D, Hussain R, O'Gara C. Addiction psychiatry and COVID-19: impact on patients and service provision. Ir J Psychol Med 2020 Sep;37(3):164-168 [FREE Full text] [doi: 10.1017/ipm.2020.47] [Medline: 32434597]
8. King DL, Delfabbro PH, Billieux J, Potenza MN. Problematic online gaming and the COVID-19 pandemic. J Behav Addict 2020 Jun;9(2):184-186. [doi: 10.1556/2006.2020.00016] [Medline: 32352927]
9. Javed A. Economic Impact of Coronavirus and Revival Measures: Way Forward for Pakistan. Asian Development Bank Institute. 2020. URL: https://think-asia.org/handle/11540/11771 [accessed 2021-10-26]
10. Colley RC, Bushnik T, Langlois K. Exercise and screen time during the COVID-19 pandemic. Health Rep 2020 Jul 15;31(6):3-11 [FREE Full text] [doi: 10.25318/82-003-x202000600001-eng] [Medline: 32672923]
11. Sultana A, Tasnim S, Hossain MM, Bhattacharya S, Purohit N. Digital screen time during the COVID-19 pandemic: a public health concern. F1000Res 2021 Feb 8;10:81. [doi: 10.12688/f1000research.50880.1]
12. Allen MS, Walter EE, Swann C. Sedentary behaviour and risk of anxiety: A systematic review and meta-analysis. J Affect Disord 2019 Jan 01;242:5-13. [doi: 10.1016/j.jad.2018.08.081] [Medline: 30170238]
13. Huang Y, Li L, Gan Y, Wang C, Jiang H, Cao S, et al. Sedentary behaviors and risk of depression: a meta-analysis of prospective studies. Transl Psychiatry 2020 Jan 22;10(1):26 [FREE Full text] [doi: 10.1038/s41398-020-0715-z] [Medline: 32066686]
14. Teychenne M, Costigan SA, Parker K. The association between sedentary behaviour and risk of anxiety: a systematic review. BMC Public Health 2015 Jun 19;15:513 [FREE Full text] [doi: 10.1186/s12889-015-1843-x] [Medline: 26088005]
15. Olason DT, Hayer T, Meyer G, Brosowski T. Economic Recession Affects Gambling Participation But Not Problematic Gambling: Results from a Population-Based Follow-up Study. Front Psychol 2017;8:1247 [FREE Full text] [doi: 10.3389/fpsyg.2017.01247] [Medline: 28790946]
16. Jiménez-Murcia S, Fernández-Aranda F, Granero R, Menchón JM. Gambling in Spain: update on experience, research and policy. Addiction 2014 Oct;109(10):1595-1601. [doi: 10.1111/add.12232] [Medline: 23718704]
17. Economou M, Souliotis K, Malliori M, Peppou LE, Kontoangelos K, Lazaratou H, et al. Problem Gambling in Greece: Prevalence and Risk Factors During the Financial Crisis. J Gambl Stud 2019 Dec;35(4):1193-1210. [doi: 10.1007/s10899-019-09843-2] [Medline: 31165324]
18. Brodeur M, Audette-Chapdelaine S, Savard A, Kairouz S. Gambling and the COVID-19 pandemic: A scoping review. Prog Neuropsychopharmacol Biol Psychiatry 2021 Dec 20;111:110389. [doi: 10.1016/j.pnpbp.2021.110389] [Medline: 34146652]
19. Auer M, Malischnig D, Griffiths MD. Gambling Before and During the COVID-19 Pandemic Among European Regular Sports Bettors: An Empirical Study Using Behavioral Tracking Data. Int J Ment Health Addict 2020 May 29:1-8 [FREE Full text] [doi: 10.1007/s11469-020-00327-8] [Medline: 32837423]
20. Lindner P, Forsström D, Jonsson J, Berman AH, Carlbring P. Transitioning Between Online Gambling Modalities and Decrease in Total Gambling Activity, but No Indication of Increase in Problematic Online Gambling Intensity During the First Phase of the COVID-19 Outbreak in Sweden: A Time Series Forecast Study. Front Public Health 2020;8:554542 [FREE Full text] [doi: 10.3389/fpubh.2020.554542] [Medline: 33117770]
21. Donati MA, Guido CA, De Meo G, Spalice A, Sanson F, Beccari C, et al. Gaming among Children and Adolescents during the COVID-19 Lockdown: The Role of Parents in Time Spent on Video Games and Gaming Disorder Symptoms. Int J Environ Res Public Health 2021 Jun 21;18(12):6642 [FREE Full text] [doi: 10.3390/ijerph18126642] [Medline: 34205533]
22. Håkansson A, Widinghoff C. Changes of Gambling Patterns during COVID-19 in Sweden, and Potential for Preventive Policy Changes. A Second Look Nine Months into the Pandemic. Int J Environ Res Public Health 2021 Feb 27;18(5):2342 [FREE Full text] [doi: 10.3390/ijerph18052342] [Medline: 33673575]
23. Håkansson A. Changes in Gambling Behavior during the COVID-19 Pandemic-A Web Survey Study in Sweden. Int J Environ Res Public Health 2020 Jun 05;17(11):4013 [FREE Full text] [doi: 10.3390/ijerph17114013] [Medline: 32516880]
24. Frisone F, Alibrandi A, Settineri S. Problem gambling during Covid-19. Mediterr J Clin Psychol 2020;8(3). [doi: 10.6092/2282-1619/mjcp-2457]
25. Kessler R, Hwang I, LaBrie R, Petukhova M, Sampson N, Winters K, et al. DSM-IV pathological gambling in the National Comorbidity Survey Replication. Psychol Med 2008 Sep;38(9):1351-1360 [FREE Full text] [doi: 10.1017/S0033291708002900] [Medline: 18257941]
26. Lorains FK, Cowlishaw S, Thomas SA. Prevalence of comorbid disorders in problem and pathological gambling: systematic review and meta-analysis of population surveys. Addiction 2011 Mar;106(3):490-498. [doi: 10.1111/j.1360-0443.2010.03300.x] [Medline: 21210880]
27. Potenza MN, Balodis IM, Derevensky J, Grant JE, Petry NM, Verdejo-Garcia A, et al. Gambling disorder. Nat Rev Dis Primers 2019 Jul 25;5(1):51. [doi: 10.1038/s41572-019-0099-7] [Medline: 31346179]
28. Yau YHC, Potenza MN. Gambling disorder and other behavioral addictions: recognition and treatment. Harv Rev Psychiatry 2015;23(2):134-146 [FREE Full text] [doi: 10.1097/HRP.0000000000000051] [Medline: 25747926]
29. Hing N, Cherney L, Gainsbury SM, Lubman DI, Wood RT, Blaszczynski A. Maintaining and losing control during internet gambling: A qualitative study of gamblers' experiences. New Media Soc 2014 Jan 27;17(7):1075-1095. [doi: 10.1177/1461444814521140]
30. Håkansson A, Widinghoff C. Over-Indebtedness and Problem Gambling in a General Population Sample of Online Gamblers. Front Psychiatry 2020;11:7 [FREE Full text] [doi: 10.3389/fpsyt.2020.00007] [Medline: 32116832]

XSL•FO
RenderX

31. Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, et al. Screening for serious mental illness in the general population. Arch Gen Psychiatry 2003 Feb;60(2):184-189. [doi: 10.1001/archpsyc.60.2.184] [Medline: 12578436]
32. Ferris J, Wynne H, Single E. Canadian Problem Gambling Index (CPGI). APA PsycTests. URL: https://doi.apa.org/doiLanding?doi=10.1037%2Ft00772-000 [accessed 2021-10-26]
33. Håkansson A. Impact of COVID-19 on Online Gambling - A General Population Survey During the Pandemic. Front Psychol 2020;11:568543 [FREE Full text] [doi: 10.3389/fpsyg.2020.568543] [Medline: 33101137]
34. Binde P, Romild U, Volberg RA. Forms of gambling, gambling involvement and problem gambling: evidence from a Swedish population survey. Int Gambl 2017 Aug 16;17(3):490-507. [doi: 10.1080/14459795.2017.1360928]
35. Håkansson A, Mårdhed E, Zaar M. Who Seeks Treatment When Medicine Opens the Door to Pathological Gambling Patients-Psychiatric Comorbidity and Heavy Predominance of Online Gambling. Front Psychiatry 2017;8:255 [FREE Full text] [doi: 10.3389/fpsyt.2017.00255] [Medline: 29238309]
36. Estévez A, Rodríguez R, Díaz N, Granero R, Mestre-Bach G, Steward T, et al. How do online sports gambling disorder patients compare with land-based patients? J Behav Addict 2017 Dec 01;6(4):639-647 [FREE Full text] [doi: 10.1556/2006.6.2017.067] [Medline: 29076354]
37. Abbott M, Binde P, Hodgins D, Korn D, Pereira A, Volberg R, et al. Conceptual Framework of Harmful Gambling: An International Collaboration. Guelph, ON: The Ontario Problem Gambling Research Centre (OPGRC); 2013.
38. Calado F, Griffiths MD. Problem gambling worldwide: An update and systematic review of empirical research (2000-2015). J Behav Addict 2016 Dec;5(4):592-613 [FREE Full text] [doi: 10.1556/2006.5.2016.073] [Medline: 27784180]
39. Hofmarcher T, Romild U, Spångberg J, Persson U, Håkansson A. The societal costs of problem gambling in Sweden. BMC Public Health 2020 Dec 18;20(1):1921 [FREE Full text] [doi: 10.1186/s12889-020-10008-9] [Medline: 33339531]
40. Baral S, Chandler R, Prieto RG, Gupta S, Mishra S, Kulldorff M. Leveraging epidemiological principles to evaluate Sweden's COVID-19 response. Ann Epidemiol 2021 Feb;54:21-26 [FREE Full text] [doi: 10.1016/j.annepidem.2020.11.005] [Medline: 33242596]

Abbreviations

OR: odds ratio
PGSI: Problem Gambling Severity Index
SEK: Swedish kronor

Edited by N Zary; submitted 14.07.21; peer-reviewed by I Shubina, A Tannoubi; comments to author 04.08.21; revised version received 10.08.21; accepted 26.09.21; published 03.11.21.

Please cite as:
Claesdotter-Knutsson E, Håkansson A
Changes in Self-Reported Web-Based Gambling Activity During the COVID-19 Pandemic: Cross-sectional Study
JMIR Serious Games 2021;9(4):e30747
URL: https://games.jmir.org/2021/4/e30747
doi:10.2196/30747
PMID:34730540

©Emma Claesdotter-Knutsson, Anders Håkansson. Originally published in JMIR Serious Games (https://games.jmir.org), 03.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Predicting Video Game Addiction Through the Dimensions of Consumer Video Game Engagement: Quantitative and Cross-sectional Study

Amir Zaib Abbasi[1,2*], PhD; Umair Rehman[3*], PhD; Zahra Afaq[4], MPhil; Mir Abdur Rafeh[1], MPhil; Helmut Hlavacs[5*], PhD; Mohammed A Mamun[6,7], HSC; Muhammad Umair Shah[8], PhD

[1]Department of Management Sciences, Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Islamabad, Pakistan

[2]Interdisciplinary Research Centers for Finance and Digital Economy, KFUPM Business School, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia

[3]User Experience Design Department, Wilfrid Laurier University, Brantford, ON, Canada

[4]Namal Institute Mianwali, Mianwali, Pakistan

[5]Entertainment Group, University of Vienna, Vienna, Austria

[6]Center for Health Innovation, Networking, Training, Action and Research-Bangladesh, Dhaka, Bangladesh

[7]Department of Public Health and Informatics, Jahangirnagar University, Savar, Bangladesh

[8]Department of Management Sciences, University of Waterloo, Waterloo, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Amir Zaib Abbasi, PhD
Department of Management Sciences
Shaheed Zulfikar Ali Bhutto Institute of Science and Technology
Street # 09, Plot # 67 Sector H-8/4
Islamabad, 46000
Pakistan
Phone: 92 514863363 ext 517
Email: aamir.zaib.abbasi@gmail.com

Abstract

Background: Video games are expanding exponentially with their increased popularity among users. However, this popularity has also led to an increase in reported video game addiction. There may be consumer engagement–related factors that may influence video game addiction.

Objective: This study aims to empirically examine the impact of the dimensions of consumer video game engagement on video game addiction. The dimensions are dedication, absorption, conscious attention, social connection, enthusiasm, and interaction. We utilize the uses and gratifications theory to study the video game engagement dimensions as potential factors through which gamers feel gratified and engaged in video game playing. Additionally, this study incorporates the cultivation theory to investigate how video game engagement factors trigger video game addiction.

Methods: A two-step process was applied for data analysis on valid cases of 176 gamers aged 15-25 years: video game addiction was specified and validated as a reflective-formative construct, and hypothesis testing was later performed using the WarpPLS on valid respondents.

Results: The analysis uncovered 2 dimensions of video game engagement: social connection with P=.08 and interaction with P=.49, which did not significantly contribute to video game addiction.

Conclusions: This study offers unique insights to a myriad of stakeholders, mostly psychologists and psychiatrists, who routinely prescribe behavior modification techniques to treat video game addiction.

(JMIR Serious Games 2021;9(4):e30310) doi:10.2196/30310

KEYWORDS

consumer video game engagement; dedication; absorption; social connection; interaction; conscious attention; enthusiasm; video game addiction; uses and gratifications theory; cultivation theory

XSL•FO
RenderX

Introduction

Background

Video games are one of the most popular forms of leisure activities, especially for youth (as well as other age groups) across the world [1]. The video gaming industry has come a long way since first video game, Atari, which was launched in 1972 [2]. Video game engagement has surpassed other forms of new media such as music and television owing to the rapid development of the digital gaming industry [3]. It has been estimated that digital gaming revenues rose to US $137.9 billion, with approximately 2.3 billion gamers worldwide by the end of 2018. With the rise of video gaming activities, the risk of video gaming addiction has also increased [4]. Excessive use of web-based and offline video gaming affects the gamers' physical health and has an impact on their psychological well-being [5]. Video game addiction can cause gamers to develop a range of issues such as aggression and depression [6,7], poor academic performance [8,9], declining relations with friends and families [10,11], insomnia [12], and even self-harm [13,14].

The literature on video game engagement has identified various factors that affect video gaming addiction. Studies on demographics have identified that males, young age, singles (relationship), etc are more likely to develop video game addiction [15]. Increased monetary spending on video gaming activities is positively correlated with video game addiction alongside other factors such as the presence of heightened family resentment among gamers and an increase in the average time spent on gaming activities [16]. With respect to psychographics, the factors that predict video game addiction are depression and sociability [17], hypercompetitiveness and psychological absorption [18], and dysfunctional impulsivity [19], thereby reducing extraversion, agreeableness, emotional stability, and attractiveness [20]. Increased video game engagement and maladaptive coping has helped identify individuals at risk of transitioning into video game addiction [21].

Likewise, the phenomenon of video game engagement has been investigated in different studies across multiple application areas. For instance, Abbasi et al [22] reported how gender differences variably influence video game engagement. Gabriel et al [23] reported that rich game experiences enhance engagement in video games. Additionally, Skoric et al [24] revealed that more addicted gamers perform poorly in academic assessments. Another study found that video game addiction affects mental health, causing depression, stress, and anxiety, but video game engagement is only linked to anxiety [21]. Video game engagement has been found to have no adverse effects on gamers' well-being [25,26], and several studies have differentiated addiction and engagement since they both result in different behavioral outcomes [27,28]. Although there are many studies on video game addiction, the research investigating the potential determinants of video game addiction through consumer video game engagement elements (eg, dedication, enthusiasm, conscious attention, absorption, social connection, interaction) remains limited.

Study Aims and Rationale

The purpose of this study was to expand the existing research by evaluating video game addiction from the lens of video game engagement. Our primary focus in this study was to evaluate whether different dimensions of consumer video game engagement play a role in predicting video game addiction. Our study will contribute to the literature in a number of ways. First, prior studies have not accounted for engagement factors that lead to addiction [1,28,29]. Our study adds value to this overarching literature as we uncover how engagement factors influence addiction in video games. Second, this study extends the application of Uses and Gratifications Theory (UGT) as studied in media, video game, and social media [30,31] and Cultivation Theory toward developing a theoretical model that explains the role of video game engagement dimensions in predicting video game addiction. We applied UGT to video game engagement factors (conscious attention, absorption, social connection, dedication, enthusiasm, and interaction), which give rise to a cultivation effect owing to increased engagement in video game play, resulting in video game addiction. Third, prior studies considered video game addiction as a second-order construct, but its specification and validation as a reflective-formative construct remains nebulous [32]. Considering this limitation, we add to video game addiction studies literature by specifying and validating video game addiction's construct as a reflective-formative construct.

Theoretical Underpinning

We explored the theoretical framework informing our understanding of the different dimensions of consumer engagement and video game addiction. Different theories associated with motivation, communication, and media were used to establish the theoretical framework for this research. This research takes its theoretical underpinning from the Cultivation Theory and UGT [33,34].

UGT

The principle of UGT was coined in the early 1940s. This theory came into the light to answer some questions raised concerning traditional mass communication. This theory emphasizes that media is used to satisfy a user's various needs, which motivates them to use that media [35,36]. This theory has been extensively applied to comprehend the motivation underlying media usage [37]; particularly, video games [38]. UGT is one of the most popular theories applied to assess the underlying motivation of users that engage in video gameplay [39]. With the high emphasis on video game engagement, UGT could provide a new perspective toward understanding video game addiction.

UGT consists of 3 constructs: (1) achievement, (2) enjoyment, and (3) social interaction. Achievement in the context of gaming can be defined as the yearning to achieve power and rewards as well as the desire to perform better in comparison to other gamers [40]. Sherry et al [39] found that challenge is the biggest motivation for playing video games as players are tempted to move through different stages quickly to complete the challenges [41]. In contrast, hardcore gamers who play video games with dedication [42] are inclined to preoccupy themselves with social comparison to flaunt their gaming achievements [43]. Enjoyment

XSL•FO
RenderX

in the context refers to inherent pleasure derived from accomplishing the tasks [40], and it plays a major role when it comes to assessing an individual's desire to use a specific information system [44]. Social interaction in web-based gaming opens new channels of communication since gamers have an opportunity to socialize and cultivate existing or new relationships. Web-based games, therefore, provide an avenue to gamers, especially those who experience loneliness and low self-esteem, thus supporting individual desire to play video games [45]. Similarly, the UGT perspective accentuates the self-indulgent needs and interests that influence the motivation to engage in digital games [46]. Thus, UGT appears to be a relevant theory for this research since video game players actively indulge in video games, thereby leading to video game addiction.

Cultivation Theory

The explanation of mass media as cultivation was introduced by Gerbner [47]. The Cultivation Theory uncovers how mass media shapes the perceptions of people [48]. The Cultivation Theory suggests that exposure to media influences viewers' interpretation of reality [49]. Research has already demonstrated this phenomenon by investigating individuals' social settings and observing changes in people's attitudes and social norms [50,51]. Therefore, the media carries the capacity to enculturate the masses to adapt to the changing environment around them [52].

The Cultivation Theory provides a wide lens through which long-term effects of media can be examined. The Cultivation Theory has mostly been applied to traditional forms of media such as television. Aligned with the Cultivation Theory assumptions with respect to television (similarly, if this concept is applied to video gaming), this theory posits that through continuous exposure to the video gaming world, player's views of his real-world will become more similar to that of his gaming world [53]. For that purpose, Van Mierlo and Van den Bulck [54] argued that video games have become so authentic that they have started to mirror reality, thereby making cultivation possible [34]. Hence, video game players engage with different forms of synthetic media such as in-game challenges, stories, and characters, and this involvement in the video games is associated with concentrated attention [23], because of which the game player identifies himself with the role-played character and creates a visually epitomized virtual identity [55]. Thus, video games have started resembling reality, enabling the cultivation of an immersive reality [56]. Seah and Cairns [57] have also reported that the more immersed a gamer becomes, the immersion may lead to video game addiction.

Consumer Video Game Addiction

The massive popularity of video games has significantly changed the gaming industry as it is one of the most booming entertainment businesses [58]. Several studies have documented the risks and effects of video game addiction. A recent meta-analytic review, including 136 articles with 130,000 participants from both Eastern and Western cultures, found that video game addiction is likely to cause anxiety, depression, social phobia, and scholastic decline [59]. The study suggests that harmonious passion and obsessive passion predict different ways to engage in massively multiplayer online role-playing game (MMORPG). The study further confirms that passion is a valuable paradigm to recognize diverse motivational patterns expressed by MMORPG players [60]. Wittek et al [61] posit that video game addiction has a negative relationship with conscientiousness and a positive association with neuroticism. Hull et al [17] report that the social features of video gaming, distortion of time perceptions, and happiness levels significantly predict video game addiction. Apart from this, video game addiction is also reported to be associated with lower psychological functioning, unsatisfactory academic performance, increased alcohol consumption, sleep deprivation, aggression and depression, social disconnectedness, and increased caffeine consumption [17,62-64].

Consumer Video Game Engagement

Consumer engagement is the interaction between a customer and an organizational entity through different channels whereas engagement in consumer game-playing is a concept that signals task-oriented consumer behavior in the game setting [65]. Video games are purposely intended to draw the attention of its consumer base to ensure that video game consumers will evolve into frequent game players [66,67]. This appeal is primarily engendered by the engaging experiences video games are intended to deliver [68]. These experiences are coined as playful video game consumption experience and have led to behavioral changes among its consumer base [69].

Previous studies deem engagement as a multidimensional construct that comprises of subconstructs such as immersion [70,71], presence [72], flow [73], fun [74], enjoyment [72,75], fun [74], and absorption [76]. Hollebeek et al [77] also suggest that the construct *engagement* is a multidimensional concept comprising of 3 dimensions: (1) cognitive, (2) emotional, and (3) behavioral. Researchers also accentuate that different dimensions of consumer engagement are interrelated [78]. For instance, Brodie et al [79] identified that emotional involvement can increase other video game engagement dimensions (ie, cognitive and behavioral).

Bouvier et al [41] identified 4 types of gaming engagement behaviors, that is, (1) environment-directed behavior, (2) social-directed behavior, (3) self-directed behavior, and (4) action-directed behavior. They further explain that environment-directed behavior refers to the participation of players oriented toward the setting or the ecological backdrop in the game. In contrast, social participation alludes to the social links present in the game. Games offer players the opportunity to develop and foster their social relationships with other players, thereby establishing social links in the virtual environment. Self-directed behavior refers to the relationship between players and their virtual character. For instance, we often see players have a preference to personalize their virtual presence such as customizing their avatar or opting for accessories for reasons other than performance. Lastly, action-directed behavior refers to goal-directed measures taken by players in different in-game situations. Players act for many reasons such as seeking to quickly move between stages, gain different experiences, and accomplish challenges.

XSL•FO
RenderX

Abbasi et al [22] report multidimensionality within video game engagement and demonstrates that the construct comprises of 6 dimensions, that is, (1) dedication, (2) absorption, (3) conscious attention, (4) social connection, (5) enthusiasm, and (6) interaction. This study aims to determine which of these dimensions can result in addiction to video games, which may provide empirical evidence in this sector.

Game Addiction and Game Engagement: Are They Similar or Distinct?

Video game addiction and video game engagement were previously used interchangeably [24]. However, recently, these 2 terms have been considered as 2 separate entities [15]. Distinguishing gaming addiction from gaming engagement has been a challenge for researchers [15]. Several studies have approached this concern by differentiating game addiction from game engagement; first, Charlton and Danforth [80] showed that a gaming addict group spends twice the hours as a highly engaged gamers group in playing games each week. In another study, game addiction scores were negatively related to emotional stability; however, gaming engagement scores remained unchanged [20]. Charlton and Danforth [80] also reported a considerable difference between video game addiction and video game engagement. They posited that addiction should not be confused with active engagement as addiction includes withdrawal symptoms, that is, anxiety and guilt. Brunborg et al [15] also distinguished video game addiction from video game engagement through a self-administered questionnaire. In their study, engaged gamers showed salience and mood modification, whereas addicted gamers exhibited anxiety, apprehension, touchiness, and isolation. The key differentiation between addiction and engagement included the range of adverse consequences experienced by the consumers [81,82]. Video game addiction is possibly linked to a range of adverse effects such as mental, physical, and social deterioration, but this is not the case in high engagement [82]. Video game addiction is mainly linked with growing signs of stress, anxiety, and depression [83].

Dimensions of Consumer Video Game Engagement

Dedication is the affective commitment of players toward video games [84]. Kallio et al [85] reported that some gamers play for pleasure, whereas others play to form a connection with the game. Players who are committed and loyal spend more time playing social games, ensuring that they are dedicated to the video game [85]. Similarly, it has been found that hardcore gamers on average play more than an hour per day [86,87], spending most of their leisure time dedicated to gaming activities [42,88]. On the contrary, casual gamers spend less time playing games since they play in small bursts and have a casual attitude toward gaming [89]. Conrad [90] found that games tended to be addictive if played for more extended periods. Griffiths [91] assessed the consequences of excessive gaming and found that players experienced behavioral addiction symptoms, including salience, mood modification, and tolerance. Van Rooij et al [92] reported that teenagers who spend an average of 55 hours per week on gaming tended to develop depressive moods, loneliness, social anxiety, and negative self-esteem. On the basis of the abovementioned relationships, we posit the following hypothesis:

> *H1: Dedication positively predicts video game addiction in a video game player.*

Absorption is defined as the degree of involvement and immersion in a given activity [93]. Csikszentmihalyi [94] reports that absorption in an activity is regarded as "flow experience," and video games encourage a state of flow and learning in gamers [95]. Players in this state immerse themselves and become absorbed in their activity. This way, they become oblivious to their surroundings, which narrows their focus of situation awareness. De Pasquale et al [96] reported that excessive absorption of video games leads to neglecting the surrounding environments. Furthermore, when trying to curb video game playing, absorption has been correlated with anxiety, irritability, and emotional fragility. Excessive video game absorption is unhealthy and can potentially lead to addiction [57,97]. Based on these relationships, we posit the following hypothesis:

> *H2: Absorption in video games significantly predicts video game addiction in a video game player.*

Conscious attention is the attention and time allocated toward a given activity [98-100]. Consciousness is a "complex system that has evolved in humans for selecting information from this profusion, processing it, and storing it" and determines "what to pay attention, how intense, and for how long" [101]. In the state of flow, an individual experiences conscious processes followed by attentional processes, which require concentration and unilateral focus [101]. The individual operates at full capacity [102], devotes full attention to achieve goals [94], and goes through the distortion of temporal experiences [101]. Haladjian and Montemayor [103] report that conscious attention lets us intermingle with our complicated surroundings and enables us to remain involved in the task at hand. Similarly, consistent, conscious attention over a considerable period makes the player more absorbed. Therefore, conscious attention and video game consumption carry addictive tendencies and lead to addiction [57,97]. We further posit the following hypothesis:

> *H3: Conscious attention predicts video game addiction in a video game player.*

Social connection is the relationship people have with other individuals or groups [104]. Web-based gamers may shape meaningful relationships through social connections made during interactive gameplay [105]. Digital games can often create unrestrained social environments where players connect and form connections [106,107]. According to the augmentation hypothesis, people with an active social life also endeavor to build social ties in the virtual world [108]. Sioni et al [109] reported that video game addiction is very likely to occur owing to social phobia. Cole and Griffiths [110] suggested that people often try to develop long-term friendships based on their gaming experiences, and this socialization aspect tends to keep them motivated to continue playing. On the contrary, Peters and Malesky Jr [111] reported that gamers looking for social connections in gaming contexts may struggle to form tangible social connections outside gaming environments. It has additionally been found that people who cannot develop real-life

XSL•FO
RenderX

connections tend to develop web-based virtual connections, and people who have social anxiety as well as depression also tend to use web-based media, which has been associated with addictive symptoms [15]. Accordingly, we put forward the following hypothesis:

H4: Social connections significantly predict video game addiction in a video game player.

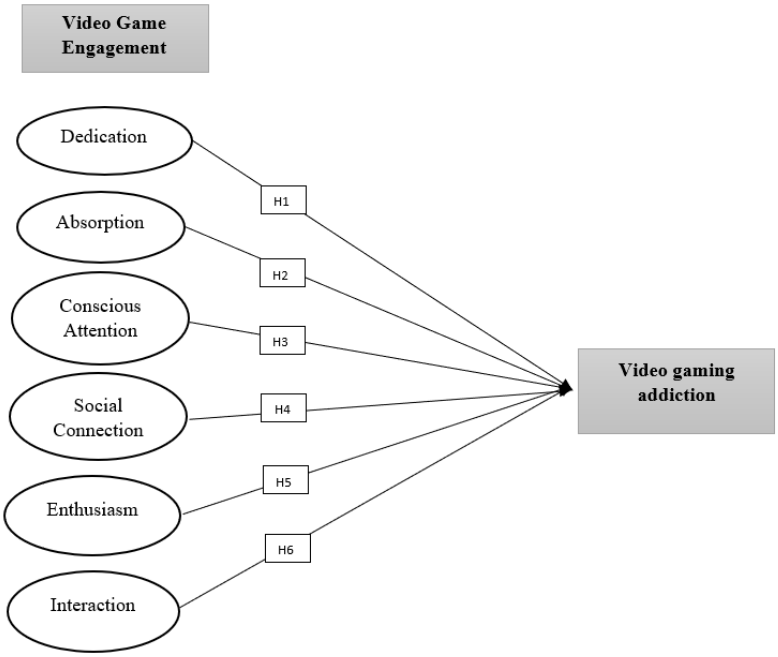
Enthusiasm is defined as the extent of eagerness and interest toward a given activity [104,112]. Enthusiasm is the inner driver that motivates individuals to play video games. Enthusiasm encourages individuals to overcome obstacles to reach higher levels in games [113]. Griffiths [91] adds that one can differentiate between enthusiasm and addiction since enthusiasm contributes to life, whereas addiction takes away life. Although enthusiasm and addiction can most certainly be different phenomena and experiences, it can be fair to conclude that excessive enthusiasm for a particular video game can potentially lead to addiction. Hence, we put forward the following hypothesis:

H5: Enthusiasm significantly predicts video game addiction in a video game player.

Interactions reflect the degree of involvement that individuals and other entities have in virtual game settings. So et al [112] report that interactions signify how players interact with synthetic game components. Interaction is positively correlated with consumer video game engagement [114] since interactive experiences influence the degree of consumer engagement [115,116]. Weibel et al [107] reported that gamers actively seek interaction with other players; this social element leads to increased duration of gameplay and heightened engagement levels. This has the potency to result in video game engagement [117]. Griffiths [91] reports that excessive interaction between human-machine components can influence different aspects of addiction, including mood, tolerance, withdrawal symptoms, etc. Similarly, it has been found that social interaction is one of the triggers of addiction [118]. Thus, we put forward the following hypothesis (see Figure 1 for the full study model and hypotheses):

H6: Interaction positively predicts video game addiction in a video game player.

Figure 1. Study conceptual model.



Methods

Study Design

This study was quantitative and cross-sectional. This approach is relatively fast in obtaining responses to develop or validate new theories [119]. Following this approach, we required a survey to collect the data (Multimedia Appendix 1). Hence, we designed the survey into 2 units: the first unit of the questionnaire included the participants' demographic profiles regarding their video gaming consumption patterns. Questions in the second unit were related to the constructs of the theoretical model, including engagement and addiction. As previously mentioned, video gaming engagement comprised 6 dimensions, whereas video gaming addiction was formulated as 7 dimensions of the second-order construct. The items measuring the video game engagement dimensions were adapted [22,28,120]. The items assessing the construct of video game addiction were adapted [121]. The questionnaire items were evaluated on a 5-point Likert Scale from 1 (strongly disagree) to 5 (strongly agree). Hair Jr et al [122] recommended applying the G*power analysis tool when the sample frame is unknown. Using the power analysis, we set the input parameters as follows: effect size (f^2)=0.15, α error probability=.05, power (1-β error probability)=.95, and predictors=6, which calculated the total sample size of 146 for the study model to conduct partial least squares-structural equation modeling (PLS-SEM) analyses. To sample the participants, we considered purposive sampling owing to its popularity in social sciences and the ability to obtain a sample that can represent the population [123].

XSL•FO
RenderX

This study was carried out among gamers who play in gaming zones (a gaming zone is a venue or an entertainment service provider, which facilitates users to play competitive games in groups to compete with each other) [124] or in organized tournaments at institutions located in the twin cities of Islamabad and Rawalpindi. The only inclusion criterion was if they play video games (eg, Counter Strike–Global Offensive, PlayerUnknown's Battlegrounds, Call of Duty, Fortnite) daily for a minimum of 1 hour. The main reason for selecting 1 hour as the minimum criterion for sample selection was that playing video games in gaming zone/cafe is an expensive activity in low-income countries such as Pakistan because gamers are charged about PKR 300-500 (US $1.5-$3) per hour/session, whereas the per capita income in Pakistan is only US $1516. Second, parents are authoritarian in Pakistan and they may limit their children from longer screen time as the age bracket from 15 to 20 years is considered as the prime age for education and career development [125]. Third, gamers older than 21 years could be engaged in some jobs with their studies to meet their needs and wants.

While collecting the data, we carefully recruited gamers who played Counter Strike–Global Offensive, PlayerUnknown's Battlegrounds, Call of Duty, and Fortnite and proceeded for data collection. Using the self-administered approach, we distributed 250 questionnaires and 190 were returned. We assessed the data quality and removed incomplete and biased responses (ie, using the particular response pattern) (Multimedia Appendix 2). Finally, we had 176 valid responses to continue with data analyses; see Table 1 for participants' details.

The structural equation modeling approach was applied for analysis, which is a well-rounded multivariate statistical analysis approach [126,127]. Structural equation modeling appears to be the most suitable approach because the research variables consist of both formative and reflective constructs [128]. For PLS-SEM analysis, WarpPLS software by Kock [129] was employed in this research.

XSL•FO
RenderX

Table 1. Characteristics of the participants (N=176).

Characteristics	Values, n (%)
Gender	
Male	128 (72.7)
Female	48 (27.3)
Age (years)	
15-20	53 (30.1)
21-25	123 (69.9)
Education	
Higher secondary education	45 (25.6)
Bachelor's	65 (36.9)
Master's	66 (37.5)
Expenditures on video games daily (exchange rate: 1 USD=PKR 153.314 in December 2019)	
PKR 0-500	88 (50)
PKR 500-1000	58 (33)
PKR 1000-5000	20 (11.4)
Above PKR 5000	10 (5.7)
Device used for gaming	
Play station	59 (33.5)
Cybercafé	45 (25.6)
Xbox	26 (14.8)
Others	46 (26.1)
Games played the most	
Counter Strike–Global Offensive	55 (31.3)
PlayerUnknown's Battlegrounds	29 (16.5)
Call of duty	68 (38.6)
Fortnite	24 (13.6)
Time spent on video games daily (hours)	
0-1	72 (40.9)
1-2	65 (36.9)
2-4	39 (22.2)

Assessing the Reflective Constructs

In the measurement model's assessment, first, the reflective constructs were assessed in which the construct's outer loadings, Cronbach alpha, composite reliability, and convergent validity were evaluated. According to the criteria, the outer loadings should be greater than at least 0.40, the Cronbach alpha value should be greater than .70, the composite reliability should be greater than 0.70, and the convergent validity should be greater than 0.5 [126,128]. The values for all the measures met the benchmark, thereby ensuring the reliability and the validity of the constructs; hence, the measurement model of the reflective measures was satisfactorily fulfilled (Table 2).

Additionally, the discriminant validity was evaluated applying the Fornell and Larcker (1981) criterion. The criterion suggests that the square root of the average variance extracted of each dimension must be larger than its corresponding correlation coefficient. Heterotrait-monotrait ratio of correlations is another criterion that is the most recent and recommended validity test to assess the discriminant validity for reflective constructs [130]. The heterotrait-monotrait ratio of correlations values are deemed to be satisfactory if lower than 0.90 and best if lower than 0.85. The results in Table 3 and Table 4 reported that discriminant validity had attained satisfactory values.

XSL•FO
RenderX

Table 2. Measurement model: first-order constructs.

Scale, items	Loading	Composite reliability	Cronbach alpha	Average variance extracted
Conscious attention		0.935	.916	0.707
Item 1	0.84			
Item 2	0.872			
Item 3	0.855			
Item 4	0.902			
Item 5	0.842			
Item 6	0.724			
Absorption		0.93	.906	0.728
Item 1	0.838			
Item 2	0.866			
Item 3	0.895			
Item 4	0.863			
Item 5	0.802			
Dedication		0.928	.903	0.72
Item 1	0.851			
Item 2	0.863			
Item 3	0.856			
Item 4	0.842			
Item 5	0.83			
Enthusiasm		0.893	.85	0.628
Item 1	0.688			
Item 2	0.832			
Item 3	0.844			
Item 4	0.816			
Item 5	0.771			
Social connection		0.9	.833	0.75
Item 1	0.865			
Item 2	0.871			
Item 3	0.862			
Interaction		0.93	.906	0.727
Item 1	0.832			
Item 2	0.878			
Item 3	0.866			
Item 4	0.858			
Item 5	0.827			
Salience		0.91	.851	0.771
Item 1	0.873			
Item 2	0.872			
Item 3	0.889			
Tolerance		0.893	.82	0.736
Item 1	0.841			
Item 2	0.894			

XSL•FO
RenderX

Scale, items	Loading	Composite reliability	Cronbach alpha	Average variance extracted
Item 3	0.837			
Mood modification		0.902	.836	0.753
Item 1	0.859			
Item 2	0.862			
Item 3	0.883			
Relapse		0.913	.857	0.778
Item 1	0.879			
Item 2	0.897			
Item 3	0.87			
Withdrawal		0.897	.828	0.745
Item 1	0.831			
Item 2	0.908			
Item 3	0.848			
Conflict		0.912	.855	0.775
Item 1	0.873			
Item 2	0.896			
Item 3	0.872			
Problem		0.952	.924	0.869
Item 1	0.928			
Item 2	0.958			
Item 3	0.909			

Table 3. Discriminant validity using Fornell and Larcker criterion.

Dimension	Conscious attention	Absorption	Dedication	Enthusiasm	Social connection	Interaction	Salience	Tolerance	Mood modification	Relapse	Withdrawal	Conflict	Problem
Conscious attention	0.841												
Absorption	0.500	0.853											
Dedication	0.402	0.538	0.849										
Enthusiasm	0.439	0.496	0.544	0.792									
Social connection	0.313	0.440	0.275	0.601	0.866								
Interaction	0.487	0.454	0.405	0.573	0.663	0.853							
Salience	0.402	0.510	0.462	0.566	0.527	0.565	0.878						
Tolerance	0.336	0.456	0.373	0.479	0.435	0.435	0.614	0.858					
Mood modification	0.323	0.310	0.239	0.403	0.300	0.322	0.395	0.461	0.868				
Relapse	0.248	0.379	0.288	0.401	0.288	0.265	0.499	0.413	0.587	0.882			
Withdrawal	0.276	0.334	0.334	0.304	0.221	0.229	0.397	0.424	0.535	0.736	0.863		
Conflict	0.365	0.201	0.248	0.212	0.146	0.133	0.241	0.291	0.447	0.392	0.535	0.88	
Problem	0.265	0.357	0.353	0.296	0.195	0.217	0.333	0.415	0.286	0.366	0.374	0.426	0.932

XSL•FO
RenderX

Table 4. Discriminant validity using heterotrait-monotrait ratio of correlations.[a]

	Conscious attention	Absorption	Dedication	Enthusiasm	Social connection	Interaction	Salience	Tolerance	Mood modification	Relapse	Withdrawal	Conflict	Problem
Conscious attention													
Absorption	0.550												
Dedication	0.441	0.596											
Enthusiasm	0.499	0.571	0.625										
Social connection	0.357	0.507	0.316	0.715									
Interaction	0.537	0.504	0.448	0.649	0.763								
Salience	0.455	0.580	0.526	0.668	0.625	0.645							
Tolerance	0.387	0.529	0.433	0.576	0.525	0.504	0.733						
Mood modification	0.371	0.355	0.275	0.480	0.359	0.370	0.468	0.558					
Relapse	0.278	0.430	0.327	0.470	0.341	0.301	0.584	0.492	0.692				
Withdrawal	0.314	0.387	0.385	0.363	0.267	0.268	0.475	0.516	0.647	0.876			
Conflict	0.410	0.227	0.282	0.254	0.173	0.152	0.281	0.347	0.529	0.458	0.634		
Problem	0.290	0.390	0.386	0.339	0.221	0.237	0.374	0.476	0.325	0.411	0.427	0.480	

[a]Heterotrait-monotrait ratio of correlations: good if <0.90; best if <0.85.

Assessing the Formative Constructs

To create the second-order formative construct, that is, video game addiction, we employed the two-stage approach of Becker et al [131] to calculate the latent variable scores of first-order reflective constructs of salience, problem, withdrawal, mood modification, relapse, conflict, and tolerance. Afterwards, the latent variables scores were used to create the video game addiction construct as specified in Figure 1 and Figure 2. The quality checks for assessing the formative construct were quite different from those for assessing reflective constructs. Assessing the formative construct for reliability and validity, researchers suggest to use variance inflation factor (the values should be <5 or <3.3, the most restrictive one) and calculate the weights and its significance [29,132]. We analyzed the formative construct and reported the results in Table 5, thereby showing the formative construct's validity. It met the threshold values, that is, variance inflation factor values are lower than 3.3 and indicator weights are significant.

XSL•FO
RenderX

Figure 2. The empirical-based study model.

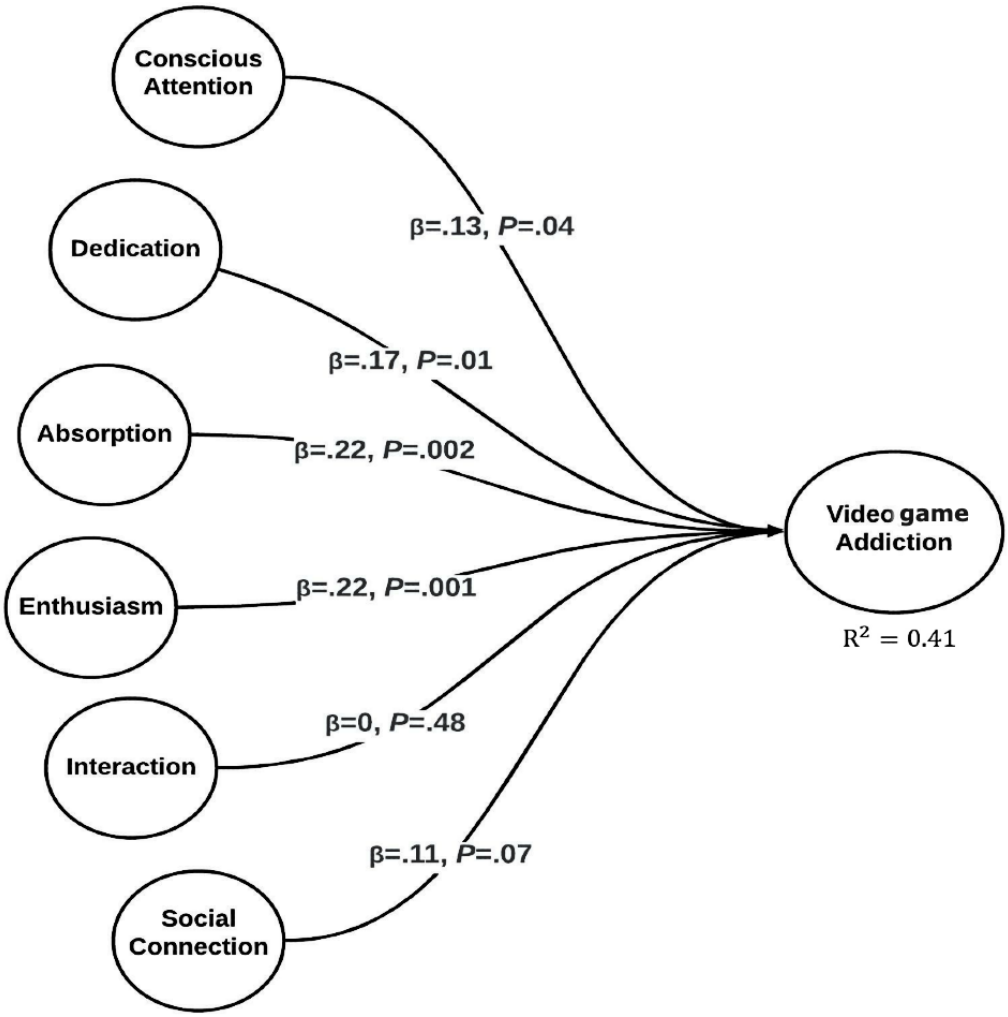


Table 5. Evaluation of the formative measurement model on the second-order constructs of video game addiction.

Items	Scale type	Weights	Significance	Full collinearity
Salience	Formative	0.188	.005	2.383
Tolerance	Formative	0.195	.004	1.987
Mood modification	Formative	0.205	.003	1.892
Relapse	Formative	0.223	.001	2.867
Withdrawal	Formative	0.222	.001	2.744
Conflict	Formative	0.178	.008	1.780
Problems	Formative	0.167	.01	1.485

Results

PLS-SEM analysis was conducted in 2 stages comprising the measurement model and structural model assessment. Once the reliability and validity of the reflective and formative constructs were ensured, researchers recommended evaluating the structural model through path coefficients, effect size, *t* value and calculated the predictive relevance, that is, Q^2 and R^2 coefficient value for the endogenous constructs [28,29,128]. WarpPLS was used to validate the hypotheses. For H1, the analysis results revealed a significant relation with B=0.129, *P*=.04, and *t* statistics of 1.76 with f^2=0.056. Hence, H1 was supported. For H2, the results revealed a significant impact of absorption on the consumer video game addiction with B=0.215, *P*=.002, and *t* statistics of 2.98 with f^2>=0.113. For H3, the results were supported with B=0.166, *P*=.01, and *t* statistics of 2.27 with f^2=0.079. For H4, the hypothesis was also supported with B=0.220, *P*=.001, and *t* statistics of 3.05 with f^2=0.118. For H5, the analysis results revealed that the relationship between both the constructs was insignificant with B=0.107, *P*=.08, and *t* value of 1.44. Hence H5 was rejected. Similarly, interaction with consumer video game addiction was also insignificant with B=–0.003, *P*=.49, and *t* value of 0.04. We also calculated the

XSL•FO
RenderX

R^2 and Q^2 for the endogenous construct, that is, video game addiction, which resulted in R^2=0.409 and Q^2=0.413. Our results showed that the study model has sufficient explanatory capacity and predictive relevance. Figure 2 and Table 6 show the results of the structural model assessment.

Table 6. Structural model assessment.

Dimensions of video game addiction	B	SE	*t* value[a]	*P* value	F^2
Conscious attention	0.129	0.073	1.76	.04	0.056
Absorption	0.215	0.072	2.98	.002	0.113
Dedication	0.166	0.073	2.27	.01	0.079
Enthusiasm	0.220	0.072	3.05	.001	0.118
Social connection	0.107	0.074	1.44	.08	0.045
Interaction	–0.003	0.075	0.04	.49	0.001

[a]Degree of freedom not provided because WarpPLS software was used.

Discussion

Principal Findings

In this study, we intended to empirically test the theoretical model illustrating the different dimensions of video game engagement, that is, (1) dedication, (2) absorption, (3) conscious attention, (4) social connection, (5) enthusiasm, and (6) interaction in a digital game, that predict consumer video game addiction. The underlying constructs were supported by theoretical underpinnings from the Cultivation Theory [47] and UGT [37]. Our results indicated that 4 out of 6 dimensions (ie, dedication, absorption, conscious attention, and enthusiasm in consumer video game engagement) significantly predict consumer video game addiction; however, social connection and interaction in consumer video game engagement had no significant effect on video game addiction.

Previous research suggests various personal, emotional, and psychological needs that contribute to overindulgence in video gameplay [133,134]. It has also been reported that gaming environments provide safe and interactive avenues for individuals who experience social disconnection or isolation in real-life environments [15]. This is also a means of open interaction for such individuals since the consequences for actions in gaming environments are minimal; therefore, the fear of interaction subsides [107]. Peer pressure in social groups is also a significant reason for excessive indulgence in video games. The media hype of video games have led some people to assume that behavioral problems are triggered by video games and contribute to video game addiction [135]. Further, rigorous advertising on different media plays a role in convincing individuals to play video games and experience instant gratification and other psychological needs. Cultivation Theory unravels the grounds explaining why individuals are addicted to video gaming, primarily because increased exposure to video games carries negative consequences. Such results are consistent with those reported in previous literature, where addiction contributes to different problems in the individual's personality and social life [136].

H1 showed the relationship between dedication and the resulting video game addiction. The results supported that there is a positive relationship between the 2 constructs. They showed that dedicated players (and attached) to the gameplay for long hours put in more energy, commitment, and loyalty than casual players [85,137]. H2 showed the relationship between absorption and video game addiction. Our results supported the positive relationship between the 2 constructs—consistent with that reported in previous literature [138]. They indicated that absorption leads to deep involvement in the game, which results in neglecting the surrounding environments and resulting in various problematic behaviors [96]. H3 showed the relationship between conscious attention and video gaming addiction. The results supported this testable statement (showed a positive effect between conscious attention and video gaming addiction). Through conscious attention, players tend to develop deep concentration [101], which results in neglecting the surroundings [103], leading toward video game addiction. H4 showed the relation between enthusiasm and video gaming addiction. As reported previously by Seay and Kraut [139], we found that excessive consumption of entertaining electronic gadgets can create interference in the real lives of the individuals.

Social connection and interaction were hypothesized to positively affect video game addiction (H4 and H6, respectively). However, both the constructs had no significant relationship with video game addiction. These results seem to contradict the prior literature [109,110,140-142]. Based on the context of Pakistan, this study is not in line with these 2 concepts. Pakistan, being a collectivist society, encourages people to interact socially, and excessive video gaming tends to make them socially isolated. The social connections developed during video gaming sessions seem to have little or no effect on video gaming addiction. The sense of isolation due to excessive use of technology has also been confirmed in different studies concerning internet usage [143,144]. This contradictory finding in our study opens a new door for future research as to why video gaming social connectedness does not predict video gaming addiction in Pakistan and whether the findings of this study can be applied to other contexts. We also concluded that the interaction does not predict video game addiction. Sedig and Parsons [114] report that video games are purposefully designed for a highly interactive experience. Although interactive video games lead to a more profound interest and increased engagement, they do not necessarily result

XSL•FO
RenderX

in gaming addiction. Additional research is required to investigate this finding further.

Theoretical Implications

The results of this study contribute to the understanding of the Cultivation Theory in the context of video games, that is, video game addiction [145]. While the focus of Cultivation Theory has mainly been restricted to the effects of video-based media, for example, television and films, this research applies it in the context of video games, especially how video game engagement factors cultivate addiction in video game players. This expands the scope of the theory to media that is not merely passive but also interactive in video games. This may help future researchers understand how behaviors, especially negative ones, are influenced by video game material, that is, video game engagement dimensions. Similarly, this study also helps expand the scope of the UGT. It does this by exploring the different video game engagement factors comprising absorption, conscious-attention, dedication, enjoyment, and social interaction in video games as user's gratifications are derived from video game play. This study's conceptual model highlights factors that motivate users of pleasure-oriented information systems, specifically video games, to continue prolonged usage. Specifically, this study offers dedication, absorption, conscious attention, and enthusiasm as significant, positively impacting factors that lead to addiction. This adds another dimension to the understanding of video game addiction by combining cultivation theory, UGT, and these unique factors that lead to habit-building that turns into an addiction. Moreover, this study offers insight into the reasons underlying video game usage, mostly hedonic, and how it is reinforced due to a specific gameplay characteristic. In doing so, it helps highlight the reasons underlying engagement with pleasure-oriented information systems to construct a holistic conceptual model where usage is observed as a form of meeting certain needs (UGT) and shaping worldviews (cultivation theory). Previously, authors conceptualized and specified the gaming addiction construct as a reflective-formative construct, but its empirical validation remains nebulous to date [32]. Addressing this knowledge gap, we employ the hierarchical component model approach as recommended by Sarstedt et al [132] to operationalize, specify, and validate the construct of video game addiction as a reflective-formative construct.

Practical Implications

This study offers valuable insight for practitioners in game development, psychology, and social policy as well. First, an understanding of why consumers continue to use games up to the point of addiction can highlight factors that help make games engaging. This can be used to cultivate brand loyalty as game developers roll out updates of their games. Moreover, the insignificant correlation between social connections and video game addiction will help developers understand how social bonds may not facilitate engagement for specific regions such as Pakistan, in this case. At the same time, this information can also be used to curb factors that turn repeated or prolonged usage into a harmful addiction. This can help game developers ensure that they are producing games conducive to healthy usage to meet health care and social concerns. Similarly, these findings can be used by mental health practitioners who aim to find treatment solutions for video game addictions. By understanding the aspects of video game play that leads to addiction, these practitioners may be better able to prescribe changes to playing behavior and patterns to reduce addiction. In terms of policymaking, this study helps policymakers understand what aspects of video game playing lead to the development of problematic real-life behaviors that are considered unhealthy or are discouraged.

Conclusion

This study seeks to examine the individual effect of different dimensions of video gaming engagement on video game addiction. The findings of our study add value to the existing literature. The factors that lead to video gaming addiction can be empirically stated through the findings of this study. This research also provides valuable insights into the consumption pattern of video gamers. Future research directions could involve incorporating the psychographic mechanisms as moderators such as personality traits, values and attitudes, interests, lifestyles, and gender in understanding video game addiction [146]. Additionally, the theoretical model could be applied to investigate the antecedents of video gaming engagement, that is, the factors that cause people to engage in video gaming consumption and that video game interaction does not have a significant relationship with video gaming addiction. However, the literature suggests that video games are designed to have an interactive interface to engage consumers. In this regard, a mediating mechanism of absorption, immersion, and pleasure could be empirically tested to show how the interactive interfaces convert the video gaming engagement into video game addiction.

Authors' Contributions

UR and AZA worked on the idea development and conceptual design. ZA and MAR worked on the literature review, which was further edited by UR, AZA, and HH. AZA worked on the data analyses. MM and MUS improved the contents of the manuscript. All authors contributed to this paper and approved the submitted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Questionnaire used in this study.
[DOCX File , 25 KB - games_v9i4e30310_app1.docx]

Multimedia Appendix 2
Data analysis file.
[XLSX File (Microsoft Excel File), 43 KB - games_v9i4e30310_app2.xlsx]

References

1. Abbasi AZ, Shah MU, Rehman U, Hlavacs H, Ting DH, Nisar S. The Role of Personality Factors Influencing Consumer Video Game Engagement in Young Adults: A Study on Generic Games. IEEE Access 2021;9:17392-17410. [doi: 10.1109/access.2021.3053207]
2. Goethe O. Visual aesthetics in games and gamification. In: Gamification Mindset. Cham: Springer; 2019:85-92.
3. Santasärkkä S. The digital games industry and its direct and indirect impact on the economy. case study: supercell and Finland. Theseus. 2017. URL: https://www.theseus.fi/handle/10024/128453 [accessed 2021-11-15]
4. Zhu S, Zhuang Y, Lee P, Li JC, Wong PWC. Leisure and Problem Gaming Behaviors Among Children and Adolescents During School Closures Caused by COVID-19 in Hong Kong: Quantitative Cross-sectional Survey Study. JMIR Serious Games 2021 May 07;9(2):e26808 [FREE Full text] [doi: 10.2196/26808] [Medline: 33960954]
5. Mamun MA, Griffiths MD. The Psychosocial Impact of Extreme Gaming on Indian PUBG Gamers: the Case of PUBG (PlayerUnknown's Battlegrounds). Int J Ment Health Addiction 2019 May 22:1-5. [doi: 10.1007/s11469-019-00102-4]
6. Asghar M, Noor M. Relationship Between Temperament and Video Game Addiction among Youth. pjpbs 2017 Jan 05;2(2):285-298. [doi: 10.32879/pjpbs.2016.2.2.285-298]
7. Wang J, Rost DH, Qiao R, Monk R. Academic stress and smartphone dependence among Chinese adolescents: A moderated mediation model. Children and Youth Services Review 2020 Nov;118:105029. [doi: 10.1016/j.childyouth.2020.105029]
8. Anand V. A study of time management: the correlation between video game usage and academic performance markers. Cyberpsychol Behav 2007 Aug;10(4):552-559. [doi: 10.1089/cpb.2007.9991] [Medline: 17711364]
9. Bai C, Chen X, Han K. Mobile phone addiction and school performance among Chinese adolescents from low-income families: A moderated mediation model. Children and Youth Services Review 2020 Nov;118:105406. [doi: 10.1016/j.childyouth.2020.105406]
10. Wang L, Zhu S. Online game addiction among university students. University of Gavle. 2013. URL: https://www.diva-portal.org/smash/get/diva2:602320/FULLTEXT01.pdf [accessed 2021-11-15]
11. Liu Q, Yang X, Hu Y, Zhang C, Nie Y. How and when is family dysfunction associated with adolescent mobile phone addiction? Testing a moderated mediation model. Children and Youth Services Review 2020 Apr;111:104827. [doi: 10.1016/j.childyouth.2020.104827]
12. Lam LT. Internet gaming addiction, problematic use of the internet, and sleep problems: a systematic review. Curr Psychiatry Rep 2014 Apr;16(4):444. [doi: 10.1007/s11920-014-0444-1] [Medline: 24619594]
13. Chandrima RM, Kircaburun K, Kabir H, Riaz BK, Kuss DJ, Griffiths MD, et al. Adolescent problematic internet use and parental mediation: A Bangladeshi structured interview study. Addict Behav Rep 2020 Dec;12:100288 [FREE Full text] [doi: 10.1016/j.abrep.2020.100288] [Medline: 32637564]
14. Wong RSM, Ho FKW, Tung KTS, Fu K, Ip P. Effect of Pokémon Go on Self-Harm Using Population-Based Interrupted Time-Series Analysis: Quasi-Experimental Study. JMIR Serious Games 2020 Jun 12;8(2):e17112 [FREE Full text] [doi: 10.2196/17112] [Medline: 32530429]
15. Brunborg GS, Mentzoni RA, Melkevik OR, Torsheim T, Samdal O, Hetland J, et al. Gaming Addiction, Gaming Engagement, and Psychological Health Complaints Among Norwegian Adolescents. Media Psychology 2013 Jan;16(1):115-128. [doi: 10.1080/15213269.2012.756374]
16. 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]
17. Hull DC, Williams GA, Griffiths MD. Video game characteristics, happiness and flow as predictors of addiction among video game players: A pilot study. Journal of Behavioral Addictions 2013 Sep;2(3):145-152. [doi: 10.1556/jba.2.2013.005]
18. Groves SJ, Skues JL, Wise LZ. Examining the Role of Personality Factors in Problematic Video Game Play Associated with Facebook Games. SN 2015;04(03):80-95. [doi: 10.4236/sn.2015.43010]
19. Blinka L, Škařupová K, Mitterova K. Dysfunctional impulsivity in online gaming addiction and engagement. Cyberpsychology 2016 Oct 01;10(3):1-9. [doi: 10.5817/cp2016-3-5]
20. Charlton JP, Danforth ID. Validating the distinction between computer addiction and engagement: online game playing and personality. Behaviour & Information Technology 2010 Nov;29(6):601-613. [doi: 10.1080/01449290903401978]
21. Loton D, Borkoles E, Lubman D, Polman R. Video Game Addiction, Engagement and Symptoms of Stress, Depression and Anxiety: The Mediating Role of Coping. Int J Ment Health Addiction 2015 Jul 31;14(4):565-578. [doi: 10.1007/s11469-015-9578-6]

XSL•FO
RenderX

22. Abbasi A, Ting D, Hlavacs H. Engagement in Games: Developing an Instrument to Measure Consumer Videogame Engagement and Its Validation. International Journal of Computer Games Technology 2017;2017:1-10. [doi: 10.1155/2017/7363925]
23. Gabriel Chong YM, Scott Teng KZ, Amy Siew SC, Skoric MM. Cultivation Effects of Video Games: A Longer-Term Experimental Test of First- and Second-Order Effects. Journal of Social and Clinical Psychology 2012 Nov;31(9):952-971. [doi: 10.1521/jscp.2012.31.9.952]
24. Skoric MM, Teo LLC, Neo RL. Children and video games: addiction, engagement, and scholastic achievement. Cyberpsychol Behav 2009 Oct;12(5):567-572. [doi: 10.1089/cpb.2009.0079] [Medline: 19624263]
25. Krossbakken E, Pallesen S, Mentzoni RA, King DL, Molde H, Finserås TR, et al. A Cross-Lagged Study of Developmental Trajectories of Video Game Engagement, Addiction, and Mental Health. Front Psychol 2018;9:2239 [FREE Full text] [doi: 10.3389/fpsyg.2018.02239] [Medline: 30519203]
26. Brunborg GS, Mentzoni RA, Frøyland LR. Is video gaming, or video game addiction, associated with depression, academic achievement, heavy episodic drinking, or conduct problems? Journal of Behavioral Addictions 2014 Mar;3(1):27-32. [doi: 10.1556/jba.3.2014.002]
27. 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. Computers in Human Behavior 2012 Nov;28(6):2143-2151. [doi: 10.1016/j.chb.2012.06.019]
28. Abbasi A, Ting D, Hlavacs H, Costa L, Veloso A. An empirical validation of consumer video game engagement: A playful-consumption experience approach. Entertainment Computing 2019 Mar;29:43-55 [FREE Full text] [doi: 10.1016/j.entcom.2018.12.002]
29. Abbasi AZ, Nisar S, Rehman U, Ting DH. Impact of HEXACO Personality Factors on Consumer Video Game Engagement: A Study on eSports. Front Psychol 2020;11:1831 [FREE Full text] [doi: 10.3389/fpsyg.2020.01831] [Medline: 32849078]
30. Kaur P, Dhir A, Chen S, Malibari A, Almotairi M. Telematics and Informatics 2020 Oct;53:101376. [doi: 10.1016/j.tele.2020.101376]
31. Pham TTL, Huang H, Cheng T, Wong M, Liao Y, Yang Y, et al. The need for exercise in exergaming perspective of the uses and gratifications theory. IMDS 2020 May 23;120(6):1085-1099. [doi: 10.1108/imds-07-2019-0377]
32. van Rooij AJ, Van Looy J, Billieux J. Internet Gaming Disorder as a formative construct: Implications for conceptualization and measurement. Psychiatry Clin Neurosci 2017 Jul;71(7):445-458 [FREE Full text] [doi: 10.1111/pcn.12404] [Medline: 27193187]
33. Katz E, Blumler J, Gurevitch M. The Uses of Mass Communications: Current Perspectives on Gratifications Research. Oaks, CA: Sage Publications; 1974:0803903405.
34. Gerbner G. Cultivation Analysis: An Overview. Mass Communication and Society 1998 Jun;1(3-4):175-194. [doi: 10.1080/15205436.1998.9677855]
35. Wu J, Wang S, Tsai H. Falling in love with online games: The uses and gratifications perspective. Computers in Human Behavior 2010 Nov;26(6):1862-1871. [doi: 10.1016/j.chb.2010.07.033]
36. Rubin A. Uses and gratifications. In: The SAGE Handbook of Media Processes and Effects. Oaks, CA: Sage Publication; 2009:147-159.
37. Ruggiero TE. Uses and Gratifications Theory in the 21st Century. Mass Communication and Society 2000 Feb;3(1):3-37. [doi: 10.1207/s15327825mcs0301_02]
38. Funk J, Buchman D. Playing violent video and computer games and adolescent self-concept. Journal of communication 1996;46(2):19-32. [doi: 10.1111/j.1460-2466.1996.tb01472.x]
39. Sherry J, Lucas K, Greenberg B, Lachlan K. Video game uses and gratifications as predictors of use and game preference. Playing video games: Motives, responses, and consequences 2006;24(1):213-224. [doi: 10.4324/9780203873700]
40. Sampat B, Krishnamoorthy B. Journal of International Technology and Information Management. 2016. URL: https://scholarworks.lib.csusb.edu/jitim/vol25/iss3/5/ [accessed 2021-11-15]
41. Bouvier P, Lavoué E, Sehaba K. Defining Engagement and Characterizing Engaged-Behaviors in Digital Gaming. Simulation & Gaming 2014 Nov 06;45(4-5):491-507. [doi: 10.1177/1046878114553571]
42. Bosser A, Nakatsu R. Hardcore gamers and casual gamers playing online together. In: Entertainment Computing. Berlin, Heidelberg: Springer; 2006:374-377.
43. Poels Y, Annema JH, Verstraete M, Zaman B, De Grooff D. Are you a gamer? a qualitative study on the parameters for categorizing casual and hardcore gamers. IJWI 2012:1-16. [doi: 10.33965/ijwi]
44. Hsu C, Lu H. Consumer behavior in online game communities: A motivational factor perspective. Computers in Human Behavior 2007 May;23(3):1642-1659. [doi: 10.1016/j.chb.2005.09.001]
45. Van Rooij AJ, Schoenmakers T, Vermulst A, Van den Eijnden RJJM, Van de Mheen D. Online video game addiction: identification of addicted adolescent gamers. Addiction 2011 Jan;106(1):205-212. [doi: 10.1111/j.1360-0443.2010.03104.x] [Medline: 20840209]
46. Palmgreen P, Wenner L, Rosengren K. Uses and gratifications research: The past ten years. Media gratifications research: Current perspectives 1985;11:37.

XSL•FO
RenderX

47. Gerbner G. An institutional approach to mass communications research. Communication Theory and Research. 1967. URL: https://web.asc.upenn.edu/gerbner/Asset.aspx?assetID=302 [accessed 2021-11-15]
48. Borah P. Media Effects Theory. The international encyclopedia of political communication 2016:1-12. [doi: 10.1002/9781118541555.wbiepc156]
49. Riddle K. Cultivation theory revisited: the impact of childhood television viewing levels on social reality beliefs and construct accessibility in adulthood (conference papers). International Communication Association. 2009. URL: https://tinyurl.com/6ra66fns [accessed 2021-11-15]
50. Cialdini R, Trost M. Social influence: social norms, conformity and compliance. In: The Handbook of Social Psychology. Boston: McGraw-Hill; 1998:151-192.
51. Fazio R, Roskos-Ewoldsen D, Powell M. Attitudes, perception, and attention. In: The Heart's Eye: Emotional Influences in Perception and Attention. San Diego: Academic Press; 1994:197-216.
52. Gerbner G, Gross L. Living with television: The violence profile. Journal of Communication 1976;26(2):172-199. [doi: 10.1111/j.1460-2466.1976.tb01397.x]
53. Valkenburg PM, Peter J, Walther JB. Media Effects: Theory and Research. Annu Rev Psychol 2016;67:315-338. [doi: 10.1146/annurev-psych-122414-033608] [Medline: 26331344]
54. Van Mierlo J, Van den Bulck J. Benchmarking the cultivation approach to video game effects: a comparison of the correlates of TV viewing and game play. J Adolesc 2004 Feb;27(1):97-111. [doi: 10.1016/j.adolescence.2003.10.008] [Medline: 15013263]
55. Jin SA. “It Feels Right. Therefore, I Feel Present and Enjoy”: The Effects of Regulatory Fit and the Mediating Roles of Social Presence and Self-Presence in Avatar-Based 3D Virtual Environments. Presence: Teleoperators and Virtual Environments 2011 Apr 01;20(2):105-116. [doi: 10.1162/pres_a_00038]
56. Shang C, Williams T, Beavis A, Ward J, Sims C, Phillips R. SU-FF-T-117: Can Current Prostate IMRT Be Further Improved with Immersive Virtual Reality Simulation? Med. Phys 2006 Jul 11;33(6Part8):2075-2076. [doi: 10.1118/1.2241042]
57. Seah M, Cairns P. From immersion to addiction in videogames. 2008 Presented at: 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction; September 1-5, 2008; Liverpool, UK p. 55-63. [doi: 10.14236/ewic/hci2008.6]
58. Quwaider M, Alabed A, Duwairi R. The Impact of Video Games on the Players Behaviors: A Survey. Procedia Computer Science 2019;151:575-582. [doi: 10.1016/j.procs.2019.04.077]
59. Prot S, Gentile DA, Anderson CA, Suzuki K, Swing E, Lim KM, et al. Long-term relations among prosocial-media use, empathy, and prosocial behavior. Psychol Sci 2014 Feb;25(2):358-368. [doi: 10.1177/0956797613503854] [Medline: 24335350]
60. Fuster H, Chamarro A, Carbonell X, Vallerand RJ. Relationship between passion and motivation for gaming in players of massively multiplayer online role-playing games. Cyberpsychol Behav Soc Netw 2014 May;17(5):292-297. [doi: 10.1089/cyber.2013.0349] [Medline: 24611801]
61. Wittek CT, Finserås TR, Pallesen S, Mentzoni RA, Hanss D, Griffiths MD, et al. Prevalence and Predictors of Video Game Addiction: A Study Based on a National Representative Sample of Gamers. Int J Ment Health Addict 2016;14(5):672-686 [FREE Full text] [doi: 10.1007/s11469-015-9592-8] [Medline: 27688739]
62. Porter JR. Understanding and addressing real-world accessibility issues in mainstream video games. SIGACCESS Access. Comput 2014 Jan(108):42-45. [doi: 10.1145/2591357.2591364]
63. 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. Journal of Behavioral Addictions 2014 Sep;3(3):157-165. [doi: 10.1556/jba.3.2014.013]
64. Asghar M, Noor M. Relationship Between Temperament and Video Game Addiction among Youth. pjpbs 2017 Jan 05;2(2):285-298. [doi: 10.32879/pjpbs.2016.2.2.285-298]
65. Banyte J, Gadeikiene A. The Effect of Consumer Motivation to Play Games on Video Game-playing Engagement. Procedia Economics and Finance 2015;26:505-514 [FREE Full text] [doi: 10.1016/s2212-5671(15)00880-1]
66. Brown E, Cairns P. A grounded investigation of game immersion. 2004 Presented at: CHI EA '04; April; Vienna, Austria. [doi: 10.1145/985921.986048]
67. Rigby S, Ryan R. Glued to Games: How Video Games Draw Us In and Hold Us Spellbound. USA: Praeger; Illustrated edition; Feb 18, 2011.
68. Przybylski AK, Rigby CS, Ryan RM. A Motivational Model of Video Game Engagement. Review of General Psychology 2010 Jun 01;14(2):154-166. [doi: 10.1037/a0019440]
69. Renninger K, Hidi S, Krapp A, Renninger A. The Role of Interest in Learning and Development. New York: Taylor & Francis; 2014.
70. Adams E, Rollings A. Fundamentals of Game Design. Berkeley, CA: Pearson Prentice Hall; 2007.
71. 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 Sep;66(9):641-661. [doi: 10.1016/j.ijhcs.2008.04.004]
72. Lombard M, Ditton T. At the heart of it all: The concept of presence. Journal of computer-mediated communication 1997;3(2). [doi: 10.1111/j.1083-6101.1997.tb00072.x]

XSL•FO
RenderX

73. Chen J. Flow in games (and everything else). Commun. ACM 2007 Apr;50(4):31-34. [doi: 10.1145/1232743.1232769]
74. Koster R. Theory of Fun for Game Designers. Scottsdale, Arizona: Paraglyph Press; 2005.
75. IJsselsteijn W, Van DHW, Klimmt C, De KY, Lindley C, Mathiak K. Measuring the experience of digital game enjoyment. Psychology of Digital Games. 2008. URL: https://www.researchgate.net/publication/254415787_Measuring_the_experience_of_digital_game_enjoyment [accessed 2021-11-15]
76. Calleja G. In-game: From Immersion to Incorporation. Harvard: MIT Press; 2011.
77. Hollebeek LD, Glynn MS, Brodie RJ. Consumer Brand Engagement in Social Media: Conceptualization, Scale Development and Validation. Journal of Interactive Marketing 2014 May;28(2):149-165. [doi: 10.1016/j.intmar.2013.12.002]
78. Kuvykaitė R, Tarutė A. A Critical Analysis of Consumer Engagement Dimensionality. Procedia - Social and Behavioral Sciences 2015 Dec;213:654-658. [doi: 10.1016/j.sbspro.2015.11.468]
79. Brodie RJ, Ilic A, Juric B, Hollebeek L. Consumer engagement in a virtual brand community: An exploratory analysis. Journal of Business Research 2013 Jan;66(1):105-114. [doi: 10.1016/j.jbusres.2011.07.029]
80. Charlton JP, Danforth ID. Distinguishing addiction and high engagement in the context of online game playing. Computers in Human Behavior 2007 May;23(3):1531-1548. [doi: 10.1016/j.chb.2005.07.002]
81. Griffiths MD. Videogame Addiction: Further Thoughts and Observations. Int J Ment Health Addiction 2007 Sep 18;6(2):182-185. [doi: 10.1007/s11469-007-9128-y]
82. Seok S, DaCosta B. Distinguishing Addiction From High Engagement. Games and Culture 2014 Jul 13;9(4):227-254. [doi: 10.1177/1555412014538811]
83. Mentzoni RA, Brunborg GS, Molde H, Myrseth H, Skouverøe KJM, Hetland J, et al. Problematic video game use: estimated prevalence and associations with mental and physical health. Cyberpsychol Behav Soc Netw 2011 Oct;14(10):591-596. [doi: 10.1089/cyber.2010.0260] [Medline: 21342010]
84. Eklund L. Playing video games together with others: Differences in gaming with family, friends and strangers. journal of gaming & virtual worlds 2015 Sep 01;7(3):259-277. [doi: 10.1386/jgvw.7.3.259_1]
85. Kallio KP, Mäyrä F, Kaipainen K. At Least Nine Ways to Play: Approaching Gamer Mentalities. Games and Culture 2010 Dec 31;6(4):327-353. [doi: 10.1177/1555412010391089]
86. Van den Bulck H, Van Aelst P. Het Etmaal van de Communicatiewetenschap: Een lokale anomalie in een globaliserend academisch veld? Tijdschrift voor Communicatiewetenschappen 2007 Dec;35(4):269-274. [doi: 10.1347/tijd.35.4.269]
87. Schuurman D, De MK, De ML, Van LJ, editors. Fanboys, competers, escapists and time-killers: a typology based on gamers' motivations for playing video games. 2008 Presented at: 3rd International Conference on Digital Interactive Media in Entertainment and Arts; September 10-12, 2008; Athens, Greece. [doi: 10.1145/1413634.1413647]
88. Juul J. A Casual Revolution: Reinventing Video Games and Their Players. Cambridge, MA: MIT press; 2010.
89. Kuittinen J, Kultima A, Niemel J, Paavilainen J. Casual games discussion. 2007 Presented at: Proceedings of the 2007 conference on Future Play; November; Toronto, Canada p. 166-173. [doi: 10.1145/1328202.1328221]
90. Conrad D. The compatibility of artworks and games. Trinity University. 2016. URL: https://digitalcommons.trinity.edu/cgi/viewcontent.cgi?article=1031&context=infolit_usra [accessed 2020-11-12]
91. Griffiths M. A ‘components’ model of addiction within a biopsychosocial framework. Journal of Substance Use 2009 Jul 12;10(4):191-197. [doi: 10.1080/14659890500114359]
92. van RA, Schoenmakers T, Vermulst A, van DER, van DMD. Online video game addiction: identification of addicted adolescent gamers. Addiction 2010;8:11. [doi: 10.1111/j.1360-0443.2010.03104.x]
93. Vivek SD, Beatty SE, Morgan RM. Customer Engagement: Exploring Customer Relationships Beyond Purchase. Journal of Marketing Theory and Practice 2014 Dec 08;20(2):122-146. [doi: 10.2753/mtp1069-6679200201]
94. Csikszentmihalyi M. Flow. In: The Psychology of Optimal Experience. New York: Cambridge University Press; 1990.
95. Bowman N, Kowert R, Ferguson C. The impact of video game play on human (and orc) creativity. In: Video Games and Creativity. Amsterdam: Academic Press, Elsvier; 2015.
96. De Pasquale C, Dinaro C, Sciacca F. Relationship of Internet gaming disorder with dissociative experience in Italian university students. Ann Gen Psychiatry 2018;17:28 [FREE Full text] [doi: 10.1186/s12991-018-0198-y] [Medline: 29983724]
97. Chou T, Ting C. The role of flow experience in cyber-game addiction. Cyberpsychol Behav 2003 Dec;6(6):663-675. [doi: 10.1089/109493103322725469] [Medline: 14756934]
98. Hollebeek LD. Demystifying customer brand engagement: Exploring the loyalty nexus. Journal of Marketing Management 2011 Jul;27(7-8):785-807. [doi: 10.1080/0267257x.2010.500132]
99. Mollen A, Wilson H. Engagement, telepresence and interactivity in online consumer experience: Reconciling scholastic and managerial perspectives. Journal of Business Research 2010 Sep;63(9-10):919-925. [doi: 10.1016/j.jbusres.2009.05.014]
100. Hollebeek L. Exploring customer brand engagement: definition and themes. Journal of Strategic Marketing 2011 Dec;19(7):555-573. [doi: 10.1080/0965254x.2011.599493]
101. Nakamura J, Csikszentmihalyi M. The concept of flow. In: Flow and the Foundations of Positive Psychology. Cham: Springer; 2014:239-263.
102. DeCharms R. Personal Causation: The Internal Affective Determinants of Behavior. New York: Academic Press; 1968.

XSL•FO
RenderX

103. Haladjian H, Montemayor C. Unifying theories of consciousness, attention, and conscious attention. 2013 Presented at: Annual Meeting of the Cognitive Science Society; July 31-August 3, 2013; Berlin, Germany. [doi: 10.7551/mitpress/9780262028974.003.0004]
104. Vivek SD, Beatty SE, Dalela V, Morgan RM. A Generalized Multidimensional Scale for Measuring Customer Engagement. Journal of Marketing Theory and Practice 2014 Dec 07;22(4):401-420. [doi: 10.2753/mtp1069-6679220404]
105. Jansz J, Martens L. Gaming at a LAN event: the social context of playing video games. New Media & Society 2016 Jun 30;7(3):333-355. [doi: 10.1177/1461444805052280]
106. Krotoski A. Chicks and joysticks: an exploration of women and gaming. Psychology. 2004. URL: https://tinyurl.com/3j37bpjh [accessed 2021-11-15]
107. Weibel D, Wissmath B, Habegger S, Steiner Y, Groner R. Playing online games against computer- vs. human-controlled opponents: Effects on presence, flow, and enjoyment. Computers in Human Behavior 2008 Sep;24(5):2274-2291. [doi: 10.1016/j.chb.2007.11.002]
108. Valkenburg P, Peter J. Online communication and adolescent well-being: Testing the stimulation versus the displacement hypothesis. Journal of Computer-Mediated Communication 2007;12(4):1169-1182. [doi: 10.1111/j.1083-6101.2007.00368.x]
109. Sioni SR, Burleson MH, Bekerian DA. Internet gaming disorder: Social phobia and identifying with your virtual self. Computers in Human Behavior 2017 Jun;71:11-15. [doi: 10.1016/j.chb.2017.01.044]
110. Cole H, Griffiths MD. Social interactions in massively multiplayer online role-playing gamers. Cyberpsychol Behav 2007 Aug;10(4):575-583. [doi: 10.1089/cpb.2007.9988] [Medline: 17711367]
111. Peters CS, Malesky LA. Problematic usage among highly-engaged players of massively multiplayer online role playing games. Cyberpsychol Behav 2008 Aug;11(4):481-484. [doi: 10.1089/cpb.2007.0140] [Medline: 18721098]
112. So KKF, King C, Sparks B. Customer Engagement With Tourism Brands. Journal of Hospitality & Tourism Research 2012 Jul 03;38(3):304-329. [doi: 10.1177/1096348012451456]
113. Morrison M, Krugman DM. A Look At Mass and Computer Mediated Technologies: Understanding the Roles of Television and Computers in the Home. Journal of Broadcasting & Electronic Media 2001 Mar;45(1):135-161. [doi: 10.1207/s15506878jobem4501_9]
114. Sedig K, Parsons P. Interaction Design for Complex Cognitive Activities with Visual Representations: A Pattern-Based Approach. THCI 2013 Jun 30;5(2):84-133. [doi: 10.17705/1thci.00055]
115. Granic I, Lobel A, Engels RCME. The benefits of playing video games. Am Psychol 2014 Jan;69(1):66-78. [doi: 10.1037/a0034857] [Medline: 24295515]
116. Krueger AB. How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989. The Quarterly Journal of Economics 1993 Feb 01;108(1):33-60. [doi: 10.2307/2118494]
117. Wan RW, Hamid X, Sabri M. Video games: issues and problems. : IEEE xplore; 2007 Presented at: 5th International Conference on Information and Communications Technology; Dec 16-18, 2007; Cairo, Egypt. [doi: 10.1109/itict.2007.4475651]
118. Griffiths M, Hussain Z, Grüsser S, Thalemann R, Cole H, Davies M. Social interactions in online gaming. International Journal of Game-Based Learning 2011;1(4):20-36. [doi: 10.4018/ijgbl.2011100103]
119. Sedgwick P. Cross sectional studies: advantages and disadvantages. BMJ 2014 Mar 26;348:g2276-g2276. [doi: 10.1136/bmj.g2276]
120. Abbasi A, Ting D, Hlavacs H. A New Scale Development. : Springer, Cham; 2016 Presented at: International Conference on Entertainment Computing; September 2016; Austria p. 247-252. [doi: 10.1007/978-3-319-46100-7_25]
121. Lemmens JS, Valkenburg PM, Peter J. Development and Validation of a Game Addiction Scale for Adolescents. Media Psychology 2009 Mar 05;12(1):77-95. [doi: 10.1080/15213260802669458]
122. Hair JJ, Hult G, Ringle C, Sarstedt M. An introduction to structural equation modeling. In: Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R. Cham: Springer; 2021.
123. Emmel N. Purposeful sampling. Sampling and choosing cases in qualitative research: A realist approach 2013:33-45. [doi: 10.4135/9781473913882.n3]
124. Abbasi AZ, Asif M, Hollebeek LD, Islam JU, Ting DH, Rehman U. The effects of consumer esports videogame engagement on consumption behaviors. Journal of Product & Brand Management. 2020 Dec 07. URL: https://www.researchgate.net/publication/345061354_The_effect_of_consumer_esports_videogame_engagement_on_consumption_behaviors [accessed 2021-11-15]
125. Abbasi A, Shamim A, Hooi TD, Hlavacs H, Rehman U. Playful-consumption experiences and subjective well-being: Children?s smartphone usage. Entertainment Computing 2021:100390 [FREE Full text] [doi: 10.1016/j.entcom.2020.100390]
126. Hair J, Hollingsworth CL, Randolph AB, Chong AYL. An updated and expanded assessment of PLS-SEM in information systems research. IMDS 2017 Apr 10;117(3):442-458. [doi: 10.1108/imds-04-2016-0130]
127. Chang C, Yeh S, Li M. The Adoption of a Virtual Reality-Assisted Training System for Mental Rotation: A Partial Least Squares Structural Equation Modeling Approach. JMIR Serious Games 2020 Jan 17;8(1):e14548 [FREE Full text] [doi: 10.2196/14548] [Medline: 31804184]
128. Hair JF, Risher JJ, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. EBR 2019 Jan 14;31(1):2-24. [doi: 10.1108/ebr-11-2018-0203]

XSL•FO
RenderX

129. Kock N. WarpPLS user manual: Version 6. ScriptWarp Systems. 2017. URL: http://cits.tamiu.edu/WarpPLS/UserManual_v_6_0.pdf [accessed 2021-11-15]
130. Henseler J, Ringle CM, Sarstedt M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. J. of the Acad. Mark. Sci 2014 Aug 22;43(1):115-135. [doi: 10.1007/s11747-014-0403-8]
131. Becker J, Klein K, Wetzels M. Hierarchical Latent Variable Models in PLS-SEM: Guidelines for Using Reflective-Formative Type Models. Long Range Planning 2012 Oct;45(5-6):359-394. [doi: 10.1016/j.lrp.2012.10.001]
132. Sarstedt M, Hair JF, Cheah J, Becker J, Ringle CM. How to Specify, Estimate, and Validate Higher-Order Constructs in PLS-SEM. Australasian Marketing Journal 2021 Jan 27;27(3):197-211. [doi: 10.1016/j.ausmj.2019.05.003]
133. Bhagat S, Jeong EJ, Kim DJ. The Role of Individuals' Need for Online Social Interactions and Interpersonal Incompetence in Digital Game Addiction. International Journal of Human–Computer Interaction 2019 Aug 29;36(5):449-463. [doi: 10.1080/10447318.2019.1654696]
134. von der Heiden JM, Braun B, Müller KW, Egloff B. The Association Between Video Gaming and Psychological Functioning. Front Psychol 2019;10:1731 [FREE Full text] [doi: 10.3389/fpsyg.2019.01731] [Medline: 31402891]
135. Wood RTA. Problems with the Concept of Video Game "Addiction": Some Case Study Examples. Int J Ment Health Addiction 2007 Oct 23;6(2):169-178. [doi: 10.1007/s11469-007-9118-0]
136. Funk D, Mahony D, Ridinger L. Characterizing consumer motivation as individual difference factors: Augmenting the sports interest inventory (SII) to explain level of spectator support. Sport Marketing Quarterly 2002;11(1):A.
137. Corcoran R. Combined group proposal for individuals experiencing online video game addiction. California State University. 2016. URL: https://scholarworks.csun.edu/bitstream/handle/10211.3/183082/Corcoran-Rory-thesis-2017.pdf?sequence=1 [accessed 2020-10-11]
138. Barnes SJ, Pressey AD, Scornavacca E. Mobile ubiquity: Understanding the relationship between cognitive absorption, smartphone addiction and social network services. Computers in Human Behavior 2019 Jan;90:246-258. [doi: 10.1016/j.chb.2018.09.013]
139. Seay A, Kraut R. Project massive: Self-regulation and problematic use of online gaming. 2007 Presented at: SIGCHI conference on human factors in computing systems; April 28-May 3, 2007; San Jose, California, USA. [doi: 10.1145/1240624.1240749]
140. Van RA. Online video game addiction: exploring a new phenomenon. Erasmus MC: University Medical Center Rotterdam. 2011. URL: https://repub.eur.nl/pub/23381/ [accessed 2021-01-02]
141. Schmitt ZL, Livingston MG. Video game addiction and college performance among males: results from a 1 year longitudinal study. Cyberpsychol Behav Soc Netw 2015 Jan;18(1):25-29. [doi: 10.1089/cyber.2014.0403] [Medline: 25584728]
142. Bhagat S, Jeong EJ, Kim DJ. The Role of Individuals' Need for Online Social Interactions and Interpersonal Incompetence in Digital Game Addiction. International Journal of Human–Computer Interaction 2019 Aug 29;36(5):449-463. [doi: 10.1080/10447318.2019.1654696]
143. Yayan EH, Suna Dağ Y, Düken ME. The effects of technology use on working young loneliness and social relationships. Perspect Psychiatr Care 2019 Apr;55(2):194-200. [doi: 10.1111/ppc.12318] [Medline: 30044497]
144. 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]
145. Stein J, Krause E, Ohler P. Every (Insta)Gram counts? Applying cultivation theory to explore the effects of Instagram on young users' body image. Psychology of Popular Media 2021 Jan;10(1):87-97. [doi: 10.1037/ppm0000268]
146. LaGrotta C. Gender Differences in Addiction. Curr Addict Rep 2021 Jan 18;8(1):64-70. [doi: 10.1007/s40429-021-00355-x]

Abbreviations

MMORPG: massively multiplayer online role-playing game
PLS-SEM: partial least squares-structural equation modeling
UGT: uses and gratifications theory

Edited by N Zary; submitted 10.05.21; peer-reviewed by D Parry, O Blanson Henkemans; comments to author 03.08.21; revised version received 28.09.21; accepted 12.10.21; published 26.11.21.

Please cite as:
Abbasi AZ, Rehman U, Afaq Z, Rafeh MA, Hlavacs H, Mamun MA, Shah MU
Predicting Video Game Addiction Through the Dimensions of Consumer Video Game Engagement: Quantitative and Cross-sectional Study
JMIR Serious Games 2021;9(4):e30310
URL: https://games.jmir.org/2021/4/e30310
doi:10.2196/30310
PMID:34842539

XSL•FO
RenderX

©Amir Zaib Abbasi, Umair Rehman, Zahra Afaq, Mir Abdur Rafeh, Helmut Hlavacs, Mohammed A Mamun, Muhammad Umair Shah. Originally published in JMIR Serious Games (https://games.jmir.org), 26.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Relationship Between Illness Representations and Symptoms of Internet Gaming Disorder Among Young People: Cross-Lagged Model

Xue Yang[1], PhD; Kei Man Wong[1], BSc; Rui She[1], PhD; Chengjia Zhao[2], MA; Nani Ding[3], MA; Huihui Xu[2], MA; Xiaolian Tu[4], MA; Xinyi Lai[2], MA; Guohua Zhang[5,6], PhD

[1]JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)
[2]Department of Psychology, School of Mental Health, Wenzhou Medical University, Wenzhou, China
[3]School of Public Health and Management, Wenzhou Medical University, Wenzhou, China
[4]Renji College, Wenzhou Medical University, Wenzhou, China
[5]The Affiliated Kangning Hospital, Wenzhou Medical University, Wenzhou, China
[6]Institute of Aging, Key Laboratory of Alzheimer's Disease of Zhejiang Province, Wenzhou Medical University, Wenzhou, China

Corresponding Author:
Guohua Zhang, PhD
Institute of Aging
Key Laboratory of Alzheimer's Disease of Zhejiang Province
Wenzhou Medical University
University Town
Chashan
Wenzhou, 325035
China
Phone: 86 +8613957764528
Fax: 86 0577 86689925
Email: zghcnu@wmu.edu.cn

Abstract

Background: The common-sense model of illness suggests that mental representations of health threats may affect one's behavioral reactions to them and health status. Internet gaming disorder is a newly defined mental disorder. Illness representations of internet gaming disorder may affect one's risk of internet gaming disorder. In turn, symptoms of internet gaming disorder may affect one's perceptions of the disorder.

Objective: This study aimed to investigate the relationships between illness representations and symptoms of internet gaming disorder in college students.

Methods: A 1-year longitudinal study was conducted with a convenience sample of Chinese college students (n=591; 342/591, 57.9% female).

Results: Of the participants, 10.1% (60/591) and 9.1% (54/591) were classified as having probable internet gaming disorder at baseline (T1) and follow-up (T2), respectively. The correlations between some dimensions of illness representations regarding internet gaming disorder (ie, consequence, timeline, personal control, treatment control, and concern) at T1 and symptoms of internet gaming disorder at T2 and between symptoms of internet gaming disorder at T1 and the dimensions of illness representations at T2 (ie, consequence, timeline, personal control, and emotional response) were statistically significant. The cross-lagged model fit the data well ((χ^2/df=2.28, comparative fit index=.95, root mean square error of approximation=.06) and showed that internet gaming disorder at T1 was positively associated with unfavorable illness representations at T2.

Conclusions: Individuals with more severe symptoms of internet gaming disorder had more pessimistic perceptions about the disorder. Such cognitive perceptions may affect one's emotional and behavioral reactions towards the disorder (eg, greater levels of depression and low self-control intention) and should be modified by educational programs and psychological interventions.

(JMIR Serious Games 2021;9(4):e28117) doi:10.2196/28117

XSL•FO
RenderX

KEYWORDS

illness representations; internet gaming disorder; youth; cross-lagged model

Introduction

Background

Internet gaming is the most common leisure activity among young people in East Asian countries, including China [1-4]. Although some studies have proposed positive effects of internet gaming [5,6], excessive internet game use frequently leads to various physical (eg, obesity), psychological (eg, depression, anxiety), and social problems (eg, interpersonal conflict) [1-3]. Internet gaming disorder (IGD) is defined as "persistent and recurrent use of the internet to engage in games, often with other players, leading to clinically significant impairment or distress" [7,8]. It was listed as a condition for further research in the Diagnostic and Statistical Manual of Mental Disorder (DSM-5) [7] and later as an official disorder in the 11th Revision of the International Classification of Diseases (ICD-11) [9]. Examples of IGD symptoms include preoccupation, tolerance, and inability to control gaming behaviors in the past year [7,9]. Debates have continued about the conceptualization, diagnostic criteria, and guidelines for IGD [8].

Prevalence rates of IGD vary widely across studies, ranging from 0.7% to 15.6% [10]. A recent meta-analysis found a worldwide IGD prevalence of 3.05% [11]. Youth and young adults are high-risk populations for IGD. For example, a cross-cultural study among young adults (aged 18-25 years) reported IGD prevalence rates of 26% in the United States, 21.1% in China, and 15.4% in Singapore [12]. Studies among Chinese college students found prevalences ranging from 4% to 25.7% [13-17]. A longitudinal survey with Chinese university students reported prevalence rates of 14.8% and 9.9% at baseline and the 1-year follow-up, respectively [18]. IGD has been found to be associated with other mental disorders, such as depression, anxiety, and attention-deficit/hyperactivity disorder [19-22] as well as health and life problems, including poor sleep quality, academic failure, and interpersonal problems [16,23,24]. Given the high prevalence and severe consequences of IGD in young people, efforts to understand how they perceive this newly defined disorder is warranted. The understanding of these perceptions and relationships would facilitate related research and intervention development (eg, educational programs, psychological interventions) to prevent and reduce IGD.

Concepts of Illness Representations

The lay understanding of a potential health threat (ie, illness representations) may influence individuals' responses to it (eg, IGD). The common-sense model (CSM) of illness [25] suggests that when individuals encounter a health threat, they develop mental representations of it based on their general knowledge and past experience, information from the external social environment (eg, friends), and current experience of it. Illness representations involve cognitive and emotional representations. To be specific, the cognitive representations include consequences (assessing the severity of the illness), timeline (considering whether the illness is acute/chronic/cyclical), control (considering whether the illness is under volitional control), identity (labeling the illness and identifying its symptoms), illness comprehension (overall comprehension of the illness), and cause (attributing likely causes of the illness). Emotional representations (expressing the emotions evoked by the illness) typically entail negative emotional responses such as anger and worry. These representations interact with each other and work in parallel to guide individuals' coping strategies and affect their mental and behavioral responses to the illness (eg, self-management, help-seeking). For example, it was found that high treatment control (ie, perceived treatment effectiveness) was associated with better treatment adherence among patients with bipolar disorder [26] and type 2 diabetes [27]. Illness representations are dynamic. They would be appraised and revised based on the consequences of the coping responses to health threats [25] and could be modified by interventions [28,29].

Relationships Between Illness Representations and Symptoms of IGD

Illness representations of IGD may affect one's risk of IGD. People with unfavorable illness representations of IGD (eg, perceiving and worrying about severe consequences of IGD) may be less likely to have IGD. It may be because such people are more likely to regulate their internet gaming behaviors, prevent themselves from excessive internet gaming and IGD, or make an effort to reduce their IGD symptoms. Although related studies are limited, CSM can be used to support our hypothesis theoretically, which postulates that both cognitive representations and emotional representations of the illness can affect coping and appraisal of a disease, which in turn determine health-related outcomes [25,30]. In addition, a number of studies have investigated the associations between illness representations and various health-related behaviors (eg, self-management behaviors, willingness to consult health care providers) and health outcomes (eg, physical functioning, mental health status, and clinical health outcomes), among both patients and healthy populations [30-32]. For example, a previous study reported that illness representations of H1N1 were significantly associated with H1N1 preventive behaviors, including wearing face masks, hand washing, and behavioral intention to take up influenza vaccination [32].

However, IGD symptoms may increase the levels of unfavorable representations of IGD. It may be because people with greater symptoms of IGD may have experienced negative consequences of IGD or had trouble with regulating their internet gaming behaviors, thus enhancing their unfavorable representations of IGD, while those with no or minimal symptoms of IGD have not experienced negative consequences of IGD or may not necessarily think about IGD negatively [33]. A recent cross-sectional study in an adult sample (mean age 40.42 years) in Macao, China reported that those with IGD showed more negative emotional representations, scored higher in the level of the time cyclical factor, and had better comprehensiveness of IGD than non-IGD cases [33]. However, this was the only existing study, and the cross-sectional study design could not

XSL•FO
RenderX

demonstrate a causal relationship between IGD and illness representations of IGD. To our best knowledge, no study has investigated the longitudinal and interactive relationships between illness representations and symptoms of IGD. This study attempted to fill this gap.

This Study

This 2-wave longitudinal study aimed to understand how young people cognitively and emotionally perceive IGD as a disorder and health threat (ie, illness representations). In addition, the study investigated the relationships between such perceptions and symptoms of IGD in a population of Chinese college students. It was hypothesized that (1) unfavorable illness representations (eg, severe consequence) of IGD at baseline (T1) would predict low levels of IGD symptoms at follow-up (T2); however, (2) low levels of IGD symptoms at baseline (T1) would predict high levels of unfavorable illness representations of IGD at follow-up (T2).

Methods

Participants

This longitudinal study adopted a convenience sample recruited from 2 universities in China. The inclusion criteria of this study included (1) being a first-year college student, (2) being willing to participate in the baseline and follow-up studies, and (3) playing internet games in the past year. The exclusion criterion was being a non-Chinese speaker. Data from the participants who completed the 2-wave surveys were reported in this study (n=591). As shown in Table 1, about three-fifths of the participants (342/591, 57.9%) were female, 51.4% (287/591) were from urban areas, and 48.7% (288/591) lived in one-child families. The mean hours of internet gaming per week were 7.7 (SD 9.7) at T1 and 8.4 (SD 10.7) at T2.

Table 1. Background characteristics of the participants (n=591).

Background characteristics	Participants, n (%)
Sex	
Female	342 (57.9)
Male	249 (42.1)
Residence	
Urban	304 (51.4)
Rural	287 (48.6)
One-child family	
Yes	288 (48.7)
No	303 (51.3)
Hours of internet gaming per week	
0-7	264 (44.6)
>7 to 14	181 (30.7)
>14 to 21	85 (14.3)
>21	67 (11.4)
Major	
Anesthesia	84 (14.2)
Forensic Medicine	87 (14.7)
Stomatology	165 (27.9)
Chinese Medicine	137 (23.1)
Psychology	118 (20.0)

Data Collection

The baseline survey (T1) was conducted during the first year of college study of the participants, while the 1-year follow-up survey (T2) was conducted during their second year of college study. In the absence of any teacher, the surveys were conducted in classroom settings. A research assistant with 2 years of experience in data collection and who majored in psychology assisted with data collection and answered any questions during the survey. The participants were assured that participation was voluntary and refusals would have no negative consequences. Data confidentiality was guaranteed, and only the researchers could access the data. Student IDs were used for data matching. Researchers were not able to access students' names or other identifying information. All participants were briefed on the purpose of the study and provided their informed consent to participate in this anonymous survey. They were also provided with information of local psychological services in case it was needed. The study procedures were carried out in accordance

XSL•FO
RenderX

with the Declaration of Helsinki. Ethical approval was obtained from the affiliated university of the corresponding author.

Measures

Illness Representations of IGD

The Brief Illness Perception Questionnaire (B-IPQ) was used to assess the participant's cognitive and emotional perceptions of IGD (Multimedia Appendix 1). The B-IPQ has 8 dimensions; each is measured by 1 item [34]. The items are related to consequences, timeline, personal control, treatment control, identity, illness concern, comprehension, and emotional response. A sample item is "How much does IGD affect the life of an IGD gamer?": 0 (no effect at all) to 10 (severely affects my life). The items have demonstrated adequate test-retest reliability and predictive validity [35]. The Chinese version of the IPQ regarding IGD (IPQ-IGD) has been validated in Chinese adults [33]. Confirmatory factor analysis showed acceptable model fit of B-IPQ at T1 (χ^2:df=2.50, comparative fit index [CFI]=.92, root mean square error of approximation [RMSEA]=.07, and standardized root mean square residual [SRMR]=.06) and at T2 (χ^2:df=2.91, CFI=.94, RMSEA=.06, SRMR=.05). Composite reliability was .79 at T1 and .80 at T2. Average variance extracted was .30 at T1 and .31 at T2. Cronbach α was.71 at T1 and .73 at T2.

IGD Symptoms

IGD symptoms were assessed using the 9 diagnostic criteria proposed in the DSM-5 (Multimedia Appendix 1) [7]. It is a short, user-friendly, self-report measure assessing IGD symptoms of preoccupation, tolerance, withdrawal, unsuccessful attempts to limit gaming, deception or lies about gaming, loss of interest in other activities, use despite knowledge of harm, use for escape or relief of negative mood, and harm based on DSM-5 criteria [7]. Participants answered whether they had experienced symptoms in the last 12 months (0=No, 1=Yes). Higher total scores indicate severer IGD symptoms. Participants who scored ≥5 were classified as having probable IGD [36]. The internal reliability was 0.86 at T1 and 0.80 at T2. The results showed that the intraclass correlation coefficient was 0.70, suggesting moderate test-retest reliability. The tool has been widely used in other IGD studies [1-3]. Like previous studies [2], the continuous variable of IGD was used in the data analyses [1,36,37]. The Chinese version of the scale was found to have good psychometric properties and is widely used in Chinese populations [1,36,37].

Statistical Analysis

Descriptive statistics of and Pearson's correlation analyses between the continuous variables of each dimension of B-IPQ and IGD were performed using SPSS 25.0 (IBM Corp, Armonk, NY). The level of statistical significance was set at .05. Furthermore, the cross-lagged model regarding the relationships between all the dimensions of B-IPQ and the IGD score was performed by structural equation modeling using SPSS Amos 25. Latent variables of B-IPQ were created with the scores of each dimension of B-IPQ being used as observed variables. The variables at the same time point were covariated. The goodness-of-model fit was assessed using the Chi-square:degrees of freedom (χ^2:df) ratio, CFI, RMSEA, and SRMR. A χ^2:df ratio ≤3, CFI ≥.90, RMSEA ≤.08, and/or SRMR ≤.08 would indicate acceptable model fit [38,39].

Results

Descriptions of B-IPQ at T1 and T2

As shown in Table 2, at T1, one-fifth of the participants perceived that IGD would severely affect his or her life (121/591, 20.4%) or would last forever (120/591, 20.3%); 45.9% (271/591) of the participants indicated that IGD would lead to severe symptoms. More than half (343/591, 58.1%) of the participants would be concerned about IGD. More than half (338/591, 57.2%) perceived personal control over IGD, while less than half of the participants believed that treatment of IGD could control the disease (241/591, 40.8%) and felt that they understood IGD (268/591, 45.3%). Only 37.2% (220/591) of the participants indicated that they would develop negative emotions due to IGD.

However, the percentages of item agreement increased at T2. To be specific, 43.0% (254/591) of the participants perceived that IGD would severely affect his or her life; 25.5% (151/591) believed that IGD would last forever; 60.3% (356/591) of the participants indicated that IGD would lead to severe symptoms; and 74.0% (437/591) would be concerned about IGD. Although 68.7% (406/591) perceived personal control over IGD, more than half of the participants believed that treatment of IGD could control the disease (333/591, 56.3%) and felt that they understood IGD (308/591, 52.1%). More than half (308/591, 52.1%) indicated that they would develop negative emotions due to IGD.

XSL•FO
RenderX

Table 2. Number of participants who endorsed each item of the Brief Illness Perception Questionnaire (B-IPQ; score ≥6) at T1 (baseline) and T2 (1-year follow-up; n=591).

B-IPQ items	T1, n (%)	T2, n (%)
Item 1: consequences	121 (20.4)	254 (43.0)
Item 2: timeline	120 (20.3)	151 (25.5)
Item 3: personal control	338 (57.2)	406 (68.7)
Item 4: treatment control	241 (40.8)	333 (56.3)
Item 5: identity	271 (45.9)	356 (60.3)
Item 6: concern	343 (58.1)	437 (74.0)
Item 7: comprehension	268 (45.3)	308 (52.1)
Item 8: emotional response	220 (37.2)	308 (52.1)

Descriptions of IGD at T1 and T2

Of the participants, 10.1% (60/591) and 9.1% (54/591; P=.64) of the participants were classified as having probable IGD at T1 and T2, respectively. Among those without IGD at T1, 4.3% (23/531) were classified with IGD (incidence rate), while among those with IGD at T1, 53.8% (32/60) remitted from IGD at T2 (remission rate).

Correlations Between B-IPQ and IGD Symptoms

As shown in Table 3, the dimensions of consequence (r=0.16), timeline (r=0.27), personal control (r=–0.15), and emotional response (r=0.10) of IPQ at T1 were significantly correlated with IGD at T2 (P<.05). IGD at T1 was significantly correlated with the dimensions of consequences (r=0.17), timeline (r=0.38), personal control (r=–0.30), treatment control (r=–0.10), and concern (r=–0.12) of IPQ at T2.

XSL•FO
RenderX

Table 3. Correlations between Brief Illness Perception Questionnaire (B-IPQ) dimensions and internet gaming disorder (IGD) symptoms at T1 (baseline) and T2 (1-year follow-up).

Variable	Consequence-T1	Timeline-T1	Personal control-T1	Treatment control-T1	Identity-T1	Concern-T1	Comprehension-T1	Emotional response-T1	Consequence-T2	Timeline-T2	Personal control-T2	Treatment control -T2	Identity-T2	Concern-T2	Comprehension-T2	Emotional response-T2	IGD symptoms-T1	IGD symptoms-T2
Consequence- T1																		
r	1	0.61	0.08	0.19	0.47	0.25	0.13	0.40	0.19	0.13	0.00	–0.04	0.12	0.13	0.15	0.18	0.23	0.16
P value	—[a]	<.001	.33	.02	<.001	<.001	.12	<.001	.03	.13	.99	.63	.17	.13	.09	.04	.01	.07
Timeline-T1																		
r	0.61	1	–0.11	0.06	0.18	–0.03	0.03	0.19	0.22	0.30	–0.13	–0.05	0.00	–0.08	–0.06	0.07	0.29	0.27
P value	<.001	—	.20	.51	.03	.75	.69	.02	.01	<.001	.15	.56	.98	.37	.53	.45	<.001	<.001
Personal control-T1																		
r	0.08	–0.11	1	0.49	0.38	0.36	0.36	0.29	–0.07	–0.25	0.17	–0.03	–0.09	–0.11	0.08	–0.10	–0.17	–0.15
P value	.33	.20	—	<.001	<.001	<.001	<.001	<.001	.40	<.001	.06	.76	.31	.23	.36	.23	.02	.08
Treatment control-T1																		
r	0.19	0.06	0.49	1	0.54	0.52	0.31	0.47	0.07	–0.03	–0.01	0.27	0.03	0.11	0.02	–0.04	–0.03	–0.02
P value	.02	.51	<.001	—	<.001	<.001	<.001	<.001	.42	.73	.92	<.001	.77	.20	.80	.67	.75	.79
Identity-T1																		
r	0.47	0.18	0.38	0.54	1	0.58	0.39	0.48	0.29	0.07	–0.01	–0.05	0.08	0.11	0.15	0.15	0.06	0.05
P value	<.001	.03	.00	<.001	—	<.001	<.001	<.001	<.001	.44	.88	.57	.37	.22	.10	.08	.36	.57
Concern-T1																		
r	0.25	–0.03	0.36	0.52	0.58	1	0.51	0.53	0.33	0.00	0.09	0.16	0.10	0.23	0.21	0.27	–0.09	–0.03
P value	<.001	.75	<.001	<.001	<.001	—	<.001	<.001	<.001	.96	.31	.07	.28	.01	.02	<.001	.29	.77
Comprehension-T1																		
r	0.13	0.03	0.36	0.31	0.39	0.51	1	0.39	0.21	0.05	0.12	0.09	0.00	0.16	0.27	0.21	–0.14	–0.02
P value	.12	.69	<.001	<.001	<.001	<.001	—	<.001	.02	.60	.17	.34	.99	.07	<.001	.02	.10	.86
Emotional response-T1																		
r	0.40	0.19	0.29	0.47	0.48	0.53	0.39	1	0.27	0.18	0.06	0.17	0.13	0.13	0.15	0.17	0.06	0.10
P value	<.001	.02	<.001	<.001	<.001	<.001	<.001	—	<.001	.05	.13	.06	.14	.16	.10	.05	.48	.24
Consequence- T2																		
r	0.19	0.22	–0.07	0.07	0.29	0.33	0.21	0.27	1	0.62	0.05	0.28	0.49	0.43	0.39	0.53	0.17	0.25
P value	.03	.01	.40	.42	<.001	<.001	.02	<.001	—	<.001	.53	<.001	<.001	<.001	<.001	<.001	.05	<.001

XSL•FO
RenderX

Variable	Consequence-T1	Timeline-T1	Personal control-T1	Treatment control-T1	Identity-T1	Concern-T1	Comprehension-T1	Emotional response-T1	Consequence-T2	Timeline-T2	Personal control-T2	Treatment control -T2	Identity-T2	Concern-T2	Comprehension-T2	Emotional response-T2	IGD symptoms-T1	IGD symptoms-T2
Timeline-T2																		
r	0.13	0.30	–0.25	–0.03	0.07	0.00	0.05	0.18	0.62	1	–0.12	0.20	0.42	0.22	0.13	0.40	0.38	0.45
P value	.13	<.001	<.001	.73	.44	.96	.60	.05	<.001	—	.17	.02	<.001	.01	.12	<.001	<.001	<.001
Personal control-T2																		
r	0.00	–0.13	0.17	–0.01	–0.01	0.09	0.12	0.06	0.05	–0.12	1	0.51	0.23	0.41	0.32	0.22	–0.30	–0.19
P value	.99	.15	.06	.92	.88	.31	.17	.13	.53	.17	—	<.001	.01	<.001	<.001	.01	<.001	.05
Treatment control -T2																		
r	–0.04	–0.05	–0.03	0.27	–0.05	0.16	0.09	0.17	0.28	0.20	0.51	1	0.33	0.49	0.22	0.32	–0.10	–0.01
P value	.63	.56	.76	<.001	.57	.07	.34	.06	<.001	.02	<.001	—	<.001	<.001	.01	<.001	.24	.90
Identity-T2																		
r	0.12	0.00	–0.09	0.03	0.08	0.10	0.00	0.13	0.49	0.42	0.23	0.33	1	0.52	0.32	0.59	0.06	0.16
P value	.17	.98	.31	.77	.37	.28	.99	.14	<.001	<.001	.01	<.001	—	<.001	<.001	<.001	.06	.06
Concern-T2																		
r	0.13	–0.08	–0.11	0.11	0.11	0.23	0.16	0.13	0.43	0.22	0.41	0.49	0.52	1	0.40	0.57	–0.12	–0.05
P value	.13	.37	.23	.20	.22	.01	.07	.16	<.001	.01	<.001	<.001	<.001	—	<.001	<.001	.16	.55
Comprehension-T2																		
r	0.15	–0.06	0.08	0.02	0.15	0.21	0.27	0.15	0.39	0.13	0.32	0.22	0.32	0.40	1	0.31	–0.05	0.05
P value	.09	.53	.36	.80	.10	.02	<.001	.10	<.001	.12	<.001	.01	<.001	<.001	—	<.001	.58	.60
Emotional response-T2																		
r	0.18	0.07	–0.10	–0.04	0.15	0.27	0.21	0.17	0.53	0.40	0.22	0.32	0.59	0.57	0.31	1	0.08	0.06
P value	.04	.45	.23	.67	.08	<.001	.02	.05	<.001	<.001	.01	<.001	<.001	<.001	<.001	—	.35	.49
IGD symptoms -T1																		
r	0.23	0.29	–0.17	–0.03	0.06	–0.09	–0.14	0.06	0.17	0.38	–0.30	–0.10	0.06	–0.12	–0.05	0.08	1	0.54
P value	.01	<.001	.02	.75	.36	.29	.10	.48	.05	<.001	<.001	.24	.06	.16	.58	.35	—	<.001
IGD symptoms -T2																		
r	0.16	0.27	–0.15	–0.02	0.05	–0.03	–0.02	0.10	0.25	0.45	–0.19	–0.01	0.16	–0.05	0.05	0.06	0.54	1
P value	.07	<.001	.08	.79	.57	.77	.86	.24	<.001	<.001	.05	.90	.06	.55	.60	.49	<.001	—

[a]Not applicable.

XSL•FO
RenderX

Cross-Lagged Model of B-IPQ Dimensions and IGD Symptoms at T1 and T2

The cross-lagged model fit the data well (χ^2/df=2.28, CFI=.95, RMSEA=.06, SRMR=.05; Figure 1). IGD at T1 was positively associated with B-IPQ at T2 (B=.18, β=.10, 95% CI 0.01-0.19, *P*=.04). B-IPQ at T1 and IGD symptoms at T1 explained 8.8% of the variance in B-IPQ at T2. B-IPQ at T1 was not significantly associated with IGD at T2 (*P*=.42). B-IPQ at T1 and IGD symptoms at T1 explained 23.1% of the variance in IGD symptoms at T2. Standardized factor loadings of B-IPQ ranged from .34 to .81 at T1 and from .33 to .76 at T2 (Table 4).

Figure 1. Cross-lagged model of Brief Illness Perception Questionnaire (B-IPQ) and internet gaming disorder (IGD) symptoms at T1 (baseline) and T2 (1-year follow-up) with standardized path coefficients.

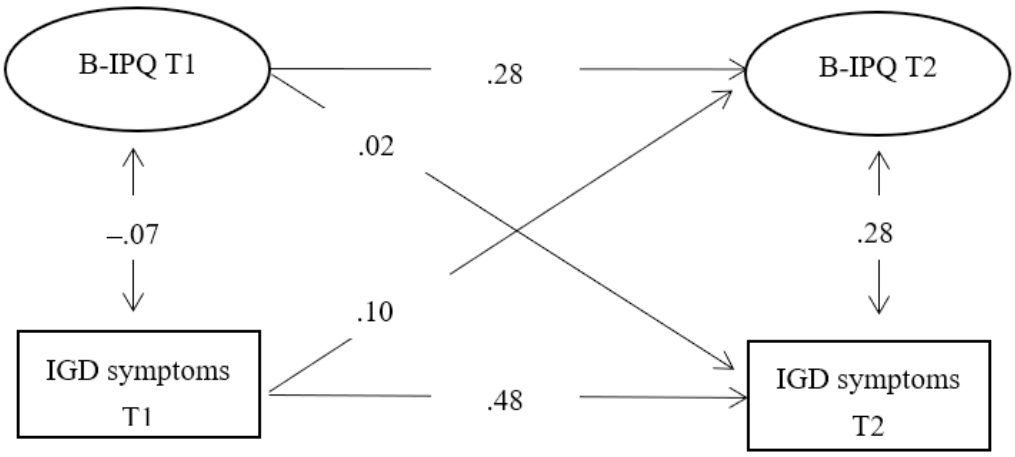


Table 4. Factor loadings of Brief Illness Perception Questionnaire (B-IPQ) at T1 (baseline) and T2 (1-year follow-up).

Factor loadings	B	β	Composite reliability	*P* value
Consequence-T1 <--- B-IPQ T1	1.00	.38	-[a]	-[a]
Timeline-T1 <--- B-IPQ T1	1.22	.34	2.42	.01
Personal control-T1 <--- B-IPQ T1	1.21	.47	7.09	<.001
Treatment control-T1 <--- B-IPQ T1	1.71	.65	8.11	<.001
Identity-T1 <--- B-IPQ T1	1.83	.73	9.13	<.001
Concern-T1 <--- B-IPQ T1	2.17	.81	8.52	<.001
Comprehension-T1 <--- B-IPQ T1	1.45	.58	7.79	<.001
Emotional tesponse-T1 <--- B-IPQ T1	1.68	.68	8.19	<.001
Consequence-T2 <--- B-IPQ T2	1.00	.68	-[a]	-[a]
Timeline-T2 <--- B-IPQ T2	.61	.45	12.41	<.001
Personal control-T2 <--- B-IPQ T2	.43	.33	7.25	<.001
Treatment control -T2 <--- B-IPQ T2	.71	.48	10.19	<.001
Identity-T2 <--- B-IPQ T2	1.07	.74	14.13	<.001
Concern-T2 <--- B-IPQ T2	1.03	.73	14.43	<.001
Comprehension-T2 <--- B-IPQ T2	.65	.48	10.27	<.001
Emotional response-T2 <--- B-IPQ T2	.96	.76	14.80	<.001

[a]In the confirmation factor analysis, the loading of this factor was fixed at 1.

Discussion

Principal Findings

This study adopted a longitudinal design and conducted cross-lagged modeling analyses to demonstrate the relationship between illness representations of IGD and IGD symptoms. The results showed that only IGD symptoms at T1 predicted high levels of unfavorable illness representations at T2. Thus, hypothesis (2) was supported but not hypothesis (1). Indeed, people with greater IGD symptoms may be more likely to experience negative consequences of IGD in their daily life. For example, people with excessive and addictive use of internet games may suffer from mental and emotional disorders (eg, depression, anxiety, social anxiety, suicidal ideation), loneliness, interpersonal conflicts, and lack of offline social connection or

XSL•FO
RenderX

support (eg, [18,19,21,24]). Since such people may have suffered from the IGD symptoms for at least 1 year, they have a better understanding of the fact that IGD can affect their daily life significantly and be chronic and cyclical. This is also consistent with the finding in the cross-sectional study in Macao [33]. In addition, people with IGD may have already experienced failure in attempting to regulate their internet gaming behaviors and thus perceive a great sense of helplessness and hopelessness as well as a low sense of self-control, self-efficacy, and personal control for their internet gaming behaviors and IGD symptoms. Such perceptions may partially explain why IGD may lead to some severe mental problems, such as depression and suicidal ideation (eg, [2,19]). However, illness representations did not show robust significant prediction for IGD symptoms in the model. According to the CSM, illness representations may directly affect risk perceptions and worry about IGD and self-regulation for IGD (eg, self-management, help-seeking), which may in turn influence one's episode of IGD. Future work should test such potential outcomes and mechanisms.

Correlation analyses showed significant correlations between greater consequence, higher timeline, and low personal control at T1 and high levels of IGD symptoms at T2; in turn, IGD symptoms at T1 were also significantly correlated with these perceptions at T2, which seem to suggest longitudinal and interactive relationships between these perceptions and IGD. Other significant correlations included the correlations between negative emotional response at T1 and IGD symptoms at T2, between IGD symptoms at T1 and treatment control at T2, and between IGD symptoms at T1 and concern at T2. Similar correlations between these dimensions of illness representations and IGD symptoms were also found at the same time point. Not all the correlations were statistically significant. Since the dimensions represent different perceptions of a disease and some perceptions may be particularly relevant to the disease, it is reasonable that not all the dimensions of illness representations of IGD were significantly correlated with IGD symptoms. Consistently, the recent cross-sectional study also reported that only timeline cyclical, personal control, illness coherence, and emotional representations were significantly correlated with IGD symptoms [33]. Similar results were also found in previous studies of other diseases (eg, [32,40]). However, most of the correlations had small effect sizes. Thus, these relationships should be interpreted with caution. Since related studies are limited, conclusions regarding the relationships between different dimensions of illness representations and IGD symptoms cannot be drawn at this stage. More empirical studies are warranted.

It is interesting to find that the percentages of item agreement for each dimension of illness representations significantly increased at T2 compared with T1. A plausible explanation for this increase may be that both international and local societies have started to emphasize the adverse consequences of excessive use of the internet and video games in recent years. For example, the World Health Organization defined IGD as a mental disorder in 2018 [9]. In 2018, the Chinese government announced the establishment of a gaming regulator to limit the number of new online games, restrict playing time, and develop an age-restriction system [41]. Also, the local public media have repeatedly criticized video games for negatively influencing young people [41]. Thus, with the changes at the international, governmental, and social levels, local citizens, including college students, may have increasingly recognized the adverse consequences of overuse of internet games and known more about IGD.

The prevalence of IGD found in our study is consistent with another study with Chinese college students (10.3%) [15]. We further found that this prevalence did not significantly change from T1 to T2; a small number of non-IGD cases became IGD cases and vice versa. It may suggest stability of IGD status over time. It is different from another longitudinal survey with 283 Chinese university students that found a significant decrease in the prevalence (14.8% at T1 versus 9.9% at T2) [18]. The different results may be due to the differences in assessments, samples, and study period. However, since we had a relatively small and nonrepresentative sample, the findings should be interpreted with caution. Longitudinal and large-scale studies that monitor the prevalence and conversion rates of IGD across time are warranted.

The levels of illness representations of IGD changed significantly over time, while IGD did not change largely. It seems contradictory to our assumptions regarding their associations. One plausible explanation may be that although the protective factors of IGD (ie, unfavorable illness representations) increased over time, there may exist increasing independent risk factors (eg, daily life stress and academic stress) among the participants that could increase their risk of IGD or moderate the effect of illness representations on IGD. Future studies should investigate both risk and protective factors of IGD to better understand the change in IGD symptoms over time. These perceptions may have great effects on mental health and behavioral outcomes. For example, pessimistic views of IGD, such as low treatment control, may reduce ones' help-seeking intention; a perception of low personal control may lead to anxiety and low self-efficacy. Future studies should test these mental health and behavioral outcomes.

Strengths and Limitations

This study represents the first longitudinal study testing the relationships between illness representations and symptoms of IGD. It is also the first attempt at understanding illness representations of IGD among young people who are a group at high risk of IGD [13-17]. The findings suggest that IGD treatment should not only address IGD symptoms but also these perceptions, such as the perception of low personal control. Cognitive behavioral therapy and motivational enhancement therapy have been promising approaches to reduce addictive behaviors [42] and may help to enhance one's motivation and skills to regulate their gaming behaviors.

Limitations of this study include the use of self-report measures, a nondiagnostic tool for IGD, and a small convenience sample. These might have affected the prevalence of IGD. The sample may differ from other populations considering some important demographic characteristics, such as age, educational level, majors, and regions. For example, the study only included medical students who might have extraordinary capabilities to assess their own health and possible threats of a disease. Thus,

XSL•FO
RenderX

their perceptions of IGD may be different from other populations' perceptions. Second, although the conduct of surveys within classroom settings helped to guarantee the survey's quality, it might come with some limitations such as social pressure to participate. Third, only a 1-year follow-up was conducted. More waves and longer years of follow-up are warranted to better understand the stability of and changes in IGD and its consequences. Fourth, this study used the short version of IPQ. Although this version is more feasible in longitudinal surveys, future studies may use the long version of IPQ to validate our findings and test the reliability of each dimension of IPQ. Also, it was the first to use the B-IPQ in the context of IGD. Future studies should validate this scale well in the context of IGD. B-IPQ and IGD explained a relatively low amount of variance for each other. Future studies should include more potential important variables in the cross-lagged model. Fifth, it is worth noting that this study only investigated illness representations among gamers. Such perceptions may be different from those of nongamers [33]. In addition, it is a limitation that we did not know whether we included diagnosed cases. Since IGD is a newly defined disorder, diagnosis is rare, and treatment is still lacking in most local mental health services, there is little likelihood to include clinically diagnosed cases or those who know that they had IGD. Future studies may test whether illness representations have different implications between gamers and nongamers and between healthy individuals and those with diagnosed IGD. Finally, although the cross-lagged model is widely used to study causal influences in longitudinal panel data, some researchers have argued that, if stability of constructs is to some extent of a trait-like and time-invariant nature, the autoregressive relationships of the cross-lagged model may fail to adequately account for the causal influences [43]. As a consequence, the lagged parameters that are obtained with the cross-lagged model may not represent the actual within-person relationships over time, and this may lead to erroneous conclusions regarding the presence, predominance, and sign of causal influences. Other approaches should be used to validate the findings in future work.

Conclusions

The findings suggest that illness representations of IGD increased over time and the level of IGD symptoms might affect ones' illness representations of IGD. Based on the CSM, these illness representations may further influence gamers' mental health, coping strategies, and behaviors related to the illness (eg, self-management, help-seeking). Future research may examine how illness representations of IGD would influence mental health and behavioral outcomes. Educational programs and psychological interventions are warranted to reduce the over-pessimistic perceptions about IGD based on the CSM structure and the self (eg, low personal control over IGD) among the gamers.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Measures.
[DOCX File , 17 KB - games_v9i4e28117_app1.docx]

References

1. Yang X, Huang B, Wong K. Prevalence and socio-demographic, anthropometric, and cognitive correlates of internet gaming disorder among children in China. Children and Youth Services Review 2021 Mar;122:105893 [FREE Full text] [doi: 10.1016/j.childyouth.2020.105893]
2. Yu Y, Yang X, Wang S, Wang H, Chang R, Tsamlag L, et al. Serial multiple mediation of the association between internet gaming disorder and suicidal ideation by insomnia and depression in adolescents in Shanghai, China. BMC Psychiatry 2020 Sep 23;20(1):460 [FREE Full text] [doi: 10.1186/s12888-020-02870-z] [Medline: 32967648]
3. Yang X, Jiang X, Mo PK, Cai Y, Ma L, Lau JT. Prevalence and interpersonal correlates of internet gaming disorders among Chinese adolescents. Int J Environ Res Public Health 2020 Jan 16;17(2):579 [FREE Full text] [doi: 10.3390/ijerph17020579] [Medline: 31963197]
4. Zhu S, Zhuang Y, Lee P, Li JC, Wong PWC. Leisure and problem gaming behaviors among children and adolescents during school closures caused by COVID-19 in Hong Kong: quantitative cross-sectional survey study. JMIR Serious Games 2021 May 07;9(2):e26808 [FREE Full text] [doi: 10.2196/26808] [Medline: 33960954]
5. Durkin K, Barber B. Not so doomed: computer game play and positive adolescent development. Journal of Applied Developmental Psychology 2002 Nov;23(4):373-392 [FREE Full text] [doi: 10.1016/s0193-3973(02)00124-7]
6. Gamberini L, Barresi G, Maier A, Scarpetta F. A game a day keeps the doctor away: A short review of computer games in mental healthcare. Journal of Cybertherapy and Rehabilitation 2008;1(2):127-145 [FREE Full text]
7. Diagnostic and Statistical Manual of Mental Disorders (DSM–5). Washington, DC: American Psychiatric Association; 2013.
8. Lee N, Kim J, Shin Y, Eom H, Kim M, Kyeong S, et al. Choice of leisure activities by adolescents and adults with internet gaming disorder: development and feasibility study of a virtual reality program. JMIR Serious Games 2020 Dec 11;8(4):e18473 [FREE Full text] [doi: 10.2196/18473] [Medline: 33306033]

XSL•FO
RenderX

9. International Statistical Classification of Diseases and Related Health Problems (ICD). World Health Organization. URL: https://www.who.int/standards/classifications/classification-of-diseases [accessed 2021-11-03]
10. Feng W, Ramo DE, Chan SR, Bourgeois JA. Internet gaming disorder: Trends in prevalence 1998-2016. Addict Behav 2017 Dec;75:17-24 [FREE Full text] [doi: 10.1016/j.addbeh.2017.06.010] [Medline: 28662436]
11. Stevens MW, Dorstyn D, Delfabbro PH, King DL. Global prevalence of gaming disorder: A systematic review and meta-analysis. Aust N Z J Psychiatry 2021 Jun 07;55(6):553-568. [doi: 10.1177/0004867420962851] [Medline: 33028074]
12. Tang CS, Koh YW, Gan Y. Addiction to internet use, online gaming, and online social networking among young adults in China, Singapore, and the United States. Asia Pac J Public Health 2017 Nov 30;29(8):673-682. [doi: 10.1177/1010539517739558] [Medline: 29191049]
13. Shi L, Wang Y, Yu H, Wilson A, Cook S, Duan Z, et al. The relationship between childhood trauma and Internet gaming disorder among college students: A structural equation model. J Behav Addict 2020 Apr 01;9(1):175-180 [FREE Full text] [doi: 10.1556/2006.2020.00002] [Medline: 32101018]
14. Ye N, Zuo B, Zhang L. The relations between internet game addiction and motivation among college students. Heilongjiang Researches on Higher Education 2009;04:97-99.
15. Peng L, Li X. A survey of Chinese college students addicted to video games. China Educ Innov Herald 2009;28:111-112.
16. Wong HY, Mo HY, Potenza MN, Chan MNM, Lau WM, Chui TK, et al. Relationships between severity of internet gaming disorder, severity of problematic social media use, sleep quality and psychological distress. Int J Environ Res Public Health 2020 Mar 13;17(6):1879 [FREE Full text] [doi: 10.3390/ijerph17061879] [Medline: 32183188]
17. Zhao Y, Yu Q, Feng X. Internet game addiction and its related factors among college students. Chinese Journal of Public Health 2009;25(11):1392-1393.
18. Zhang M, Wang X, Yu S, Wu A. Purpose in life, social support, and internet gaming disorder among Chinese university students: A 1-year follow-up study. Addict Behav 2019 Dec;99:106070-106071 [FREE Full text] [doi: 10.1016/j.addbeh.2019.106070] [Medline: 31430620]
19. Liu L, Yao Y, Li CR, Zhang J, Xia C, Lan J, et al. The comorbidity between internet gaming disorder and depression: interrelationship and neural mechanisms. Front Psychiatry 2018 Apr 23;9:154 [FREE Full text] [doi: 10.3389/fpsyt.2018.00154] [Medline: 29740358]
20. Mathews CL, Morrell HER, Molle JE. Video game addiction, ADHD symptomatology, and video game reinforcement. Am J Drug Alcohol Abuse 2019 Jun 06;45(1):67-76. [doi: 10.1080/00952990.2018.1472269] [Medline: 29874473]
21. Bonnaire C, Baptista D. Internet gaming disorder in male and female young adults: The role of alexithymia, depression, anxiety and gaming type. Psychiatry Res 2019 Feb;272:521-530 [FREE Full text] [doi: 10.1016/j.psychres.2018.12.158] [Medline: 30616119]
22. Pearcy BT, McEvoy PM, Roberts LD. Internet gaming disorder explains unique variance in psychological distress and disability after controlling for comorbid depression, OCD, ADHD, and anxiety. Cyberpsychol Behav Soc Netw 2017 Feb;20(2):126-132. [doi: 10.1089/cyber.2016.0304] [Medline: 28085490]
23. Hawi NS, Samaha M, Griffiths MD. Internet gaming disorder in Lebanon: Relationships with age, sleep habits, and academic achievement. J Behav Addict 2018 Mar 01;7(1):70-78 [FREE Full text] [doi: 10.1556/2006.7.2018.16] [Medline: 29486571]
24. Milani L, La Torre G, Fiore M, Grumi S, Gentile DA, Ferrante M, et al. Internet gaming addiction in adolescence: risk factors and maladjustment correlates. Int J Ment Health Addiction 2017 Mar 20;16(4):888-904. [doi: 10.1007/s11469-017-9750-2]
25. Leventhal H, Brissette I, Leventhal EA. The common-sense model of self-regulation of health and illness. In: Cameron LD, Leventhal H, editors. The self-regulation of health and illness behaviour. Oxfordshire, England, UK: Routledge; 2003:42-65.
26. Averous P, Charbonnier E, Lagouanelle-Simeoni M, Prosperi A, Dany L. Illness perceptions and adherence in bipolar disorder: An exploratory study. Compr Psychiatry 2018 Jan;80:109-115 [FREE Full text] [doi: 10.1016/j.comppsych.2017.10.003] [Medline: 29091776]
27. Pereira M, Ferreira G, Machado J, Pedras S. Beliefs about medicines as mediators in medication adherence in type 2 diabetes. Int J Nurs Pract 2019 Oct;25(5):e12768 [FREE Full text] [doi: 10.1111/ijn.12768] [Medline: 31328394]
28. Carletto S, Tesio V, Borghi M, Francone D, Scavelli F, Bertino G, et al. The effectiveness of a body-affective mindfulness intervention for multiple sclerosis patients with depressive symptoms: a randomized controlled clinical trial. Front Psychol 2017 Nov 30;8:2083 [FREE Full text] [doi: 10.3389/fpsyg.2017.02083] [Medline: 29250012]
29. Glattacker M, Heyduck K, Meffert C. Illness beliefs, treatment beliefs and information needs as starting points for patient information--evaluation of an intervention for patients with chronic back pain. Patient Educ Couns 2012 Mar;86(3):378-389. [doi: 10.1016/j.pec.2011.05.028] [Medline: 21719235]
30. Hagger MS, Orbell S. A meta-analytic review of the common-sense model of illness representations. Psychology & Health 2003 Jan;18(2):141-184. [doi: 10.1080/0887044031000081321]
31. Gleason C, Dowling N, Benton S, Kaseroff A, Gunn W, Edwards D. Common sense model factors affecting African Americans' willingness to consult a healthcare provider regarding symptoms of mild cognitive impairment. Am J Geriatr Psychiatry 2016 Jul;24(7):537-546 [FREE Full text] [doi: 10.1016/j.jagp.2015.08.005] [Medline: 26809602]

XSL•FO
RenderX

32. Mo PK, Lau JT. Illness representation on H1N1 influenza and preventive behaviors in the Hong Kong general population. J Health Psychol 2015 Dec 14;20(12):1523-1533. [doi: 10.1177/1359105313516031] [Medline: 24423574]
33. Lau JTF, Dang L, Cheung RYH, Zhang MX, Chen JH, Wu AMS. The psychometric properties of the Revised Illness Perception Questionnaire (IPQ-R) regarding Internet gaming disorder in a general population of Chinese adults. J Behav Addict 2020 Oct 12;9(3):654-663. [doi: 10.1556/2006.2020.00043] [Medline: 32897871]
34. Basu S, Poole J. The Brief Illness Perception Questionnaire. Occup Med (Lond) 2016 Jul 17;66(5):419-420 [FREE Full text] [doi: 10.1093/occmed/kqv203] [Medline: 27317335]
35. Broadbent E, Wilkes C, Koschwanez H, Weinman J, Norton S, Petrie KJ. A systematic review and meta-analysis of the Brief Illness Perception Questionnaire. Psychol Health 2015;30(11):1361-1385. [doi: 10.1080/08870446.2015.1070851] [Medline: 26181764]
36. Ko C, Yen J, Chen S, Wang P, Chen C, Yen C. Corrigendum to "Evaluation of the diagnostic criteria of Internet gaming disorder in the DSM-5 among young adults in Taiwan" [J Psychiat Res 53 (2014) 103–110]. Journal of Psychiatric Research 2014 Oct;57:185. [doi: 10.1016/j.jpsychires.2014.06.005]
37. Sigerson L, Li A, Cheung M, Luk J, Cheng C. Psychometric properties of the Chinese Internet Gaming Disorder Scale. Addict Behav 2017 Nov;74:20-26 [FREE Full text] [doi: 10.1016/j.addbeh.2017.05.031] [Medline: 28558336]
38. Kline RB. Principles and practice of structural equation modeling. New York, NY: Guilford Publications; 2015.
39. Hooper D, Coughlan J, Mullen M. Structural equation modelling: Guidelines for determining model fit. The Electronic Journal of Business Research Methods 2008;6(1):53-60. [doi: 10.21427/D7CF7R]
40. Figueiras MJ, Alves NC. Lay perceptions of serious illnesses: An adapted version of the Revised Illness Perception Questionnaire (IPQ-R) for healthy people. Psychology & Health 2007 Feb;22(2):143-158. [doi: 10.1080/14768320600774462]
41. Video game addiction: China imposes gaming curfew for minors. 2019 Nov 6. URL: https://www.bbc.com/news/world-asia-50315960 [accessed 2021-11-03]
42. Young KS. Cognitive behavior therapy with Internet addicts: treatment outcomes and implications. Cyberpsychol Behav 2007 Oct;10(5):671-679. [doi: 10.1089/cpb.2007.9971] [Medline: 17927535]
43. Hamaker EL, Kuiper RM, Grasman RPPP. A critique of the cross-lagged panel model. Psychol Methods 2015 Mar;20(1):102-116. [doi: 10.1037/a0038889] [Medline: 25822208]

Abbreviations

B-IPQ: Brief Illness Perception Questionnaire
CFI: comparative fit index
CSM: common-sense model
DSM: Diagnostic and Statistical Manual of Mental Disorder
ICD: International Classification of Diseases
IGD: internet gaming disorder
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual

Edited by N Zary; submitted 22.02.21; peer-reviewed by H Lukosch, M Knop; comments to author 14.05.21; revised version received 16.06.21; accepted 09.09.21; published 30.11.21.

Please cite as:
Yang X, Wong KM, She R, Zhao C, Ding N, Xu H, Tu X, Lai X, Zhang G
Relationship Between Illness Representations and Symptoms of Internet Gaming Disorder Among Young People: Cross-Lagged Model
JMIR Serious Games 2021;9(4):e28117
URL: https://games.jmir.org/2021/4/e28117
doi:10.2196/28117
PMID:34851298

©Xue Yang, Kei Man Wong, Rui She, Chengjia Zhao, Nani Ding, Huihui Xu, Xiaolian Tu, Xinyi Lai, Guohua Zhang. Originally published in JMIR Serious Games (https://games.jmir.org), 30.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Feasibility of Virtual Reality Audiological Testing: Prospective Study

Hye Yoon Seol[1,2], AuD; Soojin Kang[1,2], PhD; Jihyun Lim[3], BS; Sung Hwa Hong[2,4], MD, PhD; Il Joon Moon[2,5], MD, PhD

[1]Medical Research Institute, Sungkyunkwan University School of Medicine, Suwon, Republic of Korea

[2]Hearing Research Laboratory, Samsung Medical Center, Seoul, Republic of Korea

[3]Center for Clinical Epidemiology, Samsung Medical Center, Seoul, Republic of Korea

[4]Department of Otolaryngology-Head & Neck Surgery, Samsung Changwon Hospital, Changwon, Republic of Korea

[5]Department of Otolaryngology-Head & Neck Surgery, Sungkyunkwan University School of Medicine, Samsung Medical Center, Seoul, Republic of Korea

Corresponding Author:
Il Joon Moon, MD, PhD
Department of Otolaryngology-Head & Neck Surgery
Sungkyunkwan University School of Medicine
Samsung Medical Center
81 Irwon-ro, Gangnam-gu
Seoul, 06351
Republic of Korea
Phone: 82 2 3410 3579
Email: moonij@skku.edu

Related Article:

Correction of: https://games.jmir.org/2021/3/e26976

(JMIR Serious Games 2021;9(4):e34994) doi:10.2196/34994

In "Feasibility of Virtual Reality Audiological Testing: Prospective Study" (JMIR Serious Games 2021;9(3):e26976) one error was noted.

In the originally published paper, author affiliation 5, applying to Corresponding Author Il Joon Moon, was as follows:

Department of Otolaryngology-Head & Neck Surgery, Samsung Medical Center, Seoul, Republic of Korea

This has been corrected to:

Department of Otolaryngology-Head & Neck Surgery, Sungkyunkwan University School of Medicine, Samsung Medical Center, Seoul, Republic of Korea

Accordingly, the Corresponding Author address for Il Joon Moon has also been corrected from:

Il Joon Moon, MD, PhD
Department of Otolaryngology-Head & Neck Surgery
Samsung Medical Center
81 Irwon-ro, Gangnam-gu
Seoul, 06351
Republic of Korea
Phone: 82 2 3410 3579
Email: moonij@skku.edu

to:

Il Joon Moon, MD, PhD
Department of Otolaryngology-Head & Neck Surgery
Sungkyunkwan University School of Medicine
Samsung Medical Center
81 Irwon-ro, Gangnam-gu
Seoul, 06351
Republic of Korea
Phone: 82 2 3410 3579
Email: moonij@skku.edu

The correction will appear in the online version of the paper on the JMIR Publications website on November 25, 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.

XSL•FO
RenderX

Submitted 15.11.21; this is a non–peer-reviewed article;accepted 17.11.21; published 25.11.21.

Please cite as:
Seol HY, Kang S, Lim J, Hong SH, Moon IJ
Correction: Feasibility of Virtual Reality Audiological Testing: Prospective Study
JMIR Serious Games 2021;9(4):e34994
URL: https://games.jmir.org/2021/4/e34994
doi:10.2196/34994
PMID:34889762

©Hye Yoon Seol, Soojin Kang, Jihyun Lim, Sung Hwa Hong, Il Joon Moon. Originally published in JMIR Serious Games (https://games.jmir.org), 25.11.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on https://games.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX