
JMIR Serious Games

Journal Impact Factor (JIF) (2023): 3.8
Volume 14 (2026) ISSN 2291-9279 Editor-in-Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

- Socially Prescribed Perfectionism, Resilience, and Internet Gaming Disorder in Adolescents: 3-Wave Longitudinal Study ([e93412](#))
Peng Zheng, Zi Tao, Tao Yao, Luoxiang Fang, Min Zhao. 5
- Development and Usability of the “FORTEe Get Strong” App to Promote Physical Activity and Health Awareness in Children and Adolescents With Cancer During Intensive Treatment Using an App-Based Approach: Mixed Methods Study ([e75653](#))
Mareike Kuehn, Lena Wypyrsczyk, Francesca Lanfranconi, Adriana Balduzzi, Hayley Marriott, Alba Solera-Sanchez, Joachim Wiskemann, Nikolai Bauer, Nina Karguth, Barbara Heierer, Norbert Paul, Rodolf Mongondry, Wilhelm Bloch, Katie Rizvi, Martin Fridh, Eva S C Ramos, Carmen Fiuza-Luces, Miriam Gtte, Gabriele Gau, Filippo Spreafico, Barbara Konda, Milica Stefanovi , Elias Dreismickenbecker, Marie Neu, Joerg Faber, FORTEe Consortium. 23
- Usability Study of Augmented Reality Visualization Modalities on Localization Accuracy in the Head and Neck: Randomized Crossover Trial ([e75962](#))
Yao Li, Gijs Luijten, Christina Gsxner, Kim Grunert, Alexis Bader, Frank Hlzle, Rainer Rhrig, Matas de la Fuente, Jan Egger, Kunpeng Xie, Behrus Hinrichs-Puladi. 48
- Development of a Novel Dietary Assessment Method Using Gamification Concepts: Exploratory and Application Study ([e72387](#))
Yulin Zhang, Tianwu Lin, Yuefeng Tan, Xiaojin Shi, Yucheng Yang, Haibing Yang, Xiaona Na, Junhan Zhang, Ai Zhao. 68
- Longitudinal Association Between Internet Gaming Disorder and School Refusal Among Adolescents Using a Random Intercept Cross-Lagged Panel Model: Three-Wave Prospective Cohort Study ([e89619](#))
Peng Zheng, Zi Tao, Luoxiang Fang, Tao Yao, Lingling Zhou, Yutian Wang, Yuannan Lin. 83
- Quantifying Slowness in Parkinson Disease Using a Serious Game: Cross-Sectional Study ([e79463](#))
Luanne Mendes, Ariana Cabral, Camille Alves, Isabela Marques, Alberto Delis, Yann Morre, Adriano Andrade. 99
- Co-Designing Mobile Serious Games to Support Patients With Psoriatic Arthritis and Chronic Pain: Mixed Methods Study ([e75072](#))
Brbara Ramalho, Samuel Gomes, Marta Silva Vicente, Filipa Magalhes, Rodolfo Gonalves Costa, Sandra Gama, Vasileios Charisis, Leontios Hadjileontiadis, Sofia B Dias, iPROLEPSIS Consortium. 121
- Efficacy and Safety of a Video Game–Like Digital Therapy Intervention for Chinese Children With Attention-Deficit/Hyperactivity Disorder: Single-Arm, Open-Label Pre-Post Study ([e76114](#))
Mengyi Bao, Shang Feng, Jiaheng Wang, Jia Ye, Jiangping Wang, Wenyu Li, Kewen Jiang, Lin Yao. 147

Effects of Virtual Nature Embodiment on Compassion, Empathy, Nature Connectedness, and Symptom Burden in Patients With Psychosis or Depression: An Explorative Clinical Study (e74337) Alva Lütt, Kristina Lütgens, Christiane Montag, Stefan Gutwinski, Georg Reuth, Steve Nebel, Felix BERPohl, Pia Spangenberg.	164
Mental Health Professionals' Views on Gaming to Inform Game-Based Interventions: Qualitative Cross-Sectional Study (e69236) Lauri Lukka, Veli-Matti Karhulahti, J Palva.	179
Analysis of Machine Learning–Based Investigation Into Multivariate Factors of Team Performance in Serious Games: Cross-Sectional Retrospective Study (e83478) Gruyff Abdul-Rahman, Freark de Lange, Andrej Zwitter, Noman Haleem.	197
The Effects of Dance-Based Exergaming on Mental Rotation, General Motor Coordination, and Math Achievement in Adolescent Students: Nonrandomized Controlled Pilot Study (e82610) Patrick Fargier, Valérian Cece, Vanessa Lentillon-Kaestner, Cédric Roure.	217
User Experiences of Behavioral and Psychological Change Techniques in a Walking-Based Mobile Exergame: Cross-Sectional Qualitative Study (e78776) Marianna Antoniadou, Helena Orädd, Aseel Berglund, Erik Berglund, Anna Strömberg, Tiny Jaarsma, Leonie Klompstra.	236
Gamified Physical Education and Cognitive Performance Among Chinese Secondary School Students: Cross-Sectional Moderation Mediation Study (e81086) Jingtong Zhang, Sae-Sook Oh, Yaru Xu.	251
Effects of Exergaming Tennis on Players' Tennis Skills and Mental State Compared to Regular Tennis in Adult Players: Quasi-Experimental Study (e73732) Jia Ngo, See Hoe, Maziah Mat Rosly.	276
Design and Iterative Development of Serious Exergames for Children With Autism Spectrum Disorder: Formative Multiple-Case Pilot Study (e77727) Won Kim, Minwoo Seong, SeungJun Kim.	293
Immersive Tai Chi for Home-Based Exercise in Older Adults: Usability and Feasibility Study (e79453) XiaCheng Song, Nazlena Ali, Mohamad Salim, Muhammad Rezaldi.	309
Effects of Game-Based Learning on Piano Music Knowledge Among Elementary School Pupils: Pretest-Posttest Quasi-Experimental Study (e80766) Yun Wang, Wee Tan, Qiaoling Ye, Tingting Gu.	326
Feasibility and User Experience of Immersive Virtual Reality–Based Rehabilitation in Patients With Stroke: Single-Arm Pretest-Posttest Pilot Study (e79584) Sara Arlati, Marta Mondellini, Isabella Martinelli, Eleonora Guanziroli, Mauro Rossini, Marco Sacco, Franco Molteni.	340
Home-Based Immersive Virtual Reality to Improve Motor Performance in Children and Adolescents With Developmental Coordination Disorder: Crossover Study (e84995) Mohammed Alharbi, David Harris, Helen Dodd, Greg Wood, Gavin Buckingham.	357
User Experience of a Virtual Reality–Based Treadmill for Children With a Chronic Disease Affecting Physical Health: Cross-Sectional Feasibility Study (e82953) Capucine Hennequin, Lena Carcreff, Adélie Christiaens, Mickaël Dinomais, Josselin Demas.	374

Analysis of Spatiotemporal Features in a Virtual Navigation Game Across Different Age Groups: Quantitative Research (e83128)
 Xiaofeng Qiao, Shan Tian, Min Tang, Shipai He, Jinghui Wang, Linyuan Fan, Yuanjie Zhu, Zhiyang Zhang, Songjun Du, Chaojie Dong, Yepu Chen, Xiaoyu Liu. 391

Gamified Versus Nongamified Metaverse Learning for Breast Health Knowledge in Women: Randomized Controlled Trial (e75318)
 Rui Li, Sabzali Kahn, Bingyu Duan, Seng Wong. 411

Balance Assessment Using Gamified Digital Technology in Community-Dwelling Older Adults: Mixed Methods Validation Study and Randomized Controlled Trial (e80747)
 Jianan Zhao, Yaqin Xia, Yahui Zhang, Jihong Yu, Chen Chu. 426

A Novel Virtual Reality Intervention Combining Movement Exercises and Body Illusions for the Treatment of Chronic Back Pain: Prospective Feasibility Study (e81051)
 Isabel Neumann, Stefan Lindner, Yevgeniya Nedilko, Ralitsa Zhivova, Michael Gödde, Tobias Tischer-Zeitz, Heike Rittner, Ivo Käthner. 443

Virtual Reality–Based Pain Modulation in Subacute Musculoskeletal Injury: Functional Near-Infrared Spectroscopy Study of Neural and Behavioral Correlates (e77713)
 Ryan Mace, Ziyang Wu, Christine Sieberg, Keerthana Karunakaran, Ke Peng, David Borsook. 463

Design Requirements for Gamified Exercise Apps for Adults With Prehypertension Based on the Octalysis Framework and Self-Determination Theory: Qualitative Interview Study (e86793)
 Sijia Li, Genying Liu, Li Qiao, Hong-Yu Yu, Yu-Yang Zhang, Runyuan Pei, Yue-Tao Lin, Hao-Ming Ma, Mei-Hua Piao. 484

A Haptic-Driven Serious Game for Cognitive Stimulation and Visual Impairment Mitigation in Older Adults Based on the Design-Play-Experience Framework: Cross-Sectional Mixed Methods Pilot Study (e86290)
 Xin Huang, Nazlena Mohamad Ali, Shafrida Sahrani, Yue Zhang. 529

Associations Between Parental Gaming Behaviors and Conversion From Internet Gaming Disorder Noncases to Cases Among Adolescents: Prospective Longitudinal Cohort Study (e80061)
 Joseph Lau, Yanqiu Yu. 545

Conceptualization and Measurement of Withdrawal Symptoms in Gaming Disorder and Development and Psychometric Validation of the Gaming Withdrawal Symptoms Questionnaire: Cross-Sectional Study (e81407)
 Sylwia Starzec, Sławomir piewak, Jolanta Starosta, Paweł Strojny. 558

Reciprocal Within-Person Dynamics Between Internet Gaming Disorder Symptoms, Physical Activity, and Loneliness Among Chinese Adolescent Gamers: Three-Wave Prospective Cohort Study (e87847)
 Chang Hu, Xuan Meng, Xi Wang, Jilin Huang. 580

Viewpoints

The Mechanism and Design Principles of Serious Games in Enhancing Adolescents’ Internet Adaptability (e82505)
 Shijie Gao, Min Jia, Weijun Wang, Jianping Hu, Shihao Ma. 499

Is Gamification the New Panacea for Health Behavioral Changes? Implications for the Health and Life Insurance Industry (e80684)
 Abbas Salami, Tasos Papastilianou, Marvellous Adeoye, John Ronayne, Honor Bixby, Robert Stawski, Bernard Fromson, Matt Doltis, Osama Mahmoud, Mariachiara Di Cesare. 513

Corrigenda and Addendas

Correction: Tongue Muscle Training App for Middle-Aged and Older Adults Incorporating Flow-Based Gameplay: Design and Feasibility Pilot Study (e90410) Kuan-Chu Su, Ko-Chiu Wu, Kuei-Ru Chou, Chia-Hsu Huang.	601
Correction: Children's Improvement After Language and Rhythm Training With the Digital Medical Device Poppins for Dyslexia: Single-Arm Intervention Study (e93737) Charline Grossard, Melanie Descamps, Hugues Pellerin, François Vonthron, David Cohen.	603

Reviews

Research Design Processes in Serious Games for Adolescent Mental Health: Systematic Review (e77173) S Rooij, de, D Kuipers, J Prins, J Pierie.	605
Navigating Virtual Reality in Stroke Rehabilitation: Scoping Review of Diverse Intervention Effects (e72498) Ruixue Liu, Beibei Lin, Yanhuan Zhai, Wei Wang.	625
Effectiveness of Gamified Swallowing Exercises in Adults With Dysphagia: Systematic Review and Meta-Analysis of Randomized Controlled Trials (e82017) Jijia Liu, Meijun Ou, Mengyao Yan, Xianghua Xu, Li Qin.	638
Game Elements in Military Trauma Care Education: Systematic Review (e79163) Natalia Stathakarou, Andrzej Kononowicz, Maxine Harjani, Cara Swain, Klas Karlgren.	662

Research Letter

RedMan-GreenMan: Co-Designed Pedestrian Safety Game Prototype for Children With Autism (e69260) Jason Stanton, Nicole Peel, Caroline Mills, Paul Breen.	682
--	-----

Socially Prescribed Perfectionism, Resilience, and Internet Gaming Disorder in Adolescents: 3-Wave Longitudinal Study

Peng Zheng*; Zi Tao*; Tao Yao*; Luoxiang Fang; Min Zhao

Second Affiliated Hospital & Yuying Children's Hospital of Wenzhou Medical University, No. 109 Xueyuan West Road, Wenzhou, Zhejiang Province, China

*these authors contributed equally

Corresponding Author:

Min Zhao

Second Affiliated Hospital & Yuying Children's Hospital of Wenzhou Medical University, No. 109 Xueyuan West Road, Wenzhou, Zhejiang Province, China

Abstract

Background: Internet gaming disorder (IGD) is increasingly prevalent among adolescents. Although socially prescribed perfectionism (SPP) and resilience are both related to IGD, longitudinal evidence on their temporal relationships and underlying mechanisms remains limited.

Objective: This study aimed to examine the longitudinal associations among SPP, resilience, and IGD in Chinese adolescents; test the mediating role of resilience; and explore potential sex differences.

Methods: A 3-wave prospective longitudinal study was conducted among students from 4 middle schools in Zhejiang Province, China. Adolescents who had played online games in the past 12 months were recruited using convenience sampling. Data were collected at 6-month intervals: time 1 (T1; March 2024), time 2 (T2; September 2024), and time 3 (T3; March 2025). A total of 1332 Chinese adolescents (875/1332, 65.7% male; mean age 13.61, SD 0.70 years) participated in the baseline survey. SPP, resilience, and IGD were assessed using the Hewitt-Flett Multidimensional Perfectionism Scale-Short Form, the 10-item Connor-Davidson Resilience Scale, and the 9-item *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition) IGD Checklist, respectively. Cross-lagged panel model (CLPM) and multigroup analyses across sex were conducted.

Results: All statistical tests were 2-tailed with $\alpha=.05$. The CLPM demonstrated good fit to the data ($\chi^2_{38}=163.34$; comparative fit index [CFI]=0.945; Tucker-Lewis index [TLI]=0.932; root mean square error of approximation [RMSEA]=0.054; standardized root mean square residual [SRMR]=0.047). Higher SPP predicted later IGD (T1→T2: $\beta=0.10$, 95% CI 0.04 to 0.16, $P<.001$; T2→T3: $\beta=0.09$, 95% CI 0.03 to 0.15, $P=.004$) and lower resilience (T1→T2: $\beta=-0.09$, 95% CI -0.16 to -0.02 , $P=.007$; T2→T3: $\beta=-0.12$, 95% CI -0.18 to -0.06 , $P<.001$). In contrast, SPP was not significantly predicted by prior IGD nor resilience. Higher resilience predicted lower subsequent IGD (T1→T2: $\beta=-0.09$, 95% CI -0.15 to -0.03 , $P=.001$; T2→T3: $\beta=-0.09$, 95% CI -0.15 to -0.03 , $P=.001$), whereas higher IGD predicted lower subsequent resilience (T1→T2: $\beta=-0.19$, 95% CI -0.27 to -0.11 , $P<.001$; T2→T3: $\beta=-0.09$, 95% CI -0.15 to -0.03 , $P=.003$). Bootstrapped mediation analysis showed a significant indirect effect of SPP at T1 on IGD at T3 via resilience at T2 ($\beta=0.008$, 95% CI 0.004 to 0.012, $P=.005$). However, multigroup analyses revealed no significant sex differences.

Conclusions: This study provides novel insights into the longitudinal associations among SPP, resilience, and IGD in adolescents. Unlike previous research based mainly on cross-sectional data, this 3-wave CLPM study clarifies the temporal relationships among these variables and shows that resilience mediates the association between SPP and subsequent IGD. These findings advance the field by identifying a temporal psychological pathway underlying adolescent IGD. They also have practical implications for early screening and for developing resilience-focused interventions for adolescents at risk of IGD.

(*JMIR Serious Games* 2026;14:e93412) doi:[10.2196/93412](https://doi.org/10.2196/93412)

KEYWORDS

socially prescribed perfectionism; internet gaming disorder; resilience; sex differences; adolescents

Introduction

Background

With the rapid expansion of digital entertainment, online gaming has become one of the most prevalent leisure activities among

adolescents [1]. For a subset of individuals, however, excessive and poorly controlled gaming may develop into a maladaptive behavioral pattern with significant functional impairment [2]. Internet gaming disorder (IGD) refers to a persistent and recurrent pattern of gaming behavior marked by impaired control, prioritization of gaming over other activities, and

continuation of play despite negative consequences [3]. Recent epidemiological evidence indicates that IGD has become an increasingly salient public health concern among adolescents. A recent meta-analysis indicated that the pooled prevalence of IGD is about 10% among Chinese adolescents [4]. Empirical research further indicates that IGD is associated with a range of negative outcomes, including impaired academic performance, sleep disturbances, depression, and anxiety [5-7]. These findings highlight the importance of identifying the antecedents of IGD to inform early prevention and intervention efforts.

Socially Prescribed Perfectionism and IGD

Perfectionism is widely recognized as a multidimensional construct [8]. According to the Hewitt-Flett model, it comprises self-oriented perfectionism, which involves imposing exceptionally high standards on oneself; other-oriented perfectionism, which refers to placing unrealistic standards on others; and socially prescribed perfectionism (SPP), which reflects the belief that others expect one to be perfect and that acceptance depends on meeting these external standards [9-11]. Although all 3 dimensions may be relevant to adolescent development, this study focused specifically on SPP. Among these dimensions, SPP is most directly linked to external pressure, perceived social demands, and evaluative concerns, making it particularly salient during adolescence, a developmental period marked by heightened sensitivity to interpersonal evaluation and social approval [10]. In this study, SPP was conceptualized as a socially driven cognitive-personality vulnerability factor that reflects adolescents' perceived external expectations and evaluative pressure [12]. Prior research has consistently identified SPP as the most maladaptive dimension of perfectionism, showing robust associations with poorer psychological well-being, greater psychological distress, and behavioral maladjustment [13,14]. In addition, SPP among adolescents and young adults has increased substantially over time, highlighting its growing developmental significance [15].

The focus on SPP is also theoretically relevant for understanding adolescents' vulnerability to IGD. According to the Perfectionism Social Disconnection Model, individuals with high levels of SPP tend to perceive intense external demands and fear harsh evaluation from others [12]. To avoid criticism and preserve a flawless image, they may engage in self-concealment and emotional suppression [16], which can undermine authentic interpersonal relationships and contribute to persistent feelings of social disconnection [12]. These interpersonal and emotional difficulties may, in turn, increase adolescents' reliance on online environments as a means of coping [17]. From the perspective of compensatory internet use theory [18], adolescents may turn to online gaming to escape real-life stress, regulate negative emotions, and compensate for unmet social and psychological needs. For adolescents with high levels of SPP, online games may provide a relatively controllable context in which evaluation can be avoided or managed more easily while also offering temporary experiences of competence, achievement, and social connection [19]. Over time, this pattern of using gaming to cope with distress and interpersonal insecurity may increase vulnerability to

problematic gaming behaviors and IGD [20]. Therefore, SPP may be understood as a socially driven vulnerability factor that contributes to the development of IGD in adolescents.

Empirical evidence supports the association between perfectionistic tendencies and problematic online behaviors among adolescents and university students. Studies have shown that maladaptive perfectionism is positively associated with IGD symptoms and higher levels of internet-related behavioral problems [21,22]. Research also indicates that individuals with stronger perfectionistic concerns rely more on online activities to regulate negative affect [23]. Additionally, a 12-month follow-up study involving 465 highly engaged Australian gamers found that higher baseline levels of perfectionistic gaming-related cognitions significantly predicted subsequent increases in problematic gaming behaviors [24]. Despite these observations, most studies examining perfectionism and IGD remain cross-sectional, limiting the ability to determine temporal direction or causal mechanisms. Besides, few studies have directly examined the association between SPP and IGD. Therefore, it is necessary to use longitudinal methods to clarify such association among adolescents.

Resilience as a Potential Mediator

Another notable gap in the literature is the insufficient research exploring the psychological mechanisms linking SPP to IGD. Although previous studies have identified SPP as a risk factor for various maladaptive behavioral outcomes, including problematic internet and gaming behaviors, the mediating pathway that explains this association remains underexamined. One potential mediator is resilience. Resilience is generally defined as a positive psychological trait that enables individuals to adapt successfully to stress, adversity, or failure through effective coping strategies and emotional regulation and may play a key role in this relationship [25,26]. Rather than representing a fixed personal attribute, resilience reflects a dynamic capacity to recover from setbacks, maintain psychological stability, and mobilize internal and external resources when confronted with perfectionism-related stress [27].

The model of compensatory internet use [18] provides a useful theoretical framework for understanding why resilience may mediate the association between SPP and IGD. It proposes that individuals may engage in online activities not only for enjoyment but also to compensate for offline stress, negative emotions, and unmet psychological needs [18,28]. In other words, problematic online behaviors may emerge when the internet is used as a coping tool to manage distress that individuals feel unable to handle effectively in real life [18]. From this perspective, resilience is highly relevant because it reflects the capacity to cope adaptively with stress and emotional challenges [29]. Adolescents with higher resilience are generally better able to regulate negative emotions, recover from setbacks, and rely on effective coping strategies [30], which may reduce the need to use online gaming as an escape or emotional outlet [31]. In contrast, adolescents with lower resilience may be less able to manage perfectionism-related stress in adaptive ways and therefore more likely to turn to online gaming as a compensatory strategy to avoid pressure, relieve distress, or

obtain temporary feelings of competence and control [20]. This process may be particularly relevant for adolescents with high levels of SPP and who tend to perceive elevated external expectations, fear negative evaluation, and experience greater emotional vulnerability in academic, social, and performance contexts [32]. Over time, repeated reliance on gaming for emotional relief may reinforce maladaptive coping patterns and increase the risk of IGD, especially given the immersive and performance-based features of online games, which provide immediate feedback, achievement, and social affirmation [20].

This interpretation is also consistent with the stress-vulnerability-protective factors model, which emphasizes that psychological outcomes are shaped by the interaction between stressors, vulnerability factors, and protective resources [33]. Within this framework, SPP may be understood as a stress-related vulnerability factor [12], whereas resilience may function as a protective resource that helps adolescents maintain psychological adjustment in the face of external pressure and evaluative stress [30]. Accordingly, adolescents with high levels of SPP may be at greater risk of IGD partly because perfectionism-related stress undermines resilience, thereby increasing reliance on maladaptive coping behaviors such as online gaming.

Existing empirical findings provide support for these theoretical accounts. Prior research has shown that higher levels of SPP are associated with lower resilience, as individuals who fear external evaluation or criticism tend to exhibit heightened stress sensitivity and reduced adaptive coping skills [34]. Resilience, in turn, has been found to buffer against addictive or compulsive technology use. Several studies have demonstrated that lower resilience predicts greater vulnerability to problematic gaming and internet behaviors among adolescents and young adults [31,35]. A large-scale study found that stress originating from adverse childhood experiences increased the risk of IGD by undermining resilience, supporting a stress-resilience pathway to IGD [36]. Complementary findings suggest that resilience and stress jointly influence IGD tendencies [37], highlighting the importance of coping resources in shaping susceptibility to gaming-related problems. Despite these advances, important gaps remain. Most existing studies have relied on cross-sectional designs, which capture only static associations and cannot establish temporal precedence among SPP, resilience, and IGD. In addition, few studies have integrated the compensatory internet use framework with longitudinal tests of resilience as a mediator between SPP and IGD.

Sex Differences

Researchers have highlighted that sex plays a significant role in shaping vulnerability to problematic gaming. Specifically, males tend to report higher levels of IGD than females [38], show greater preference for performance-oriented gaming activities, and exhibit stronger competitiveness in digital environments [39]. These sex differences may reflect distinct psychological needs that drive gaming engagement. Females are more likely to engage in digital activities for interpersonal connection and emotional satisfaction, whereas males tend to use gaming to pursue competition, achievement, and self-enhancement motives [40]. Moreover, researchers have

suggested that the situational factors contributing to addictive gaming may vary by sex. For instance, SPP has been shown to predict problematic online behaviors through heightened concern with external evaluation, and this effect appears particularly salient among males [41]. Similarly, the protective effects of psychological resilience against problematic internet use or problematic gaming have been found to be stronger among females than males [41]. However, most of these studies have been cross-sectional, limiting our understanding of whether these sex-related mechanisms persist over time. Although sex differences in general problematic internet behaviors have been widely examined, very few studies have specifically investigated whether these differences extend to the longitudinal associations between SPP, resilience, and IGD. Given that IGD strongly emphasizes achievement-oriented rewards and performance feedback, it is important to examine whether similar sex-specific patterns exist in these pathways.

This Study

Given this background, this study used a cross-lagged panel model (CLPM) with a 3-wave longitudinal design to examine the longitudinal association between SPP and IGD and further test whether resilience mediates this relationship. It also explored potential sex differences in the associations among SPP, resilience, and IGD in adolescents. The following hypotheses were proposed: (H1) Higher levels of SPP positively predict subsequent IGD across adjacent waves; (H2) resilience mediates the longitudinal relationship between SPP and IGD; and (H3) the associations among SPP, resilience, and IGD differ by sex.

Methods

Study Design, Participants, and Setting

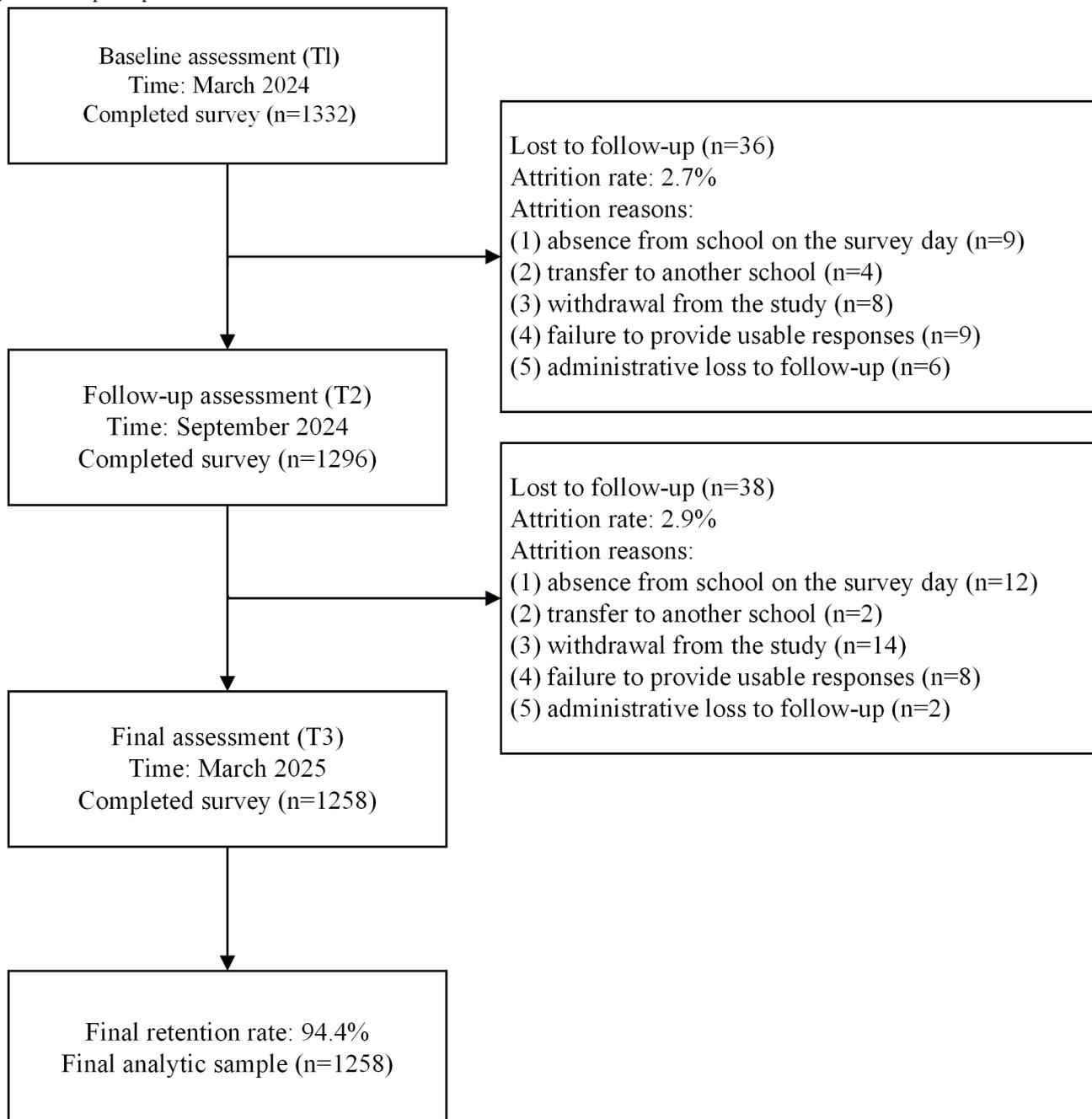
This study was designed as a 1-year prospective cohort study and conducted among students from 4 middle schools in Zhejiang Province, China. The recruitment setting was school-based, and participants were recruited from regular classroom populations in these 4 schools. The schools were selected based on accessibility and willingness to participate in the study. Data were collected at three 6-month intervals: time 1 (T1; March 2024), time 2 (T2; September 2024), and time 3 (T3; March 2025). The 6-month interval between waves was informed by prior adolescent longitudinal studies using similar follow-up periods [42-44] and was considered appropriate for capturing meaningful short-term changes in psychological and behavioral variables in a school-based design. Participants were recruited through convenience sampling, with all eligible students in the selected schools invited to participate during regular class hours through coordination with teachers and school administrators. Trained field researchers administered the surveys during regular class hours using self-report questionnaires. Prior to survey administration, both students and their parents (or legal guardians) were provided with comprehensive information about the study's aims, procedures, and confidentiality safeguards. Written informed consent was obtained from parents or guardians, while students were informed that submitting the completed questionnaire would be interpreted as assent. It was emphasized that participation was voluntary and that nonparticipation or withdrawal would

have no consequences on their academic standing or school records.

At baseline (T1), 1332 adolescents (875/1332, 65.7% males; mean age 13.61, SD 0.70 years) completed the survey. At T2, 1296 participants remained in the study, indicating that 36 participants were lost between T1 and T2. The reasons for attrition from T1 to T2 were as follows: absence from school on the survey day (n=9), transfer to another school (n=4), withdrawal from the study (n=8), failure to provide usable responses (n=9), and administrative loss to follow-up (n=6). By T3, 1258 participants had completed all 3 assessments,

indicating that a further 38 participants were lost between T2 and T3. The reasons for attrition from T2 to T3 were as follows: absence from school on the survey day (n=12), transfer to another school (n=2), withdrawal from the study (n=14), failure to provide usable responses (n=8), and administrative loss to follow-up (n=2). Overall, 74 participants were lost across the 3 waves. The participant flowchart is presented in [Figure 1](#). Participants who provided valid data at least once were retained in the analyses to maximize statistical power and preserve sample representativeness. Missing data patterns and the rationale for the handling of missingness are reported in the Results section.

Figure 1. The participant flowchart.



Ethical Considerations

This study was approved by the Ethics Committee of the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University (approval number: XMSQ-2024 - 347). Before data collection, both students and their parents or legal guardians were provided with detailed information about the study aims, procedures, voluntary nature of participation, and confidentiality protections. Written informed consent was obtained from parents or legal guardians, and students were informed that submission of the completed questionnaire would be considered assent to participate. To protect privacy and confidentiality, student ID numbers were collected only for the purpose of matching responses across the 3 survey waves; these identifiers were stored securely and removed from the analytic dataset before analysis to ensure de-identification. Participation was voluntary, and participants could decline or withdraw without any academic consequences. No financial nor material compensation was provided for participation. In addition, no images in the manuscript nor supplementary materials contain identifiable information about individual participants.

Studied Variables and Measurements

Demographic Variables

Demographic variables included age, sex, parental educational attainment, self-reported household income level, academic performance, single-parent family status, and average gaming time per month.

Internet Gaming Disorder

IGD was measured using the 9-item *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition; *DSM-5*) IGD Checklist [45], which assesses gaming-related symptoms occurring within the past 12 months using a binary (yes/no) response format. The instrument evaluates core diagnostic features, including excessive preoccupation with gaming, withdrawal symptoms, tolerance, impaired control, prioritization of gaming over other activities, continued gaming despite adverse consequences, misrepresentation of gaming involvement, use of gaming as an escape from negative moods, and functional impairment resulting from gaming behavior. The checklist has been validated in adolescent populations and has demonstrated adequate reliability and validity [46]. For analysis, the 9 items were summed to create an observed total score reflecting IGD symptom severity, with higher scores indicating greater severity. In this study, the Cronbach α of this scale was 0.70 at T1, 0.75 at T2, and 0.81 at T3.

Resilience

Resilience was measured using the 10-item short form of the Connor-Davidson Resilience Scale, which assesses individuals' capacity to adapt to stressors such as life changes, personal challenges, illness, pressure, failure, and emotional distress [47]. The Chinese version of the scale has demonstrated strong psychometric reliability and validity in prior research [48,49]. Items are rated on a 5-point Likert scale ranging from 0 ("not true at all") to 4 ("true nearly all of the time"), yielding total scores between 0 and 40, with higher scores indicating greater resilience. In this study, the Cronbach α of this scale was 0.82 at T1, 0.85 at T2, and 0.90 at T3.

Socially Prescribed Perfectionism

SPP was assessed using the corresponding 5-item subscale of the Hewitt-Flett Multidimensional Perfectionism Scale Short Form [50]. Participants rated each item on a 7-point Likert scale ranging from 1 ("strongly disagree") to 7 ("strongly agree"), with higher total scores indicating greater levels of SPP. The Chinese version of this subscale has demonstrated good reliability and validity in previous studies [51]. In this study, the Cronbach α of this scale was 0.83 at T1, 0.86 at T2, and 0.88 at T3.

Preliminary Data Analyses

Attrition analyses were performed using χ^2 tests for categorical variables and independent-samples t tests for continuous variables. Pearson correlation coefficients were calculated to examine the relationships among the main study variables. In the descriptive and correlational analyses conducted in SPSS (IBM Corp), listwise deletion was applied to address missing data. Given that the proportion of missing data was minimal (approximately 1%), this approach was considered suitable and unlikely to bias the results or substantially reduce the effective sample size.

Measurement Invariance

Longitudinal measurement invariance was examined for SPP, resilience, and IGD. An acceptable configural invariance model was determined based on standard model fit criteria, including a comparative fit index (CFI) ≥ 0.90 and root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) values ≤ 0.08 [52]. Metric and scalar invariance were supported when changes in model fit indices remained within recommended thresholds ($\Delta\text{CFI} \leq 0.01$; $\Delta\text{RMSEA} \leq 0.015$), indicating stability of factor loadings and item intercepts across time points and sex groups. Because scalar invariance was supported, the constructs were considered comparable across waves and were retained for the subsequent longitudinal structural analyses.

CLPM and Mediation Analysis

In the longitudinal models, IGD was treated as an observed continuous summed score representing symptom severity rather than a categorical diagnostic variable. To investigate the longitudinal relationships among SPP, resilience, and IGD, an a priori CLPM was specified, which is well suited for testing directional and mediational effects in longitudinal datasets [52,53]. The model included autoregressive paths for SPP, resilience, and IGD across adjacent waves; cross-lagged paths among the 3 constructs across adjacent waves; and within-wave correlations among SPP, resilience, and IGD at each time point. Indirect effects were evaluated using bootstrapping procedures with 5000 resamples to generate 95% bias-corrected confidence intervals. Model adequacy was assessed using multiple goodness-of-fit indices, including $\chi^2(df) \leq 5$, CFI and Tucker-Lewis index (TLI) ≥ 0.90 , and RMSEA and SRMR ≤ 0.08 , consistent with established structural equation modeling guidelines [54].

Multigroup Analyses

Multigroup analyses were conducted to examine sex differences in the longitudinal associations among SPP, resilience, and IGD. A series of nested models, each constraining a specific pathway, were compared with an unconstrained model in which all parameters were freely estimated. Sex differences were evaluated by comparing the fit of the unconstrained and constrained models. Following commonly recommended criteria for measurement and structural invariance testing, ΔCFI and ΔRMSEA were used as the primary indicators of model differences, with $\Delta\text{CFI}>0.01$ and $\Delta\text{RMSEA}>0.015$ indicating meaningful deterioration in model fit and, therefore, significant sex differences in the examined associations.

Software, Estimation, and Missing Data Handling

Preliminary analyses were performed using SPSS version 26.0, and all longitudinal structural analyses were performed using Mplus 8.3. Before model estimation, the distributions of SPP, resilience, and IGD were examined across the 3 waves. Because no severe departures from normality were identified, maximum likelihood estimation was considered appropriate for the longitudinal analyses. For the longitudinal models estimated in Mplus, full information maximum likelihood (FIML) analysis was used to address missing data by retaining all available observations in model estimation [55]. No post hoc model respecification was undertaken for the primary longitudinal model. Statistical significance was set at a 2-tailed P value $<.05$.

Sensitivity Analysis

In addition, multiple imputation was performed as a sensitivity analysis to examine the robustness of the findings to the handling of missing data. This approach has been widely used in longitudinal studies [56-58]. The results obtained from the imputed datasets were highly similar to those from the primary FIML analyses, indicating that the main findings were robust.

Results

Missing Data

At T1, the sample included 1332 adolescents; missing data were observed for age (11/1332, 0.8%), sex (4/1332, 0.3%), father's educational level (11/1332, 0.8%), mother's educational level (8/1332, 0.6%), academic performance (11/1332, 0.8%), SPP (7/1332, 0.5%), resilience (10/1332, 0.8%), and IGD (14/1332,

1.1%). The following variables did not have any missing data: self-reported household income, single-parent family status, and gaming time per month. At T2, the sample included 1296 adolescents.

Missing data at T2 were as follows: age, 9/1296, 0.7%; sex, 6/1296, 0.5%; father's educational level, 12/1296, 0.9%; mother's educational level, 9/1296, 0.7%; self-reported household income, 15/1296, 1.2%; academic performance, 7/1296, 0.5%; single-parent family status, 14/1296, 1.1%; gaming time per month, 8/1296, 0.6%; SPP, 14/1296, 1.1%; resilience, 8/1296, 0.6%; and IGD, 9/1296, 0.7%.

At T3, the sample included 1258 adolescents. Missing data at T3 were as follows: age, 6/1258, 0.5%; sex, 4/1258, 0.3%; father's educational level, 7/1258, 0.6%; mother's educational level, 7/1258, 0.6%; self-reported household income, 10/1258, 0.8%; academic performance, 6/1258, 0.5%; single-parent family status, 11/1258, 0.9%; gaming time per month, 6/1258, 0.5%; SPP, 10/1258, 0.8%; resilience, 12/1258, 1%; and IGD, 9/1258, 0.7%. The Little test for missing completely at random was nonsignificant ($\chi^2_{18}=15.55$, $P=.62$), indicating that the missing data pattern was consistent with the missing completely at random assumption. Therefore, FIML was considered appropriate and was used in the longitudinal analyses to retain all available data.

Distributions

Before fitting the CLPM, the distributions of the main study variables were examined across the 3 waves. The skewness and kurtosis values for SPP, resilience, and IGD indicated no severe departures from normality. Therefore, the use of maximum likelihood estimation in the subsequent longitudinal analyses was considered appropriate.

Attrition Analyses

As shown in Table 1, attrition analyses comparing participants who remained in the study with those lost to follow-up revealed no statistically significant differences in age, sex, parental educational level, self-reported household income, academic performance, single-parent family status, gaming time per month, SPP, resilience, and IGD at baseline (all $P>.05$). These findings suggest that attrition was unlikely to substantially influence the subsequent longitudinal analyses.

Table . Attrition analyses comparing adolescents who completed all 3 waves of the longitudinal study and those lost to follow-up in a study of socially prescribed perfectionism (SPP), resilience, and internet gaming disorder (IGD).

Variable	Follow-up (n=1258)	Lost to follow-up (n=74)	<i>P</i> value ^a
Sex, n (%)			.92
Male	826 (65.8)	49 (67.1)	
Female	429 (34.2)	24 (32.9)	
Father's educational level, n (%)			.83
Middle school or less	819 (65.7)	46 (62.2)	
Senior high school	338 (27.1)	22 (29.7)	
College or higher	90 (7.2)	6 (8.1)	
Mother's educational level, n (%)			.19
Middle school or less	835 (66.8)	57 (77)	
Senior high school	309 (24.7)	13 (17.6)	
College or higher	106 (8.5)	4 (5.4)	
Self-reported household income level, n (%)			.11
Below average	260 (20.7)	11 (14.9)	
Average	711 (56.5)	51 (68.9)	
Above average	287 (22.8)	12 (16.2)	
Self-reported academic performance (percentile), n (%)			.09
0-20th	169 (13.6)	6 (8.1)	
21st-40th	271 (21.7)	11 (14.9)	
41st-60th (average)	324 (26)	17 (23)	
61st-80th	280 (22.5)	21 (28.4)	
81st-100th	203 (16.3)	19 (25.7)	
Single-parent family status, n (%)			.64
No	961 (76.4)	53 (71.6)	
Yes	137 (10.9)	10 (13.5)	
Not reported	160 (12.7)	11 (14.9)	
Gaming time per month (hours), n (%)			.78
<4	557 (44.3)	35 (47.3)	
4 - 8	352 (28)	20 (27)	
8 - 12	170 (13.5)	8 (10.8)	
12 - 16	152 (12.1)	8 (10.8)	
>16	27 (2.1)	3 (4.5)	
Age (years), mean (SD)	13.61 (0.69)	13.67 (0.74)	.49
SPP, mean (SD)	16.83 (4.39)	17.07 (4.39)	.65
Resilience, mean (SD)	25.65 (9.90)	26.85 (8.45)	.32
IGD, mean (SD)	1.61 (1.78)	1.85 (1.87)	.24

^aBased on available cases for each variable; missing data were excluded from the corresponding analysis.

Descriptive Statistics of Main Variables

Across the 3 waves, mean values at T1, T2, and T3 were 16.83 (SD 4.39), 16.76 (SD 4.37), and 16.93 (SD 4.43), respectively,

for SPP; 25.67 (SD 9.91), 25.55 (SD 9.61), and 25.59 (SD 10.18), respectively, for resilience; and 1.60 (SD 1.78), 1.52 (SD 1.86), and 1.66 (SD 2.12), respectively, for IGD.

Correlation Analyses

As shown in Table 2, SPP was significantly and positively correlated with IGD across the 3 waves ($r=0.24$ to 0.35 , all

$P<.001$). Resilience was significantly and negatively correlated with SPP and IGD across 3 waves ($r=-0.28$ to -0.11 , all $P<.001$).

Table . Pearson correlation analyses among socially prescribed perfectionism (SPP), resilience, and internet gaming disorder (IGD) across the 3 waves of the longitudinal study among Chinese adolescents.

	SPP at T1	SPP at T2	SPP at T3	Resilience at T1	Resilience at T2	Resilience at T3	IGD at T1	IGD at T2	IGD at T3
SPP at T1									
<i>r</i>	1	0.57	0.52	-0.28	-0.24	-0.17	0.30	0.27	0.24
<i>P</i> value	— ^a	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001
SPP at T2									
<i>r</i>	0.57	1	0.55	-0.21	-0.24	-0.20	0.23	0.35	0.29
<i>P</i> value	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001	<.001
SPP at T3									
<i>r</i>	0.52	0.55	1	-0.23	-0.24	-0.28	0.22	0.27	0.32
<i>P</i> value	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001	<.001
Resilience at T1									
<i>r</i>	-0.28	-0.21	-0.23	1	0.45	0.39	-0.13	-0.12	-0.11
<i>P</i> value	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001	<.001
Resilience at T2									
<i>r</i>	-0.24	-0.24	-0.24	0.45	1	0.48	-0.12	-0.12	-0.13
<i>P</i> value	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001	<.001
Resilience at T3									
<i>r</i>	-0.17	-0.20	-0.28	0.39	0.48	1	-0.13	-0.14	-0.13
<i>P</i> value	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001	<.001
IGD at T1									
<i>r</i>	0.30	0.23	0.22	-0.13	-0.12	-0.13	1	0.50	0.44
<i>P</i> value	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001	<.001
IGD at T2									
<i>r</i>	0.27	0.35	0.27	-0.12	-0.12	-0.14	0.50	1	0.55
<i>P</i> value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—	<.001
IGD at T3									
<i>r</i>	0.24	0.29	0.32	-0.11	-0.13	-0.13	0.44	0.55	1
<i>P</i> value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	—

^aNot applicable.

Longitudinal Invariance Test

Table 3 presents the results of the longitudinal measurement invariance tests for SPP, resilience, and IGD. The configural invariance models demonstrated satisfactory model fit across all 3 constructs (CFI and TLI values >0.90 ; RMSEA and SRMR values <0.08), indicating that the underlying factor structures of SPP, resilience, and IGD remained stable across the 3 measurement occasions. Metric invariance was further supported, as evidenced by minimal changes in model fit indices

($\Delta\text{CFI}\leq 0.01$, $\Delta\text{RMSEA}\leq 0.015$, and $\Delta\text{SRMR}\leq 0.01$), confirming the stability of factor loadings and ensuring the comparability of the relationships between each latent construct and its indicators over time. Furthermore, scalar invariance was achieved, with similarly negligible changes in fit indices across waves, demonstrating the invariance of item intercepts across time points. These findings suggest that the longitudinal variations observed in SPP, resilience, and IGD reflect genuine temporal or developmental changes rather than artifacts arising from measurement inconsistency.

Table . Longitudinal measurement invariance tests for socially prescribed perfectionism (SPP), resilience, and internet gaming disorder (IGD) across 3 waves of the study among Chinese adolescents.

Model and variables	CFI ^a	TLI ^b	RMSEA ^c	SRMR ^d	ΔCFI	ΔRMSEA	ΔSRMR
SPP							
Configural invariance	0.933	0.925	0.045	0.031	— ^e	—	—
Metric invariance	0.928	0.923	0.044	0.029	0.005	0.001	0.002
Scalar invariance	0.926	0.921	0.046	0.031	0.002	0.002	0.002
Resilience							
Configural invariance	0.954	0.945	0.025	0.029	—	—	—
Metric invariance	0.953	0.943	0.024	0.023	0.001	0.001	0.006
Scalar invariance	0.953	0.944	0.023	0.021	0.000	0.001	0.002
IGD							
Configural invariance	0.943	0.912	0.048	0.047	—	—	—
Metric invariance	0.942	0.908	0.041	0.046	0.001	0.007	0.007
Scalar invariance	0.943	0.910	0.041	0.048	0.001	0.002	0.000

^aCFI: comparative fit index.^bTLI: Tucker-Lewis index.^cRMSEA: root mean square error of approximation.^dSRMR: standardized root mean square residual.^eNot applicable.

CLPM Results

The CLPM demonstrated a good model fit to the data ($\chi^2_{38}=163.34$; RMSEA=0.054; CFI=0.945; TLI=0.932; SRMR=0.047). The final interpreted model was the originally specified model, and no post hoc model respecification was undertaken. As illustrated in Figure 2, SPP at T1 significantly and positively predicted IGD at T2 ($\beta=0.10$, SE=0.03, 95% CI 0.04 to 0.16, $P<.001$), and SPP at T2 significantly and positively predicted IGD at T3 ($\beta=0.09$, SE=0.03, 95% CI 0.03 to 0.15, $P=.004$). In contrast, IGD did not significantly predict subsequent SPP across waves. With respect to resilience, SPP at T1 significantly and negatively predicted resilience at T2 ($\beta=-0.09$, SE=0.04, 95% CI -0.16 to -0.02, $P=.007$), and SPP at T2 significantly and negatively predicted resilience at T3 ($\beta=-0.12$, SE=0.03, 95% CI -0.18 to -0.06, $P<.001$). However,

resilience did not significantly predict subsequent SPP. Resilience at T1 significantly and negatively predicted IGD at T2 ($\beta=-0.09$, SE=0.03, 95% CI -0.15 to -0.03, $P=.001$), and resilience at T2 significantly and negatively predicted IGD at T3 ($\beta=-0.09$, SE=0.03, 95% CI -0.15 to -0.03, $P=.001$). Meanwhile, IGD at T1 significantly and negatively predicted resilience at T2 ($\beta=-0.19$, SE=0.04, 95% CI -0.27 to -0.11, $P<.001$), and IGD at T2 significantly and negatively predicted resilience at T3 ($\beta=-0.09$, SE=0.03, 95% CI -0.15 to -0.03, $P=.003$).

Regarding longitudinal mediation, SPP at T1 significantly predicted resilience at T2, which in turn significantly predicted IGD at T3, indicating a significant indirect effect of SPP on IGD via resilience over time. Specifically, the indirect effect of SPP at T1 on IGD at T3 through resilience at T2 was significant ($\beta=0.008$, SE=0.002, 95% CI 0.004 to 0.012, $P=.005$).

on external approval [60,61]. For adolescents with high levels of SPP, online gaming may become especially appealing because it provides a relatively controllable environment in which achievement, feedback, and social recognition can be obtained more immediately and, at least temporarily, with less fear of direct real-world evaluation [62,63]. In this way, gaming may serve not only as an escape from pressure but also as a compensatory context for restoring threatened competence and self-worth [64].

The predictive role of SPP in increasing IGD risk also highlights the importance of supportive offline environments [65,66]. Constructive academic and familial feedback may alleviate excessive self-criticism and reduce the tendency to seek external validation through online gaming platforms [65]. This implication is especially salient in the contemporary digital era, where online gaming is highly accessible and deeply integrated into adolescents' daily routines [67]. Accordingly, interventions aimed at modifying maladaptive perfectionistic beliefs and strengthening adaptive coping strategies represent promising avenues for prevention and early intervention in problematic gaming behaviors [68,69]. Notably, this study focused primarily on a socially driven cognitive-interpersonal vulnerability (ie, SPP); future research may incorporate additional contextual and interpersonal variables, such as peer norms, academic stress, and parenting styles, to further indicate the multifaceted pathways contributing to IGD [66,70].

The Mediating Role of Resilience

In support of H2, this study identified a significant mediating role of resilience in the longitudinal relationship between SPP and IGD. More importantly, this finding helps clarify how a socially driven perfectionistic vulnerability may be translated into problematic gaming over time [10,71]. The result is broadly consistent with the compensatory internet use framework [71,72], which proposes that maladaptive online behavior is more likely when individuals experience offline stress but lack sufficient psychological resources to cope adaptively with that stress [73]. Within this framework, resilience is not merely a general positive trait; rather, it functions as a regulatory resource that may determine whether perfectionism-related distress is managed through adaptive coping or displaced into compensatory gaming behavior [31]. Adolescents with high levels of SPP are likely to experience chronic evaluative stress, fear of negative judgment, and heightened dependence on external approval [10]. When resilience is weakened under such conditions, they may become less able to tolerate frustration, recover from perceived failure, and regulate negative affect in offline contexts [74]. Online gaming may then become especially attractive because it offers immediate feedback, structured goals, a temporary distraction, and a controllable sense of competence [75]. In this sense, our findings extend the theoretical framework by suggesting that the pathway from perfectionistic vulnerability to IGD may operate partly through erosion of a key psychological coping resource rather than through distress alone [30,72].

At the same time, our findings suggest that the role of resilience is not purely unidirectional. In addition to the hypothesized pathway from lower resilience to later IGD, IGD also negatively

predicted subsequent resilience. This reverse path is theoretically meaningful because it indicates that problematic gaming may not only reflect preexisting coping difficulties but also further deplete adolescents' psychological resources over time [76,77]. Excessive gaming may displace restorative offline experiences, weaken real-world problem-solving, disrupt sleep and daily routines, and reduce opportunities for supportive social interaction, all of which may undermine the capacity to recover from stress adaptively [78,79]. Thus, the findings point to a potentially reciprocal process: Lower resilience may increase vulnerability to gaming as a compensatory strategy, whereas higher levels of IGD may, in turn, erode resilience, thereby contributing to a self-reinforcing cycle [76,77]. However, this reciprocal dynamic did not appear to extend back to SPP, as resilience did not significantly predict subsequent SPP. This asymmetry is theoretically informative. It suggests that, although resilience is malleable and closely tied to adolescents' ongoing coping capacity [80], SPP may represent a more stable socially driven cognitive-personality vulnerability rooted in internalized external expectations and evaluative concerns [10]. In other words, problematic gaming may weaken adolescents' coping resources, but such depletion may not be sufficient to alter their broader perfectionistic beliefs over the relatively short time frame of the study [10,76,81]. These findings refine the proposed model by suggesting that SPP may function primarily as an upstream vulnerability factor, whereas resilience and IGD may constitute a more dynamic downstream process characterized by partial reciprocity [70,76].

It is important to recognize that, although the magnitude of the longitudinal indirect effect was small, this pattern is consistent with prior methodological and developmental research indicating that prospective effects in multiwave psychological models are typically modest in size [82]. Longitudinal mediation analyses inherently involve residualization of prior levels and multiplication of coefficients, both of which tend to attenuate effect estimates even when the underlying theoretical process is reliable [83]. Moreover, the current 3-wave design, which incorporated temporal ordering and statistical control of prior levels of both the mediator and outcome variables, followed established recommendations for testing mediational mechanisms using structural equation modeling and enables stronger causal inference than cross-sectional approaches [84]. The relatively small indirect effect may also reflect the fact that SPP is a comparatively distal cognitive-personality vulnerability, whereas IGD is a downstream behavioral outcome likely shaped by multiple intervening mechanisms. In this context, it is theoretically plausible that only a limited portion of the association between SPP and IGD would be transmitted through any single mediator, including resilience. Importantly, developmental research has emphasized that even modest cross-lagged or indirect effects may accumulate over time and retain theoretical and practical significance, particularly in adolescent psychosocial processes characterized by substantial temporal stability [85]. Accordingly, the mediation effect found in this study should not be overstated, but neither should it be dismissed as trivial. Rather, it suggests that resilience represents one meaningful, but not exhaustive, pathway through which SPP may contribute to later IGD.

Notably, resilience is unlikely to be the sole mediating pathway linking SPP and IGD. Other psychological processes, such as emotion regulation difficulties and depressive symptoms, may also play important intermediary roles [86]. For example, individuals with high levels of SPP may experience heightened emotional distress or impaired self-regulation, which may prompt them to seek immediate feedback and perceived controllability through immersive gaming environments [2,37]. The compensatory internet use framework further suggests that personality-related vulnerabilities and psychosocial resources interact dynamically in shaping excessive gaming behaviors, underscoring the value of incorporating multiple mediating variables for a more comprehensive understanding [18]. Future research should therefore examine additional mediators, such as self-esteem, coping styles, and broader personality traits, using longitudinal designs to clarify how these factors jointly influence the developmental trajectory of problematic gaming behaviors [87].

No Significant Sex Differences in the Longitudinal Associations

Contrary to H3, the results revealed no significant sex differences in the longitudinal pathways linking SPP, resilience, and IGD. This finding suggests that the temporal associations among socially driven cognitive interpersonal vulnerability, psychological resources, and problematic gaming behaviors are largely comparable for male and female adolescents [88,89]. Importantly, although prior studies have reported sex differences in the prevalence or mean levels of IGD [88], such mean-level disparities do not necessarily indicate differences in structural or longitudinal relationships among psychological constructs. Consistent with this distinction, our results demonstrate that the cross-wave relationships among SPP, resilience, and IGD operate similarly across sex groups [90]. One plausible explanation is that the core psychological processes underlying externally imposed perfectionistic expectations and maladaptive coping, including heightened self-criticism, increased stress sensitivity, and reduced coping flexibility, may function in comparable ways for both male and female adolescents [89,91]. In addition, adolescents often encounter shared developmental challenges and environmental pressures, such as academic competition, social evaluation, and uncertainty about the future, which may contribute to relatively uniform patterns of psychological vulnerability and behavioral responses regardless of sex [92,93]. Accordingly, sex may be less influential in differentiating longitudinal psychological mechanisms than in shaping overall levels of gaming involvement or emotional symptoms [88,93]. Future research should therefore consider whether other individual or contextual factors, such as age, gaming motivation, or peer norms, explain greater variation in these longitudinal associations than sex alone.

Implications

This study has several important implications. First, it advances the theoretical understanding of socially driven cognitive interpersonal vulnerability and behavioral addiction by demonstrating the longitudinal role of SPP in predicting IGD. Although previous studies have reported associations between perfectionism and problematic internet or gaming behaviors

[23], our findings provide temporal evidence that SPP prospectively predicts IGD over time. Second, this study contributes to theoretical development in IGD research by incorporating resilience frameworks into the examination of SPP and IGD. The results indicate that adolescents with elevated SPP may experience reduced psychological resilience, which subsequently increases their susceptibility to problematic gaming behaviors. This mediating pattern supports compensatory theoretical perspectives within the behavioral addiction literature. In addition, the findings extend prior work by showing that SPP not only directly predicts IGD but also indirectly influences IGD through resilience, highlighting the importance of psychological resources as a key intermediary mechanism in the developmental pathway of problematic gaming.

From a practical perspective, these findings suggest that interventions addressing socially driven cognitive interpersonal vulnerabilities and resilience enhancement may serve as effective strategies for preventing and reducing IGD among adolescents. Because elevated SPP can weaken resilience and increase gaming risk, intervention efforts should focus on reshaping maladaptive perfectionistic beliefs, promoting flexible self-evaluation, and reducing excessive self-criticism [23]. Counseling and cognitive behavioral programs that encourage self-compassion, realistic performance standards, and adaptive responses to external expectations may alleviate the psychological pressure that fuels excessive gaming behaviors [94,95]. Furthermore, the results emphasize the value of resilience-focused initiatives in mitigating IGD risk. Schools, families, and mental health professionals may consider implementing programs that strengthen emotional regulation, coping flexibility, and problem-solving skills among adolescents [96]. Providing offline opportunities for meaningful achievement and social engagement, including group activities, mentorship programs, and counseling services, can further reinforce psychological resources and reduce reliance on online gaming as a compensatory outlet [97]. Finally, the findings highlight the importance of promoting digital self-regulation and healthy gaming education. Educational initiatives that cultivate balanced gaming habits, effective time management, and awareness of compulsive gaming tendencies may help adolescents establish a more sustainable relationship with digital entertainment and decrease the likelihood of problematic gaming involvement [98,99].

Limitations

This study has several limitations. First, all variables were assessed using self-report measures, which may introduce social desirability and recall biases. Future research would benefit from incorporating multi-informant or behavioral assessments, such as parent or teacher reports and objective usage records. Second, although a 3-wave longitudinal design was used, this study relied on a traditional CLPM, which has important limitations. In particular, CLPM may confound stable between-person differences with within-person change processes over time [100]. Therefore, the observed cross-lagged associations should not be interpreted as definitive evidence of causal or developmental effects. Future research could use more advanced longitudinal models, such as random-intercept CLPMs, as well as experimental or intervention-based designs

to provide a clearer test of within-person processes and causal relations. Third, IGD was measured using a questionnaire rather than clinical diagnostic interviews, which may limit the clinical interpretability of the findings. Replication in clinical or high-risk populations would strengthen external validity and enhance clinical relevance. Fourth, this study focused only on SPP and did not include other dimensions of perfectionism. Future longitudinal studies should examine multiple perfectionism dimensions simultaneously to compare their unique roles in the development of IGD. Finally, participants were recruited through nonrandom sampling within a limited geographic region, which may introduce sampling bias and restrict generalizability. Future research should therefore use broader and more diverse samples across cultural and socioeconomic contexts to verify the robustness of the observed associations.

Conclusions

This study extends existing research by using a 3-wave longitudinal design to examine how SPP, resilience, and IGD are associated over time in adolescents. Unlike prior studies that have relied primarily on cross-sectional data, these findings highlight resilience as one temporal pathway through which SPP may be associated with later IGD. These results contribute to the literature by refining understanding of how socially driven perfectionistic vulnerability and psychological resources jointly shape adolescent problematic gaming. In practical terms, the findings suggest that adolescents exposed to high external evaluative pressure may benefit from early identification and prevention efforts that target both maladaptive perfectionistic beliefs and resilience-related coping capacities.

Acknowledgments

We thank the participants who participated in the study.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: PZ, TY, MZ

Data curation: PZ, ZT, TY, LF, MZ

Formal analysis: PZ, TY, LF, MZ

Investigation: PZ, ZT, TY, MZ

Methodology: PZ, ZT, TY, LF, MZ

Visualization: PZ, ZT, TY, MZ

Writing-original draft: PZ, ZT, TY, MZ

Conflicts of Interest

None declared.

Multimedia Appendix 1

Cross-lagged panel model of socially prescribed perfectionism (SPP), resilience, and internet gaming disorder (IGD) across the 3 waves, with missing data handled using multiple imputation.

[[DOCX File, 78 KB - games_v14i1e93412_app1.docx](#)]

References

1. Griffiths MD. Online gaming addiction in youth: some comments on Rosendo-Rios et al. (2022). *Addict Behav* 2022 Jul;130:107311. [doi: [10.1016/j.addbeh.2022.107311](https://doi.org/10.1016/j.addbeh.2022.107311)] [Medline: [35339731](https://pubmed.ncbi.nlm.nih.gov/35339731/)]
2. Paulus FW, Ohmann S, von Gontard A, Popow C. Internet gaming disorder in children and adolescents: a systematic review. *Develop Med Child Neuro* 2018 Jul;60(7):645-659 [FREE Full text] [doi: [10.1111/dmcn.13754](https://doi.org/10.1111/dmcn.13754)]
3. American Psychiatric Association, DSM-5 Task Force. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*: American Psychiatric Publishing, Inc; 2013. URL: [10.1176/appi.books.9780890425596](https://doi.org/10.1176/appi.books.9780890425596)
4. Gou S, Zhang W, Tang Y, Zhang J, He Q. Prevalence of internet gaming disorder among Chinese adolescents: a systematic review and meta-analysis. *Asian J Psychiatr* 2024 Dec;102:104257. [doi: [10.1016/j.ajp.2024.104257](https://doi.org/10.1016/j.ajp.2024.104257)] [Medline: [39366164](https://pubmed.ncbi.nlm.nih.gov/39366164/)]
5. Teng Z, Pontes HM, Nie Q, Griffiths MD, Guo C. Depression and anxiety symptoms associated with internet gaming disorder before and during the COVID-19 pandemic: a longitudinal study. *J Behav Addict* 2021 Mar 10;10(1):169-180. [doi: [10.1556/2006.2021.00016](https://doi.org/10.1556/2006.2021.00016)] [Medline: [33704085](https://pubmed.ncbi.nlm.nih.gov/33704085/)]

6. Kristensen JH, Pallesen S, King DL, Hysing M, Erevik EK. Problematic gaming and sleep: a systematic review and meta-analysis. *Front Psychiatry* 2021;12:675237. [doi: [10.3389/fpsy.2021.675237](https://doi.org/10.3389/fpsy.2021.675237)] [Medline: [34163386](https://pubmed.ncbi.nlm.nih.gov/34163386/)]
7. Smith MM, Sherry SB, Ge SYJ, Hewitt PL, Flett GL, Baggley DL. Multidimensional perfectionism turns 30: a review of known knowns and known unknowns. *Canadian Psychology / Psychologie canadienne* 2022;63(1):16-31. [doi: [10.1037/cap0000288](https://doi.org/10.1037/cap0000288)]
8. Hill AP, Gotwals JK. A meta-analysis of multidimensional perfectionism and impostor phenomenon. *J Res Pers* 2025 Oct;118:104639. [doi: [10.1016/j.jrp.2025.104639](https://doi.org/10.1016/j.jrp.2025.104639)]
9. Curran T, Hill AP. Young people's perceptions of their parents' expectations and criticism are increasing over time: implications for perfectionism. *Psychol Bull* 2022;148(1-2):107-128. [doi: [10.1037/bul0000347](https://doi.org/10.1037/bul0000347)] [Medline: [35357848](https://pubmed.ncbi.nlm.nih.gov/35357848/)]
10. Flett GL, Hewitt PL, Nepon T, Sherry SB, Smith M. The destructiveness and public health significance of socially prescribed perfectionism: a review, analysis, and conceptual extension. *Clin Psychol Rev* 2022 Apr;93:102130. [doi: [10.1016/j.cpr.2022.102130](https://doi.org/10.1016/j.cpr.2022.102130)] [Medline: [35216826](https://pubmed.ncbi.nlm.nih.gov/35216826/)]
11. Nepon T, Flett GL, Hewitt PL. Perfectionism, self-image goals and compassionate goals in health and mental health: a longitudinal analysis. *J Psychoeduc Assess* 2024 Sep;42(6):650-667. [doi: [10.1177/07342829241239997](https://doi.org/10.1177/07342829241239997)] [Medline: [39157723](https://pubmed.ncbi.nlm.nih.gov/39157723/)]
12. Chen S, Saklofske DH, Flett GL, Hewitt PL. Assessing and evaluating the perfectionism social disconnection model: social support, loneliness, and distress among undergraduate, law, and, medical students. *J Psychoeduc Assess* 2024 Sep;42(6):743-767. [doi: [10.1177/07342829241244951](https://doi.org/10.1177/07342829241244951)] [Medline: [39157721](https://pubmed.ncbi.nlm.nih.gov/39157721/)]
13. Mohammed SKF, Maranzan KA, Mushquash AR. Mattering matters for students high in socially prescribed perfectionism as they transition to university: support for a moderation model. *Pers Individ Dif* 2023 Nov;214:112356. [doi: [10.1016/j.paid.2023.112356](https://doi.org/10.1016/j.paid.2023.112356)]
14. Mushquash AR, Sherry SB. Understanding the socially prescribed perfectionist's cycle of self-defeat: a 7-day, 14-occasion daily diary study. *J Res Pers* 2012 Dec;46(6):700-709. [doi: [10.1016/j.jrp.2012.08.006](https://doi.org/10.1016/j.jrp.2012.08.006)]
15. Curran T, Hill AP. Perfectionism is increasing over time: a meta-analysis of birth cohort differences from 1989 to 2016. *Psychol Bull* 2019 Apr;145(4):410-429. [doi: [10.1037/bul0000138](https://doi.org/10.1037/bul0000138)] [Medline: [29283599](https://pubmed.ncbi.nlm.nih.gov/29283599/)]
16. Abdollahi A, Hosseinian S, Beh-Pajoo A, Carlbring P. Self-concealment mediates the relationship between perfectionism and attitudes toward seeking psychological help among adolescents. *Psychol Rep* 2017 Dec;120(6):1019-1036. [doi: [10.1177/0033294117713495](https://doi.org/10.1177/0033294117713495)] [Medline: [28592165](https://pubmed.ncbi.nlm.nih.gov/28592165/)]
17. Gioia F, Rega V, Boursier V. Problematic internet use and emotional dysregulation among young people: a literature review. *Clin Neuropsychiatry* 2021 Feb;18(1):41-54. [doi: [10.36131/cnfioritieditore20210104](https://doi.org/10.36131/cnfioritieditore20210104)] [Medline: [34909019](https://pubmed.ncbi.nlm.nih.gov/34909019/)]
18. Kardefelt-Winther D. A conceptual and methodological critique of internet addiction research: towards a model of compensatory internet use. *Comput Human Behav* 2014 Feb;31:351-354. [doi: [10.1016/j.chb.2013.10.059](https://doi.org/10.1016/j.chb.2013.10.059)]
19. Luján-Barrera A, Cervera-Ortiz L, Chóliz M. In-game need satisfaction, frustration, and gaming addiction patterns across subgroups of adolescents through structural equation modeling: cross-sectional and instrument validation study of the youth gaming experience Scales. *JMIR Serious Games* 2025 Nov 12;13:e63612. [doi: [10.2196/63612](https://doi.org/10.2196/63612)] [Medline: [41223412](https://pubmed.ncbi.nlm.nih.gov/41223412/)]
20. Sinha S, Sharma MK, Narayanan G, Kashyap H, Shreedevi AU, Tadpatrikar A. A systematic review of problematic gaming: escapist strategy to cope with difficult life situations. *Ind Psychiatry J* 2024;33(2):210-218. [doi: [10.4103/ipj.ipj_99_23](https://doi.org/10.4103/ipj.ipj_99_23)] [Medline: [39898101](https://pubmed.ncbi.nlm.nih.gov/39898101/)]
21. Yang W, Morita N, Zuo Z, et al. Maladaptive perfectionism and internet addiction among Chinese college students: a moderated mediation model of depression and gender. *Int J Environ Res Public Health* 2021 Mar 9;18(5):2748. [doi: [10.3390/ijerph18052748](https://doi.org/10.3390/ijerph18052748)] [Medline: [33803092](https://pubmed.ncbi.nlm.nih.gov/33803092/)]
22. Akbari M, Mohammadaliha N, Mohammadkhani S, Seydavi M, Griffiths MD. Cognitive, metacognitive, motivational, and emotional predictors of the intensity of internet gaming disorder among adolescents. *Psychiatr Q* 2024 Sep;95(3):385-414. [doi: [10.1007/s11126-024-10075-w](https://doi.org/10.1007/s11126-024-10075-w)] [Medline: [38954306](https://pubmed.ncbi.nlm.nih.gov/38954306/)]
23. Casale S, Fioravanti G, Flett GL, Hewitt PL. From socially prescribed perfectionism to problematic use of internet communicative services: the mediating roles of perceived social support and the fear of negative evaluation. *Addict Behav* 2014 Dec;39(12):1816-1822. [doi: [10.1016/j.addbeh.2014.06.006](https://doi.org/10.1016/j.addbeh.2014.06.006)] [Medline: [25127198](https://pubmed.ncbi.nlm.nih.gov/25127198/)]
24. Forrest CJ, King DL, Delfabbro PH. Maladaptive cognitions predict changes in problematic gaming in highly-engaged adults: a 12-month longitudinal study. *Addict Behav* 2017 Feb;65:125-130. [doi: [10.1016/j.addbeh.2016.10.013](https://doi.org/10.1016/j.addbeh.2016.10.013)] [Medline: [27816037](https://pubmed.ncbi.nlm.nih.gov/27816037/)]
25. Lu H, Yang J, Zhao K, et al. Perceived risk of COVID-19 hurts mental health: the mediating role of fear of COVID-19 and the moderating role of resilience. *BMC Psychiatry* 2024 Jan 22;24(1):58. [doi: [10.1186/s12888-024-05511-x](https://doi.org/10.1186/s12888-024-05511-x)] [Medline: [38254008](https://pubmed.ncbi.nlm.nih.gov/38254008/)]
26. Lu H, Yu Y, Wang DB, et al. Association between interpersonal resources and mental health professional help-seeking among Chinese adolescents with probable depression: mediations via personal resources and active coping. *BMC Psychiatry* 2024 Nov 21;24(1):840. [doi: [10.1186/s12888-024-06271-4](https://doi.org/10.1186/s12888-024-06271-4)] [Medline: [39574049](https://pubmed.ncbi.nlm.nih.gov/39574049/)]
27. Bagdžiūnienė D, Žukauskaitė I, Bulotaitė L, Sargautytė R. Study and personal resources of university students' academic resilience and the relationship with positive psychological outcomes. *Front Psychol* 2025;16:1517359. [doi: [10.3389/fpsyg.2025.1517359](https://doi.org/10.3389/fpsyg.2025.1517359)] [Medline: [39973962](https://pubmed.ncbi.nlm.nih.gov/39973962/)]

28. Lu H, Ren Z, Zhen P, Su Z, Wu Y. The longitudinal association between family functioning and problematic social media use among Chinese university students: mediation via loneliness and a subgroup analysis by sex. *Addict Behav* 2025 Jul;166:108337. [doi: [10.1016/j.addbeh.2025.108337](https://doi.org/10.1016/j.addbeh.2025.108337)] [Medline: [40121924](https://pubmed.ncbi.nlm.nih.gov/40121924/)]
29. Herrman H, Stewart DE, Diaz-Granados N, Berger EL, Jackson B, Yuen T. What is resilience? *Can J Psychiatry* 2011 May;56(5):258-265. [doi: [10.1177/070674371105600504](https://doi.org/10.1177/070674371105600504)] [Medline: [21586191](https://pubmed.ncbi.nlm.nih.gov/21586191/)]
30. Yu Z, Liu W. The psychological resilience of teenagers in terms of their everyday emotional balance and the impact of emotion regulation strategies. *Front Psychol* 2024;15:1381239. [doi: [10.3389/fpsyg.2024.1381239](https://doi.org/10.3389/fpsyg.2024.1381239)] [Medline: [40052017](https://pubmed.ncbi.nlm.nih.gov/40052017/)]
31. Attasara P, Thaweekoon T, Napa W. Perceived stress, loneliness, and resilience in relation to game addiction among adolescents in Bangkok during the COVID-19 pandemic transition period. *J Child Adolesc Psychiatr Nurs* 2025 Nov;38(4):e70044. [doi: [10.1111/jcap.70044](https://doi.org/10.1111/jcap.70044)] [Medline: [41200867](https://pubmed.ncbi.nlm.nih.gov/41200867/)]
32. Jiang Y, Konorova E. Distinct roles of self-oriented and socially prescribed perfectionism in Chinese adolescent students' achievement goals, classroom affect, and academic achievement. *Learn Individ Differ* 2023 Aug;106:102341. [doi: [10.1016/j.lindif.2023.102341](https://doi.org/10.1016/j.lindif.2023.102341)]
33. Arango C, Dragioti E, Solmi M, et al. Risk and protective factors for mental disorders beyond genetics: an evidence-based atlas. *World Psychiatry* 2021 Oct;20(3):417-436 [FREE Full text] [doi: [10.1002/wps.20894](https://doi.org/10.1002/wps.20894)] [Medline: [34505386](https://pubmed.ncbi.nlm.nih.gov/34505386/)]
34. Abdulcebbar A, Altan-Atalay A. The relationship between perfectionism and stress generation: the moderating role of looming cognitive style. *Curr Psychol* 2025 May;44(10):8282-8290. [doi: [10.1007/s12144-025-07791-2](https://doi.org/10.1007/s12144-025-07791-2)]
35. Wiedermann CJ, Barbieri V, Piccoliori G, Engl A. Problematic internet use in adolescents is driven by internal distress rather than family or socioeconomic contexts: evidence from South Tyrol, Italy. *Behav Sci (Basel)* 2025 Nov 11;15(11):1534. [doi: [10.3390/bs15111534](https://doi.org/10.3390/bs15111534)] [Medline: [41301337](https://pubmed.ncbi.nlm.nih.gov/41301337/)]
36. Ma J, Yang B, Wang S, et al. Adverse childhood experiences predict internet gaming disorder in university students: the mediating role of resilience. *Curr Opin Psychiatry* 2024 Jan 1;37(1):29-37. [doi: [10.1097/YCO.0000000000000910](https://doi.org/10.1097/YCO.0000000000000910)] [Medline: [37972967](https://pubmed.ncbi.nlm.nih.gov/37972967/)]
37. Yen JY, Lin HC, Chou WP, Liu TL, Ko CH. Associations among resilience, stress, depression, and internet gaming disorder in young adults. *Int J Environ Res Public Health* 2019 Aug 31;16(17):3181. [doi: [10.3390/ijerph16173181](https://doi.org/10.3390/ijerph16173181)] [Medline: [31480445](https://pubmed.ncbi.nlm.nih.gov/31480445/)]
38. Stevens MW, Dorstyn D, Delfabbro PH, King DL. Global prevalence of gaming disorder: a systematic review and meta-analysis. *Aust N Z J Psychiatry* 2021 Jun;55(6):553-568. [doi: [10.1177/0004867420962851](https://doi.org/10.1177/0004867420962851)] [Medline: [33028074](https://pubmed.ncbi.nlm.nih.gov/33028074/)]
39. Kuss DJ, Griffiths MD. Internet gaming addiction: a systematic review of empirical research. *Int J Ment Health Addiction* 2012 Apr;10(2):278-296. [doi: [10.1007/s11469-011-9318-5](https://doi.org/10.1007/s11469-011-9318-5)]
40. Anthony WL, Zhu Y, Nower L. The relationship of interactive technology use for entertainment and school performance and engagement: evidence from a longitudinal study in a nationally representative sample of middle school students in China. *Comput Human Behav* 2021 Sep;122:106846. [doi: [10.1016/j.chb.2021.106846](https://doi.org/10.1016/j.chb.2021.106846)]
41. Nam CR, Lee DH, Lee JY, et al. The role of resilience in internet addiction among adolescents between sexes: a moderated mediation model. *J Clin Med* 2018 Aug 19;7(8):222. [doi: [10.3390/jcm7080222](https://doi.org/10.3390/jcm7080222)] [Medline: [30126239](https://pubmed.ncbi.nlm.nih.gov/30126239/)]
42. O'Connor RC, Rasmussen S, Hawton K. Predicting depression, anxiety and self-harm in adolescents: the role of perfectionism and acute life stress. *Behav Res Ther* 2010 Jan;48(1):52-59. [doi: [10.1016/j.brat.2009.09.008](https://doi.org/10.1016/j.brat.2009.09.008)] [Medline: [19818954](https://pubmed.ncbi.nlm.nih.gov/19818954/)]
43. Dou K, Feng XK, Wang LX, Li JB. Longitudinal association between parental involvement and internet gaming disorder among Chinese adolescents: consideration of future consequences as a mediator and peer victimization as a moderator. *J Behav Addict* 2022 Sep 26;11(3):820-830. [doi: [10.1556/2006.2022.00056](https://doi.org/10.1556/2006.2022.00056)] [Medline: [35994364](https://pubmed.ncbi.nlm.nih.gov/35994364/)]
44. Huang Y, Wu R, Huang Y, Xiang Y, Zhou W. Investigating the mechanisms of internet gaming disorder and developing intelligent monitoring models using artificial intelligence technologies: protocol of a prospective cohort. *BMC Public Health* 2024;24(1):39294602. [doi: [10.1186/s12889-024-20028-4](https://doi.org/10.1186/s12889-024-20028-4)]
45. Guha M. Diagnostic and Statistical Manual of Mental Disorders: DSM-5 (5th edition). *Reference Reviews* 2014 Mar 11;28(3):36-37. [doi: [10.1108/RR-10-2013-0256](https://doi.org/10.1108/RR-10-2013-0256)]
46. Yang X, Jiang X, Wu AMS, et al. Validation of the Internet Gaming Disorder Symptoms Checklist based on the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders in Chinese adolescents. *Child Psychiatry Hum Dev* 2023 Feb;54(1):26-33. [doi: [10.1007/s10578-021-01213-7](https://doi.org/10.1007/s10578-021-01213-7)] [Medline: [34338962](https://pubmed.ncbi.nlm.nih.gov/34338962/)]
47. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the Connor-Davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. *J Trauma Stress* 2007 Dec;20(6):1019-1028 [FREE Full text] [doi: [10.1002/jts.20271](https://doi.org/10.1002/jts.20271)] [Medline: [18157881](https://pubmed.ncbi.nlm.nih.gov/18157881/)]
48. Wang L, Shi Z, Zhang Y, Zhang Z. Psychometric properties of the 10-item Connor-Davidson Resilience Scale in Chinese earthquake victims. *Psychiatry Clin Neurosci* 2010 Oct;64(5):499-504. [doi: [10.1111/j.1440-1819.2010.02130.x](https://doi.org/10.1111/j.1440-1819.2010.02130.x)] [Medline: [20923429](https://pubmed.ncbi.nlm.nih.gov/20923429/)]
49. Ye ZJ, Qiu HZ, Li PF, et al. Validation and application of the Chinese version of the 10-item Connor-Davidson Resilience Scale (CD-RISC-10) among parents of children with cancer diagnosis. *Eur J Oncol Nurs* 2017 Apr;27:36-44. [doi: [10.1016/j.ejon.2017.01.004](https://doi.org/10.1016/j.ejon.2017.01.004)] [Medline: [28279394](https://pubmed.ncbi.nlm.nih.gov/28279394/)]
50. Hewitt PL, Flett GL, Mikail SF. *Perfectionism: A Relational Approach to Conceptualization, Assessment, and Treatment*: Guilford Publications; 2017.

51. Zhao W, Wu AMS, Feng C, Yu K, Wang Z, Jiao K. Perfectionism and suicidal ideation: the serial mediating roles of appearance-based rejection sensitivity and loneliness. *Curr Psychol* 2024 Aug;43(31):25494-25503. [doi: [10.1007/s12144-024-06230-y](https://doi.org/10.1007/s12144-024-06230-y)]
52. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling* 2002 Apr;9(2):233-255. [doi: [10.1207/S15328007SEM0902_5](https://doi.org/10.1207/S15328007SEM0902_5)]
53. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annu Rev Psychol* 2007;58(1):593-614. [doi: [10.1146/annurev.psych.58.110405.085542](https://doi.org/10.1146/annurev.psych.58.110405.085542)] [Medline: [16968208](https://pubmed.ncbi.nlm.nih.gov/16968208/)]
54. McDonald RP, Ho MHR. Principles and practice in reporting structural equation analyses. *Psychol Methods* 2002 Mar;7(1):64-82. [doi: [10.1037/1082-989x.7.1.64](https://doi.org/10.1037/1082-989x.7.1.64)] [Medline: [11928891](https://pubmed.ncbi.nlm.nih.gov/11928891/)]
55. Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling: A Multidisciplinary J* 2001 Jul 1;8(3):430-457. [doi: [10.1207/S15328007SEM0803_5](https://doi.org/10.1207/S15328007SEM0803_5)]
56. Nicholson JS, Deboeck PR, Howard W. Attrition in developmental psychology: A review of modern missing data reporting and practices. *Int J Behav Dev* 2017;41(1):143-153. [doi: [10.1177/0165025415618275](https://doi.org/10.1177/0165025415618275)]
57. Zhou J, Li T, Li S, Yang P, Gong X. Bidirectional relations between belief in a just world and internalizing and externalizing problems among Chinese secondary vocational students. *J Youth Adolesc* 2025 Oct;54(10):2644-2659. [doi: [10.1007/s10964-025-02259-9](https://doi.org/10.1007/s10964-025-02259-9)] [Medline: [40960704](https://pubmed.ncbi.nlm.nih.gov/40960704/)]
58. Ren Z, Lu H, Ma X, et al. Longitudinal association between problematic social media use and sleep problems among adolescents: a random intercept cross-lagged panel model and sex differences. *J Youth Adolesc* 2026 Jan;55(1):119-134. [doi: [10.1007/s10964-025-02310-9](https://doi.org/10.1007/s10964-025-02310-9)] [Medline: [41452419](https://pubmed.ncbi.nlm.nih.gov/41452419/)]
59. Long J, Liu TQ, Liao YH, et al. Prevalence and correlates of problematic smartphone use in a large random sample of Chinese undergraduates. *BMC Psychiatry* 2016 Nov 17;16(1):408. [doi: [10.1186/s12888-016-1083-3](https://doi.org/10.1186/s12888-016-1083-3)] [Medline: [27855666](https://pubmed.ncbi.nlm.nih.gov/27855666/)]
60. Hill AP, Hall HK, Appleton PR. The relationship between multidimensional perfectionism and contingencies of self-worth. *Pers Individ Dif* 2011 Jan;50(2):238-242. [doi: [10.1016/j.paid.2010.09.036](https://doi.org/10.1016/j.paid.2010.09.036)]
61. Luo Z, Li H, Xiong Y. The double-edged perfectionist parenting: an examination of its impact on adolescent internet gaming addiction (IGA). *BMC Psychol* 2025 Sep 26;13(1):1033. [doi: [10.1186/s40359-025-03366-9](https://doi.org/10.1186/s40359-025-03366-9)] [Medline: [41013662](https://pubmed.ncbi.nlm.nih.gov/41013662/)]
62. Casale S, Fioravanti G, Flett GL, Hewitt PL. Self-presentation styles and problematic use of internet communicative services: the role of the concerns over behavioral displays of imperfection. *Pers Individ Dif* 2015 Apr;76:187-192. [doi: [10.1016/j.paid.2014.12.021](https://doi.org/10.1016/j.paid.2014.12.021)]
63. Wang HY, Cheng C. The associations between gaming motivation and internet gaming disorder: systematic review and meta-analysis. *JMIR Ment Health* 2022 Feb 17;9(2):e23700. [doi: [10.2196/23700](https://doi.org/10.2196/23700)] [Medline: [35175204](https://pubmed.ncbi.nlm.nih.gov/35175204/)]
64. Kavanagh M, Brett C, Brignell C. The role of self-esteem in triggering and maintaining gaming for people with gaming disorder cut-off scores and above average gaming disorder scores: a quasi-experimental study. *Comput Human Behav* 2024 Mar;152:108051. [doi: [10.1016/j.chb.2023.108051](https://doi.org/10.1016/j.chb.2023.108051)]
65. Meng C, Quancai L, Kunjie C, Yanyu X, Wencai L, Yiwei X. Low self-control, perceived social support and internet gaming addiction: findings from an ethnicity minority region in China. *Front Psychiatry* 2024;15:1458626. [doi: [10.3389/fpsy.2024.1458626](https://doi.org/10.3389/fpsy.2024.1458626)] [Medline: [39310659](https://pubmed.ncbi.nlm.nih.gov/39310659/)]
66. Schneider LA, King DL, Delfabbro PH. Family factors in adolescent problematic Internet gaming: a systematic review. *J Behav Addict* 2017 Sep 1;6(3):321-333. [doi: [10.1556/2006.6.2017.035](https://doi.org/10.1556/2006.6.2017.035)] [Medline: [28762279](https://pubmed.ncbi.nlm.nih.gov/28762279/)]
67. Park JJ, Stryjewski A, Chen B, Potenza MN. Treatment of gaming disorder in children and adolescents: a systematic review. *J Korean Acad Child Adolesc Psychiatry* 2025 Jul 1;36(3):106-121. [doi: [10.5765/jkacap.250014](https://doi.org/10.5765/jkacap.250014)] [Medline: [40631626](https://pubmed.ncbi.nlm.nih.gov/40631626/)]
68. Handley AK, Egan SJ, Kane RT, Rees CS. A randomised controlled trial of group cognitive behavioural therapy for perfectionism. *Behav Res Ther* 2015 May;68:37-47. [doi: [10.1016/j.brat.2015.02.006](https://doi.org/10.1016/j.brat.2015.02.006)] [Medline: [25795927](https://pubmed.ncbi.nlm.nih.gov/25795927/)]
69. Lindenberg K, Kindt S, Szász-Janocha C. Effectiveness of cognitive behavioral therapy-based intervention in preventing gaming disorder and unspecified internet use disorder in adolescents: a cluster randomized clinical trial. *JAMA Netw Open* 2022 Feb 1;5(2):e2148995. [doi: [10.1001/jamanetworkopen.2021.48995](https://doi.org/10.1001/jamanetworkopen.2021.48995)] [Medline: [35179587](https://pubmed.ncbi.nlm.nih.gov/35179587/)]
70. Brand M, Wegmann E, Stark R, et al. The Interaction of Person-Affect-Cognition-Execution (I-PACE) model for addictive behaviors: update, generalization to addictive behaviors beyond internet-use disorders, and specification of the process character of addictive behaviors. *Neuroscience & Biobehavioral Reviews* 2019 Sep;104:1-10. [doi: [10.1016/j.neubiorev.2019.06.032](https://doi.org/10.1016/j.neubiorev.2019.06.032)]
71. Giardina A, Schimmenti A, Starcevic V, King DL, Di Blasi M, Billieux J. Problematic gaming, social withdrawal, and escapism: the Compensatory-Dissociative Online Gaming (C-DOG) model. *Comput Human Behav* 2024 Jun;155:108187. [doi: [10.1016/j.chb.2024.108187](https://doi.org/10.1016/j.chb.2024.108187)]
72. Verrastro V, Saladino V, Giordano F, Calaresi D. Longitudinal pathways from emotional abuse to problematic gaming in adolescents: the role of psychoticism. *Compr Psychiatry* 2025 Feb;137:152569. [doi: [10.1016/j.comppsy.2024.152569](https://doi.org/10.1016/j.comppsy.2024.152569)] [Medline: [39733705](https://pubmed.ncbi.nlm.nih.gov/39733705/)]
73. Çelikkaleli Ö, Ata R. The effect of basic need dissatisfaction and lack of social support on problematic Internet use in emerging adults: the mediating role of regulatory emotional self-efficacy. *Acta Psychol (Amst)* 2025 Apr;254:104857. [doi: [10.1016/j.actpsy.2025.104857](https://doi.org/10.1016/j.actpsy.2025.104857)] [Medline: [40054082](https://pubmed.ncbi.nlm.nih.gov/40054082/)]

74. Yu Z, Lai X, Wang Y, Wang Y, Zhao H. Implicit benefits of adolescents with high psychological resilience in action control of emotion regulation. *PLoS ONE* 2025;20(9):e0332384. [doi: [10.1371/journal.pone.0332384](https://doi.org/10.1371/journal.pone.0332384)]
75. Ballou N, Deterding S. The basic needs in games model of video game play and mental health. *Interact Comput* 2026 Apr 13;38(3):469-486. [doi: [10.1093/iwc/iwae042](https://doi.org/10.1093/iwc/iwae042)]
76. Amendola S, Bernath J, Presaghi F, Waller G, Hengartner MP. Bidirectional relationship between gaming disorder, internalizing psychopathology, psychological distress, and well-being: a systematic review with meta-analysis of longitudinal studies. *J Affect Disord* 2025 Aug 15;383:480-493. [doi: [10.1016/j.jad.2025.04.108](https://doi.org/10.1016/j.jad.2025.04.108)] [Medline: [40288452](https://pubmed.ncbi.nlm.nih.gov/40288452/)]
77. Yu Y, Wu AMS, Fong VWI, Zhang J, Li JB, Lau JTF. Association between internet gaming disorder and suicidal ideation mediated by psychosocial resources and psychosocial problems among adolescent internet gamers in China: cross-sectional study. *JMIR Serious Games* 2024 Sep 19;12:e48439. [doi: [10.2196/48439](https://doi.org/10.2196/48439)] [Medline: [39298753](https://pubmed.ncbi.nlm.nih.gov/39298753/)]
78. Nagata JM, Shim JE, Balasubramanian P, et al. Prospective associations between early adolescent problematic screen use, mental health, sleep, and substance use. *Am J Prev Med* 2026 Apr;70(4):108248. [doi: [10.1016/j.amepre.2025.108248](https://doi.org/10.1016/j.amepre.2025.108248)] [Medline: [41677479](https://pubmed.ncbi.nlm.nih.gov/41677479/)]
79. Canale N, Elgar FJ, Pivetta E, et al. Problem gaming and adolescents' health and well-being: evidence from a large nationally representative sample in Italy. *Comput Human Behav* 2025 Jul;168:108644. [doi: [10.1016/j.chb.2025.108644](https://doi.org/10.1016/j.chb.2025.108644)]
80. Fu C, Chien WT, Zhang Y, Lam KK. Strength-based capacity-building interventions to promote adolescents' mental health: a systematic review and meta-analysis. *Eur Child Adolesc Psychiatry* 2025 Oct;34(10):2975-3015. [doi: [10.1007/s00787-025-02741-6](https://doi.org/10.1007/s00787-025-02741-6)] [Medline: [40437197](https://pubmed.ncbi.nlm.nih.gov/40437197/)]
81. Damian LE, Negru-Subtirica O, Stoeber J, Vois D. On the longitudinal interplay between perfectionism and general affect in adolescents. *J Res Pers* 2021 Aug;93:104102. [doi: [10.1016/j.jrp.2021.104102](https://doi.org/10.1016/j.jrp.2021.104102)]
82. Adachi P, Willoughby T. Interpreting effect sizes when controlling for stability effects in longitudinal autoregressive models: implications for psychological science. *European Journal of Developmental Psychology* 2015 Jan 2;12(1):116-128. [doi: [10.1080/17405629.2014.963549](https://doi.org/10.1080/17405629.2014.963549)]
83. Jose PE. The merits of using longitudinal mediation. *Educ Psychol* 2016 Oct;51(3-4):331-341. [doi: [10.1080/00461520.2016.1207175](https://doi.org/10.1080/00461520.2016.1207175)]
84. Cole DA, Maxwell SE. Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *J Abnorm Psychol* 2003 Nov;112(4):558-577. [doi: [10.1037/0021-843X.112.4.558](https://doi.org/10.1037/0021-843X.112.4.558)] [Medline: [14674869](https://pubmed.ncbi.nlm.nih.gov/14674869/)]
85. Orth U, Robins RW, Meier LL, Conger RD. Refining the vulnerability model of low self-esteem and depression: disentangling the effects of genuine self-esteem and narcissism. *J Pers Soc Psychol* 2016 Jan;110(1):133-149. [doi: [10.1037/pspp0000038](https://doi.org/10.1037/pspp0000038)] [Medline: [25915133](https://pubmed.ncbi.nlm.nih.gov/25915133/)]
86. Estupiñá FJ, Bernaldo-de-Quirós M, Vallejo-Achón M, Fernández-Arias I, Labrador F. Emotional regulation in gaming disorder: a systematic review. *Am J Addict* 2024 Nov;33(6):605-620. [doi: [10.1111/ajad.13621](https://doi.org/10.1111/ajad.13621)] [Medline: [38878004](https://pubmed.ncbi.nlm.nih.gov/38878004/)]
87. Ropovik I, Martončík M, Babinčák P, Baník G, Vargová L, Adamkovič M. Risk and protective factors for (internet) gaming disorder: a meta-analysis of pre-COVID studies. *Addict Behav* 2023 Apr;139:107590. [doi: [10.1016/j.addbeh.2022.107590](https://doi.org/10.1016/j.addbeh.2022.107590)] [Medline: [36571943](https://pubmed.ncbi.nlm.nih.gov/36571943/)]
88. Tu S, Liu J, Zhao H, Turel O, He Q. Sex-based differences in the impacts of impulsivity and self-control on internet gaming disorder tendencies. *Computers in Human Behavior Reports* 2025 Dec;20:100821. [doi: [10.1016/j.chbr.2025.100821](https://doi.org/10.1016/j.chbr.2025.100821)]
89. Suh H, Liou PY, Jeong J, Kim SY. Perfectionism, prolonged stress reactivity, and depression: a two-wave cross-lagged analysis. *J Rat-Emo Cognitive-Behav Ther* 2024 Mar;42(1):54-68. [doi: [10.1007/s10942-022-00483-x](https://doi.org/10.1007/s10942-022-00483-x)]
90. Tsui YYY, Cheng C. Internet gaming disorder, risky online behaviour, and mental health in Hong Kong adolescents: the beneficial role of psychological resilience. *Front Psychiatry* 2021;12:722353. [doi: [10.3389/fpsy.2021.722353](https://doi.org/10.3389/fpsy.2021.722353)] [Medline: [34721101](https://pubmed.ncbi.nlm.nih.gov/34721101/)]
91. Stoeber J, Otto K. Positive conceptions of perfectionism: approaches, evidence, challenges. *Pers Soc Psychol Rev* 2006;10(4):295-319. [doi: [10.1207/s15327957pspr1004_2](https://doi.org/10.1207/s15327957pspr1004_2)] [Medline: [17201590](https://pubmed.ncbi.nlm.nih.gov/17201590/)]
92. Steare T, Gutiérrez Muñoz C, Sullivan A, Lewis G. The association between academic pressure and adolescent mental health problems: a systematic review. *J Affect Disord* 2023 Oct 15;339:302-317. [doi: [10.1016/j.jad.2023.07.028](https://doi.org/10.1016/j.jad.2023.07.028)] [Medline: [37437728](https://pubmed.ncbi.nlm.nih.gov/37437728/)]
93. Schweizer S, Lawson RP, Blakemore SJ. Uncertainty as a driver of the youth mental health crisis. *Curr Opin Psychol* 2023 Oct;53:101657. [doi: [10.1016/j.copsy.2023.101657](https://doi.org/10.1016/j.copsy.2023.101657)] [Medline: [37517166](https://pubmed.ncbi.nlm.nih.gov/37517166/)]
94. Woodfin V, Molde H, Dundas I, Binder PE. A randomized control trial of a brief self-compassion intervention for perfectionism, anxiety, depression, and body image. *Front Psychol* 2021;12:751294. [doi: [10.3389/fpsyg.2021.751294](https://doi.org/10.3389/fpsyg.2021.751294)] [Medline: [34955971](https://pubmed.ncbi.nlm.nih.gov/34955971/)]
95. Stevens MWR, King DL, Dorstyn D, Delfabbro PH. Cognitive-behavioral therapy for internet gaming disorder: a systematic review and meta-analysis. *Clin Psychology and Psychoth* 2019 Mar;26(2):191-203 [FREE Full text] [doi: [10.1002/cpp.2341](https://doi.org/10.1002/cpp.2341)]
96. Durlak JA, Weissberg RP, Dymnicki AB, Taylor RD, Schellinger KB. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Dev* 2011;82(1):405-432. [doi: [10.1111/j.1467-8624.2010.01564.x](https://doi.org/10.1111/j.1467-8624.2010.01564.x)] [Medline: [21291449](https://pubmed.ncbi.nlm.nih.gov/21291449/)]

97. Ling Y, You Y, You S, et al. The relation between social support and internet addiction among Chinese adolescents: a moderated mediation model. *Curr Psychol* 2023 Aug;42(22):19355-19363. [doi: [10.1007/s12144-022-03159-y](https://doi.org/10.1007/s12144-022-03159-y)] [Medline: [37359659](https://pubmed.ncbi.nlm.nih.gov/37359659/)]
98. Colder Carras M, Aljuboori D, Shi J, et al. Prevention and health promotion interventions for young people in the context of digital well-being: rapid systematic review. *J Med Internet Res* 2024 Dec 18;26:e59968. [doi: [10.2196/59968](https://doi.org/10.2196/59968)] [Medline: [39693138](https://pubmed.ncbi.nlm.nih.gov/39693138/)]
99. Throuvala MA, Griffiths MD, Rennoldson M, Kuss DJ. School-based prevention for adolescent internet addiction: prevention is the key. A systematic literature review. *Curr Neuropharmacol* 2019;17(6):507-525. [doi: [10.2174/1570159X16666180813153806](https://doi.org/10.2174/1570159X16666180813153806)] [Medline: [30101714](https://pubmed.ncbi.nlm.nih.gov/30101714/)]
100. Hamaker EL, Kuiper RM, Grasman R. A critique of the cross-lagged panel model. *Psychol Methods* 2015 Mar;20(1):102-116. [doi: [10.1037/a0038889](https://doi.org/10.1037/a0038889)] [Medline: [25822208](https://pubmed.ncbi.nlm.nih.gov/25822208/)]

Abbreviations

CFI: comparative fit index

CLPM: cross-lagged panel model

DSM-5: *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition)

FIML: full information maximum likelihood

H1: hypothesis 1

H2: hypothesis 2

H3: hypothesis 3

IGD: internet gaming disorder

RMSEA: root mean square error of approximation

SEM: structural equation modeling

SPP: socially prescribed perfectionism

SRMR: standardized root mean square residual

TLI: Tucker-Lewis index

Edited by S Brini; submitted 12.Feb.2026; peer-reviewed by C Hu, Y Liu; revised version received 06.Apr.2026; accepted 10.Apr.2026; published 30.Apr.2026.

Please cite as:

Zheng P, Tao Z, Yao T, Fang L, Zhao M

Socially Prescribed Perfectionism, Resilience, and Internet Gaming Disorder in Adolescents: 3-Wave Longitudinal Study

JMIR Serious Games 2026;14:e93412

URL: <https://games.jmir.org/2026/1/e93412>

doi: [10.2196/93412](https://doi.org/10.2196/93412)

© Peng Zheng, Zi Tao, Tao Yao, Luoxiang Fang, Min Zhao. Originally published in JMIR Serious Games (<https://games.jmir.org>), 30.Apr.2026. 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.

Development and Usability of the “FORTEe Get Strong” App to Promote Physical Activity and Health Awareness in Children and Adolescents With Cancer During Intensive Treatment Using an App-Based Approach: Mixed Methods Study

Mareike Kuehn^{1,2*}, MSc; Lena Wypyrsczyk^{1,2*}, MSc; Francesca Lanfranconi³, MD, PhD; Adriana Cristina Balduzzi^{4,5}, Prof Dr; Hayley Marriott⁶, MSc; Alba Solera-Sanchez⁶, PhD; Joachim Wiskemann⁷, Prof Dr Phil; Nikolai Bauer⁷, MEd; Nina Karguth⁸, MSc; Barbara Heißerer⁸, PhD; Norbert W Paul⁹, MD, PhD, Prof Dr rer medic; Rodolf Mongondry¹⁰, MSc; Wilhelm Bloch¹¹, Prof Dr Med; Katie Rizvi¹²; Martin Kaj Fridh¹³, PhD; Eva S C Ramos¹⁴, MSc; Carmen Fiuza-Luces¹⁵, PhD; Miriam Götte^{16,17}, PD Dr rer med; Gabriele Gauß¹⁷, MA; Filippo Spreafico^{18,19}, MD; Barbara Konda²⁰, BSc; Milica Stefanović²¹, MD, PhD; Elias Dreismickenbecker^{1,2}, MSc; Marie A Neu¹, Dr med; Joerg Faber¹, Prof Dr med; FORTEe Consortium^{22,23}

¹Childhood Cancer Center Mainz, University Medical Center of the Johannes Gutenberg University Mainz, Langenbeckstraße 1, Mainz, Germany

¹⁰Centre de Lutte Contre le Cancer, Centre Léon Bérard, Lyon, France

¹¹Department of Molecular and Cellular Sport Medicine, German Sport University Cologne, Cologne, Germany

¹²Youth Cancer Europe, Cluj-Napoca, Romania

¹³Department of Pediatrics and Adolescent Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

¹⁴Department of Sport Sciences, Faculty of Medicine, Health and Sports, Universidad Europea de Madrid, Madrid, Spain

¹⁵Research Institute Hospital 12 de Octubre, Madrid, Spain

¹⁶West German Cancer Center, Essen University Hospital, Essen, Germany

¹⁷Hematology and Oncology, Clinic for Pediatrics III, Essen University Hospital, Essen, Germany

¹⁸Oncology Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy

¹⁹Pediatric Oncology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

²Department of Sports Medicine, Prevention and Rehabilitation, Institute of Sport Science, Johannes Gutenberg-University Mainz, Mainz, Germany

²⁰Forma 3D Ltd, Ljubljana, Slovenia

²¹Ljubljana University Medical Centre, Ljubljana, Slovenia

²²See Acknowledgments

²³

³Fondazione Monza e Brianza per il bambino e la sua mamma, Centro Maria Letizia Verga, Monza, Italy

⁴School of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy

⁵Pediatrics, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy

⁶School of Sport, Nutrition and Allied Health Professions, Oxford Brookes University, Oxford, United Kingdom

⁷Department of Medical Oncology, Working Group Exercise Oncology, a partnership between DKFZ and University Medical Center Heidelberg, Heidelberg University Hospital and National Center for Tumor Diseases, Heidelberg, Germany

⁸Concentris research management GmbH, Fürstfeldbruck, Germany

⁹Institute for the History, Philosophy and Ethics of Medicine, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

*these authors contributed equally

Corresponding Author:

Elias Dreismickenbecker, MSc

Childhood Cancer Center Mainz, University Medical Center of the Johannes Gutenberg University Mainz, Langenbeckstraße 1, Mainz, Germany

Abstract

Background: As survival rates for children, adolescents, and young adults with cancer improve, managing treatment-related side effects is increasingly important. Enhancing physical activity levels has been shown to be effective in reducing some of these effects. Digital interventions, such as mobile apps, offer engaging tools to promote physical activity in young populations.

Objective: This article introduces the “FORTEe Get Strong” app (Nurogames GmbH) and presents a formative evaluation of its acceptability among children, adolescents, and young adults with cancer.

Methods: The “FORTEe Get Strong” app was developed within the multicenter FORTEe trial. Grounded in behavior change theories, the app uses gamification to deliver a child-friendly platform promoting physical activity and health-related knowledge. An embedded mixed methods design, with qualitative insights contextualizing quantitative findings, was applied. User experience was assessed using a self-developed questionnaire combining Likert scale items (1 - 5: “not at all satisfied” to “very satisfied,” 1 - 3 for participants aged <8 y) and open-ended questions. Quantitative data were analyzed descriptively and inferentially, including chi-square tests for differences in app usage by age, sex, and diagnosis, and Kruskal-Wallis tests to compare app feature ratings between age groups. Free text responses underwent qualitative content analysis.

Results: The app was published in May 2023; it integrates gamified features to teach health knowledge. Exercise videos address endurance, strength, flexibility, coordination, and gait. Feedback on user experience was provided by 53 patients aged 5 - 19 years (mean 11.4, SD 3.8 years, 95% CI 10.4 - 12.5; 27/53, 51% male). App use did not differ significantly by age groups ($\chi^2_3=5.1$; $P=.16$; Cramer $V=0.135$), sex ($\chi^2_1=3.4$; $P=.06$; $\phi=0.110$), or diagnosis ($P=.54$; Cramer $V=0.168$). The app was generally well-received, with a mean Likert scale score of 4.14 (SD 1.04, 95% CI 3.84 - 4.43). For individual components, mean scores ranged from 3.52 (SD 1.44, 95% CI 3.09 - 3.95) for exercise content to 4.37 (SD 0.82, 95% CI 4.14 - 4.60) for design. Kruskal-Wallis tests revealed no significant differences among age groups. Qualitative analysis identified key categories regarding overall app evaluation, specific features, target group perception, usability, and suggestions for improvement. While participants highlighted the app’s design and gamification, concerns arose regarding age-appropriateness, lack of supervision, and adaptability of exercises.

Conclusions: This study provides novel insights into user experience of children, adolescents, and young adults with cancer with a gamified exercise and health education app. Unlike previous studies focusing on survivorship, this evaluation offers a comprehensive understanding of how digital tools can support patients in maintaining physical activity during intensive treatment. Apps with interactive, gamified elements may complement clinical care by providing low-threshold access to exercise and health education. The findings advance the field by identifying key facilitators and barriers to engagement. Future research should assess adherence, behavioral outcomes, and effectiveness in larger samples to inform clinical implementation.

Trial Registration: ClinicalTrials.gov NCT05289739; <https://clinicaltrials.gov/study/NCT05289739>

(*JMIR Serious Games* 2026;14:e75653) doi:[10.2196/75653](https://doi.org/10.2196/75653)

KEYWORDS

childhood cancer; exercise; oncology; serious game; gamification; mHealth; mixed methods study; mobile phone

Introduction

Cancer is the leading cause of disease-related death among children, adolescents, and young adults. Approximately 14,000 new diagnoses were estimated in Europe in 2022 in children, adolescents, and young adults aged 0 to 19 years. Advances in treatment have significantly improved survival rates, leading to a survival rate of 81% in the mentioned age group [1]. However, this remarkable improvement in survival rates underscores the growing importance of addressing the short and long-term side effects [2,3]. Side effects can include fatigue, muscle weakness, reduced mobility, and long-term effects, such as cardiovascular diseases and metabolic disorders [2,3]. Exercise programs may offer an effective approach to addressing these issues [4]. Various studies demonstrate that exercise interventions can enhance physical aspects, such as muscle strength [5-7] and endurance capacity [5,7], as well as psychosocial outcomes, such as cancer-related fatigue [5,8] and quality of life [9,10]. Given the considerable side effects associated with the diagnosis and treatment of cancer, there is an urgent need for innovative, motivational, and age-appropriate interventions to support young patients in their physical activity and to counteract the observed side effects [11].

The increasing integration of digitalization and mobile technologies into everyday life has profoundly accelerated the

use of digital tools in various domains of health promotion, including encouraging physical activity [12-15]. Among young populations, digital interventions, such as mobile apps, have become an effective and engaging tool to promote health-related behaviors [16,17]. A substantial number of these so-called mobile health (mHealth) apps designed for children, adolescents, and young adults target a variety of health outcomes, most notably physical activity [16,17].

To ensure the effectiveness of these digital interventions and to keep users motivated to engage in health-related behaviors, it is crucial to understand the mechanisms and design features that drive behavioral change, particularly in young populations. In this context, Baumann et al [16] highlight the efficacy of gamified approaches in promoting physical activity and reducing sedentary behavior in children, adolescents, and young adults, providing that these interventions are supported by individualization, a robust theoretical framework, and the integration of behavior change techniques. The role of social connections, including family, peers, and the wider community, is identified as a key factor influencing the success of mHealth strategies, demonstrating the need for interventions tailored to developmental stages and specific needs to reduce physical inactivity and sedentary behavior. Complementing this, Ghosh et al [18] investigate critical app design features that increase continued engagement and behavioral impact among

adolescents. These include user-friendly and aesthetically appealing interfaces, personalized goal-setting, comprehensive data tracking, and access to educational content, such as exercise recommendations and nutritional advice. Additionally, the integration of electronic health records has the potential to facilitate data sharing with health care providers, and gamification and social interaction components are known to enhance intrinsic motivation and lead to a sense of group belonging [18]. The findings emphasize the importance of theoretically grounded and thoughtfully designed mHealth interventions in achieving substantial and lasting health-related behavior improvements in young populations.

The Europe-wide, multicenter FORTEe trial (Get strong to fight childhood cancer – An exercise intervention for children and adolescents undergoing anti-cancer treatment) aims to implement exercise as a therapy and increase physical activity using digital approaches in children, adolescents, and young adults with cancer. As part of the trial, the “FORTEe Get Strong” app (Nurogames GmbH) was developed with the primary aim of engaging children, adolescents, and young adults with cancer in physical activity and secondarily to teach health-related behaviors. Details of the study design have been published previously [19].

Despite growing evidence on mHealth solutions for young populations, little is known about how children, adolescents, and young adults with cancer themselves perceive such apps and which refinements might be necessary to support long-term engagement. Addressing these gaps requires both quantitative assessments of user experience and qualitative insights into individual needs and expectations. Therefore, the objectives of this article are to (1) introduce the “FORTEe Get Strong” app and describe its design and development, and (2) present a formative mixed methods evaluation assessing the app’s acceptability in a sample of children, adolescents, and young adults with cancer. The aim is to evaluate user experience ratings quantitatively and explore user perceptions qualitatively to contextualize the quantitative findings. Finally, both strands are integrated in a formative mixed methods evaluation, with the qualitative findings providing insight into, and helping to interpret, the quantitative data.

Methods

App Development

Overview

The “FORTEe Get Strong” app represents a component of the overall FORTEe exercise intervention, facilitating the transition from supervised, hospital-based training to self-directed exercise at home. The mobile app was developed by scientific and clinical experts of the FORTEe consortium and engineered by Nurogames GmbH. Nurogames GmbH is a software development company with expertise in creating games that, in addition to their entertainment value, contribute to the investigation of specific topics and the generation of scientific knowledge. To ensure the app resonated with its target group, patient representatives were actively involved in the design process. Feedback from patient interviews regarding their

preferred theme and avatar design was essential in understanding patients’ wishes and creating an appealing and relatable design. The app has been translated into 7 languages (Danish, English, French, German, Italian, Slovenian, and Spanish) for use on Android (Google LLC) and iOS (Apple Inc) devices. It is freely available for download from the App Store (Apple Inc) and Google Play Store (Google LLC), respectively.

Theoretical Framework

The development of the app is based on a combination of the social cognitive theory (SCT) [20] and self-determination theory (SDT) [21]. SCT explains the interaction between personal factors, environmental influences, and behaviors, emphasizing the importance of self-efficacy, an individual’s belief in his or her ability to succeed. This belief influences goals, outcome expectations, and behavioral adoption, with environmental facilitators and barriers shaping these processes. The app’s integration of evidence-based exercise and other health-related content (eg, on healthy diet) is intended to enhance self-efficacy in patients with cancer, enabling them to set meaningful goals, create actionable strategies, and develop skills to engage in regular physical activity [20,22,23]. SDT highlights the importance of fulfilling the psychological needs for autonomy, competence, and relatedness to enhance intrinsic motivation. When these needs are met, individuals experience greater autonomous motivation, leading to sustainable behavior change. For instance, gamification features (eg, a customizable avatar and interactive quizzes) may enhance relatedness and competence, while allowing patients to select and adapt activities to their preferences and individual needs may increase autonomy [21,23,24]. Behavior change techniques, like SCT and SDT, have previously been implemented as a guideline for an innovative mobile app designed for children, adolescents, and young adults with cancer in a study by Fueemler et al [23]. While initial evidence for actual behavior change in that study was limited, the intervention demonstrated feasibility, positive user engagement, and overall favorable reception, supporting the use of theory-based strategies as a guiding framework for developing our app.

Data Protection

For data protection reasons, no personal data (eg, duration or frequency of app usage) is processed in relation to the use of the “FORTEe Get Strong” app due to its implementation in a clinical setting. The data attained in the app are stored exclusively on the local device. Upon uninstallation, the app is restored to its original factory settings. As such, it does not collect or process personal data, nor does it facilitate the transfer of health information, such as electronic health records.

App Design, Content, and Functionality

The app’s gamification is characterized by a playful aesthetic, focusing on incentivizing physical activity as a health promotion strategy and directly providing access to individually tailored exercises. The purpose of the design of the app is to raise awareness of the importance of regular physical activity during cancer treatment and to promote intrinsic motivation for continued use of the mHealth app [18,25]. In accordance with the recommendations by Ghosh et al [18], the objective was to

provide simple and accessible guidelines for various exercises, thereby empowering young users to engage in independent physical activity and support self-efficacy. The exercises were developed by the Network Active-Onco-Kids and compiled in collaboration with exercise professionals from the FORTEe consortium, ensuring their adaptation to the needs of young patients with cancer. The compilation of all educational content, as well as the exercises themselves, was undertaken by a group of scientific researchers, as well as Network ActiveOncoKids and Youth Cancer Europe, and in line with preliminary German exercise guidelines in pediatric oncology [26], with the objective of providing content that is evidence-based, freely accessible, and easily comprehensible. To further support the gamification characteristics of the app, as recommended by Fuemmeler et al [23], Baumann et al [16], and Ghosh et al [18], various mini-games are integrated in the app. To enhance motivation and thereby support continued app use, patients are encouraged through a system of rewards to engage with the extensive knowledge content and to participate in the provided exercise sessions on a regular basis.

Evaluation of the App

Study Design and Rationale

As part of the FORTEe trial, the aim of this article is to conduct a formative evaluation of the mHealth “FORTEe Get Strong” app and gain insights into the experiences of children, adolescents, and young adults with cancer. This study represents a subproject and thus a secondary analysis within the overall FORTEe trial. This study adhered to the Mixed Methods Reporting in Rehabilitation and Health Sciences checklist (Checklist 1) [27].

A mixed methods approach was chosen in order to provide a comprehensive evaluation of the app, by combining quantitative and qualitative data. Quantitative data, such as user numbers and structured app ratings, capture levels of adoption and satisfaction and enable comparison across participant groups. Complementary qualitative analysis of open-ended responses adds depth by explaining why users rated the app as they did, highlighting specific usability issues and perceived benefits, and suggesting improvements. The embedded design, in which qualitative insights help explain and contextualize quantitative findings, allows both strands to complement each other, increasing the interpretability of user experiences and guiding targeted app refinements. Therefore, a self-developed questionnaire was administered between May 2023 and February 2025 with patients who participated in the randomized controlled FORTEe trial (prospectively registered on March 21, 2022). A self-developed tool was used as no existing validated instrument covered the specific app features that were intended to be assessed. Participants either responded in an interview format or answered the questions independently in written form. The interviews were not recorded. As the questionnaire or interview guidance was exploratory and context-specific, no formal psychometric validation was conducted. However, the items were reviewed and refined by pediatric oncologists (n=2) and exercise experts (n=4) on criteria including clarity of wording and developmental appropriateness. In order to minimize the influence of parents or study staff, participants were encouraged

to respond independently. Parents or study staff were only permitted to provide clarification when necessary. Whenever feasible, participants were seated separately from their parents to support independent responding. The English version of the questionnaire is provided as [Multimedia Appendix 1](#).

The integration of quantitative and qualitative data, expert review of the questionnaire or interview guidance items, and the standardized administration procedures used were intended to ensure methodological rigor and enhance the credibility and trustworthiness of the findings.

Participants and Recruitment

Participants were recruited within the multicenter FORTEe trial [19]. Patients were eligible if they had been diagnosed with cancer according to the International Classification of Childhood Cancer, Third edition, were between 4 and 21 years of age, and received chemotherapy and/or radiotherapy at one of the FORTEe trial sites. Participants for the app evaluation were purposively sampled from the FORTEe trial cohort, as they were the intended users of the “FORTEe Get Strong” app and could provide relevant insights into its acceptability and usability. The questionnaire has been performed at all participating recruitment centers of the FORTEe trial—University Medical Center of the Johannes Gutenberg-University Mainz, Childhood Cancer Center, Mainz (Germany); Fondazione Monza e Brianza per Il Bambino e La Sua Mamma, Monza (Italy); School of Sport, Nutrition and Allied Health Professions, Oxford Brookes University, Oxford (United Kingdom); Heidelberg University Hospital and National Center for Tumor Diseases, a partnership between DKFZ and University Medical Center Heidelberg, Germany, Department of Medical Oncology, Working Group Exercise Oncology, Heidelberg (Germany); Center de Lutte Contre le Cancer Léon Bérard, Lyon (France); Department of Pediatrics and Adolescent Medicine, Copenhagen University Hospital, Rigshospitalet, Copenhagen (Denmark); Universidad Europea de Madrid, Faculty of Medicine, Health and Sports, Department of Sport Sciences, Madrid (Spain); Research Institute of the Hospital 12 de Octubre (‘imas12’), Madrid (Spain); University Hospital Essen, West German Cancer Center, Essen (Germany); Fondazione IRCCS Istituto Nazionale dei Tumori, Pediatric Oncology Unit, Milan (Italy); and University Medical Center Ljubljana, Ljubljana (Slovenia) in cooperation with Forma 3D Ltd, Ljubljana (Slovenia).

Ethical Considerations

The FORTEe study protocol and all related documents were approved by the Ethics Committee of the Medical Chamber of Rhineland-Palatinate (application 2021 - 15904 on August 4, 2021) as well as the local ethics committees of all FORTEe trial sites and reviewed by the local data protection officer. Written informed consent was obtained from all participants, or from their legal guardians in the case of minors. Where possible, documented assent was collected from children aged 6 years and older. All participant data were pseudonymized before analysis. Identifying information was accessible only to authorized study personnel, safeguarding privacy and confidentiality. Participants in this evaluation did not receive any financial or material compensation. All procedures adhered

to the principles outlined in the Declaration of Helsinki. Neither the paper nor the supplementary materials contain images of the study participants. Any individuals depicted are nonparticipant staff members, and it is not possible to identify individual study participants.

Method of Assessment

Overview

The questionnaire was designed to gather participants' perspectives regarding their use of technology. With regard to the "FORTEe Get Strong" app, it included questions about user satisfaction with the app as a whole, the design of the app, the FORTEe avatar, the range of exercises, the FORTEe quiz, the knowledge content included and the app's user-friendliness. The interviews were conducted by local exercise professionals with experience in pediatric oncology in their native language and translated into English language for further analysis. Their previous understanding of childhood cancer treatment and exercise interventions informed probing questions, while efforts were made to minimize bias by following standardized guidelines and allowing independent participant responses. To further support comprehension, interviewers followed standardized guidelines that allowed them to adapt the wording of items to the developmental level of the participants while preserving the original meaning.

Assessment of Quantitative Data

The quantitative component of the evaluation consisted of questions presented on a 5-point or 3-point Likert scale, depending on the age of the participant. The response options ranged from "not at all satisfied" to "very satisfied" and were illustrated by child-friendly emojis. The 3-point scale, developed for use of participants under the age of 8 years, included only the response options 1 ("not at all"), 3 ("somewhat"), and 5 ("very much"). The 5-point scale, which includes the additional options 2 ("hardly") and 4 ("quite a lot"), is used by participants aged 8 years and older. This age-differentiated scaling approach was chosen because it provides younger children with fewer, more distinguishable response categories [28-31].

Assessment of Qualitative Data

Furthermore, participants were invited to provide additional comments regarding their thoughts on the "FORTEe Get Strong" app in an open-ended format, allowing for qualitative analysis of their individual experiences and perceptions.

Study Size

This formative evaluation was embedded within the main FORTEe randomized controlled trial. No separate power calculation was performed for this app evaluation. The overall FORTEe trial included an a priori sample size calculation for the primary outcome (cancer-related fatigue), targeting 450 participants [19]. The sample for this evaluation comprises FORTEe participants who took part in the app evaluation during the data collection period.

Data Analysis

Overview

The ages of the participants were recorded as continuous variables and subsequently grouped into 4 predefined categories to enable group-based comparisons. According to Sawyer et al [32], chronological age boundaries are both arbitrary and context-dependent, varying according to cognitive, social, and cultural factors. For the FORTEe trial and therefore for the "FORTEe Get Strong" app, age groups were defined based on cognitive development, literacy, and school-related competencies, in order to ensure age-appropriate interaction and consent procedures. The age groups are (1) 4 - 5 years (preschoolers), (2) 6 - 11 years (school children), (3) 12 - 16 years (adolescents), and (4) 17 - 21 years (young adults) [19].

Analysis of Quantitative Data

For quantitative analyses, descriptive continuous variables are expressed as mean values with SDs and 95% CIs, while categorical variables are presented as the actual number of participants, accompanied by the corresponding percentage. For analysis and visualization, responses from the 3-point scale (used for participants younger than 8 y) and the 5-point scale (used for participants aged 8 y and older) were combined, with equivalent points being interpreted as the same level of satisfaction. A chi-square test was performed to investigate whether app usage differed across the defined age groups, sex, or diagnosis. If the assumption that the expected frequency in at least 20% of cells was less than 5 was violated, a Fisher-Freeman-Halton exact test was performed instead. A Kruskal-Wallis test was used to investigate possible differences in the evaluation of individual app features by age group. The significance level was set at 5%.

Analysis of Qualitative Data

Complementing the quantitative evaluation, qualitative content analysis, as outlined by Mayring [33], was used to facilitate a systematic examination of the free-text responses. A pragmatic interpretive approach guided the analysis, focusing on understanding participants' experiences and perceptions of the app in context. The objective was to categorize the responses and record the frequency of occurrence. A repeated review of all responses was conducted to identify key themes, and categories were formed inductively. The process of open coding was used to extract the recurring themes and group them into overarching categories. This was performed independently by 2 researchers (MK and LW), and discrepancies were resolved through discussion. The categories identified reflect the principal topics addressed in the free text responses. Categories were developed inductively, ensuring consistent interpretation across the dataset. A quantitative analysis was then conducted, counting the frequency with which each category appeared in the responses. This provides an initial approximation of the relative importance of the different topics.

Mixed Methods Integration

Finally, both data strands were integrated using an embedded mixed methods design, in which quantitative data provided an overview of adoption and satisfaction, and qualitative insights helped explain and contextualize the quantitative findings. This

approach allows the complementary strengths of both methods to reinforce the conclusions, while acknowledging their respective limitations. The integrated analysis enables a richer interpretation of the results and informs targeted app refinement and implementation strategies, illustrating the added value of using a mixed methods design.

The descriptive statistical analyses were conducted using IBM SPSS (version 27), while qualitative content analysis was carried out using Microsoft Excel (version 2016, 16.0.5474.1002). The graphs were created with R statistical software (RStudio; version 2024.09.1).

Results

The App Development

Overview

The “FORTEe Get Strong” app ([Figure 1](#)) was published in May 2023 for use on Android (Google LLC) and iOS (Apple Inc) devices and is now freely available for download from the App Store (Apple Inc) and Google Play Store (Google LLC). Minor technical updates were made to the “FORTEe Get Strong” app during the evaluation period. These included correcting language inconsistencies, optimizing in-app texts, and resolving minor software bugs. The following versions were released, used, and evaluated by participants—version 1.0.1, version 1.1.1, and version 1.1.2. As all versions were functionally identical in terms of content and user interface, they were evaluated collectively.

Figure 1. The “FORTEe Get Strong” app.

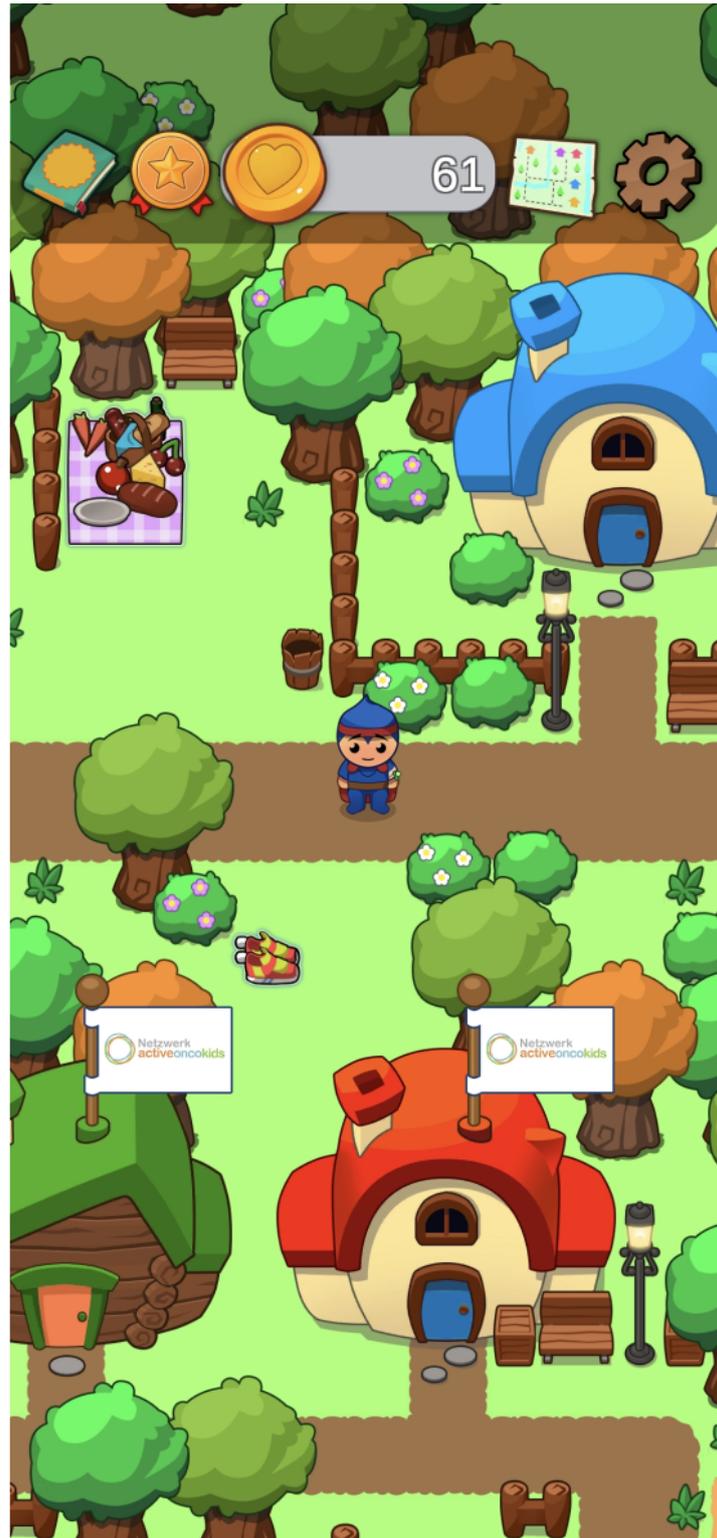


App Design, Content, and Functionality

The child-friendly app enables users to access a virtual environment, the FORTEe village (Figure 2). Within this virtual space, users can navigate their own avatar through a bright and colorful world containing educational content and entertaining

facts focusing on physical activity and providing a range of exercises to improve different physical fitness components. The avatar includes features such as an infusion needle, bandage, alopecia, and a superhero cape, symbolizing the experiences of children and adolescents with cancer and fostering a sense of identification and empowerment when using the app.

Figure 2. The FORTEe avatar can move around the FORTEe village, interact with the characters in the village, and can explore various health content.



The interaction with child-friendly characters living in the FORTEe village enables users to explore relevant knowledge regarding physical activity during cancer treatment, as well as general information on a healthy lifestyle in a playful way. The educational content is tailored for children, adolescents, and young adults with cancer and provides guidance on topics such as safe exercise during intensive treatment, changes in muscle strength, fall prevention, and healthy nutrition and an active lifestyle. This information is reinforced through quizzes and

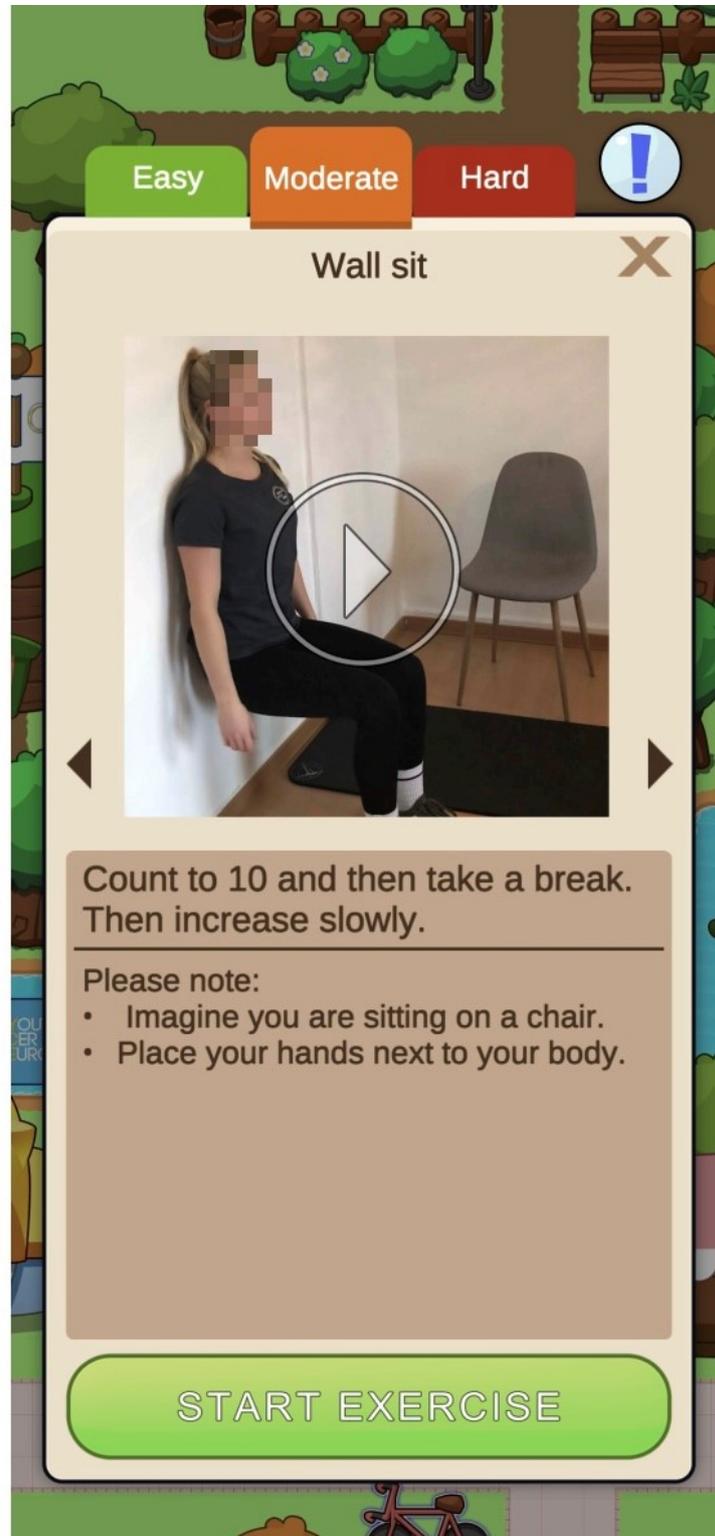
mini-games. Additionally, by visiting the FORTEe house or the numerous market stalls located in the village center, users can gather insights into the background of the app—the randomized controlled FORTEe trial.

The FORTEe village consists of a series of exercise houses, each designed to target physical fitness components, including endurance, strength, flexibility, coordination, balance, and gait. The exercises are presented as exercise videos, accompanied by instructions on correct execution and number of repetitions,

encouraging users to imitate and perform the demonstrated exercises themselves. The exercises are presented with a range of levels of difficulty, thus allowing for adaptation to the user's physical ability and current health status (Figure 3). Many

exercises can be performed while sitting or lying down, enabling participation by patients with reduced mobility or an increased risk of falling.

Figure 3. The app offers different exercises for different functions, eg, strength exercises (such as “wall sitting” with video instructions) and different levels of difficulty.



The app also provides various mini-games, including a FORTEe quiz (Figure 4), memory games, and others (Figures 5 and 6). Focusing primarily on physical activity, but also on a healthy diet, the mini-games are a fun way to introduce health-related

topics and provide an insight into the FORTEe trial and the European partners involved in the study. By integrating the smartphone's gyroscope, some interactive exercises within the app require users to move or shake the device in specific

directions to complete simple in-game challenges. This feature transforms the smartphone into an active interface that requires physical motion, thereby playfully encouraging body movement. Finally, the app contains a “Parents” section that provides guidance on how to support their child in exercising safely

during treatment. This section covers topics such as infection risk, postoperative activity, exercising with infusion stands, fall prevention, and information on peripheral neuropathy and bone metastases.

Figure 4. Dr Owls quiz is a fun way to learn about physical activity, exercise, and other health topics, as well as information about the FORTEe trial.

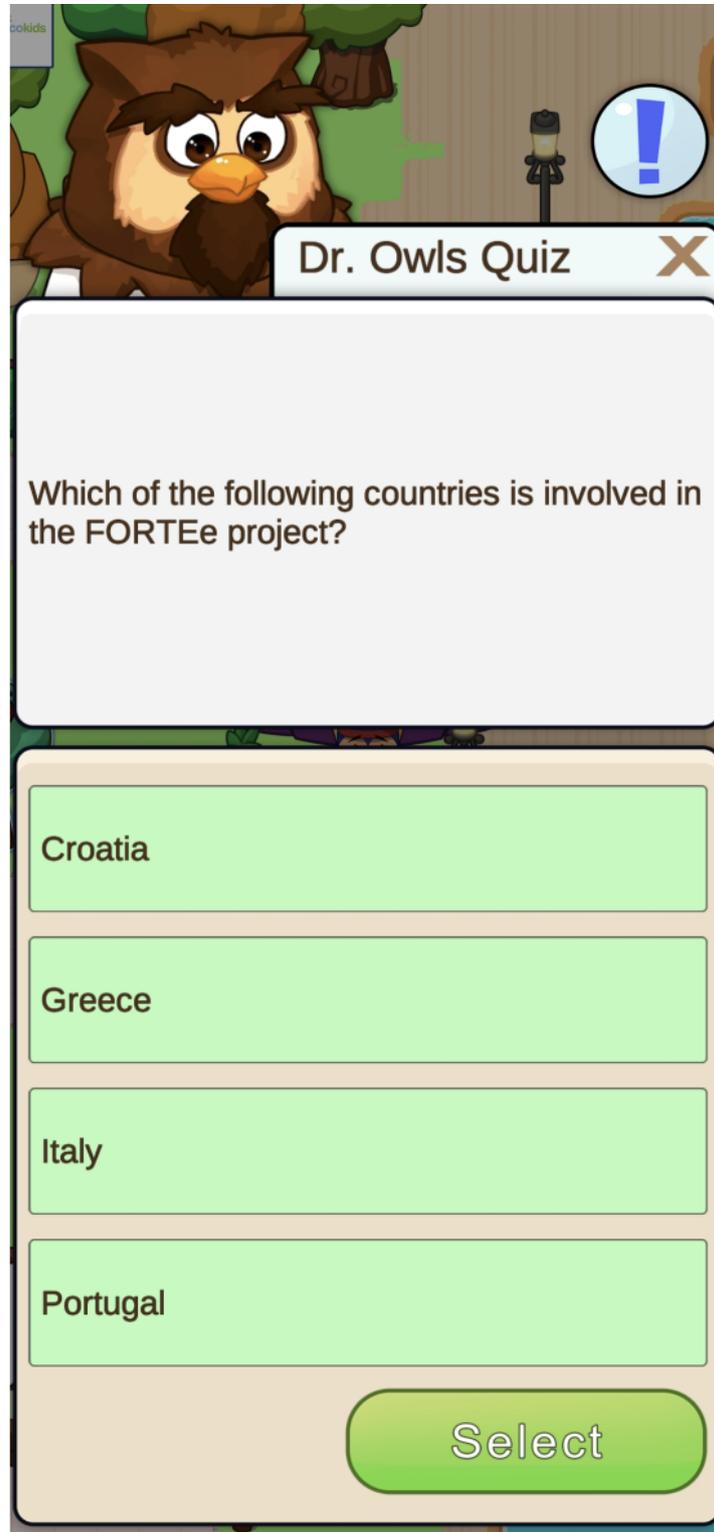


Figure 5. Other mini-games included in the app are a crossword puzzle, a memory game, and a nutrition puzzle.

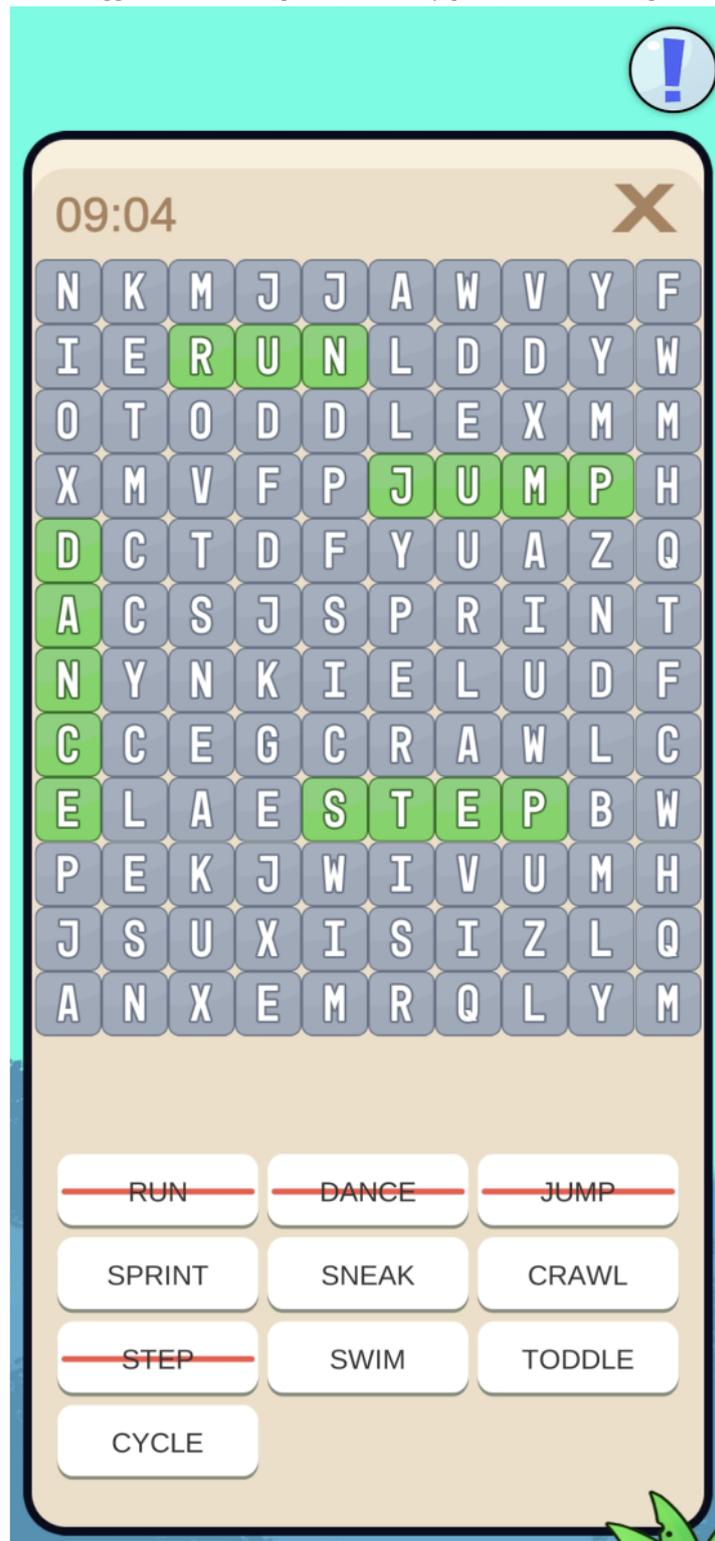


Figure 6. Some mini-games also offer the option to use the gyroscope of the device to directly encourage movement. This allows users to engage in activities such as running, cycling, or surfing with their avatar, overcoming various obstacles, helping the avatar in picking apples from a tree, or guiding a frisbee through the air by moving, tilting, or shaking the phone.



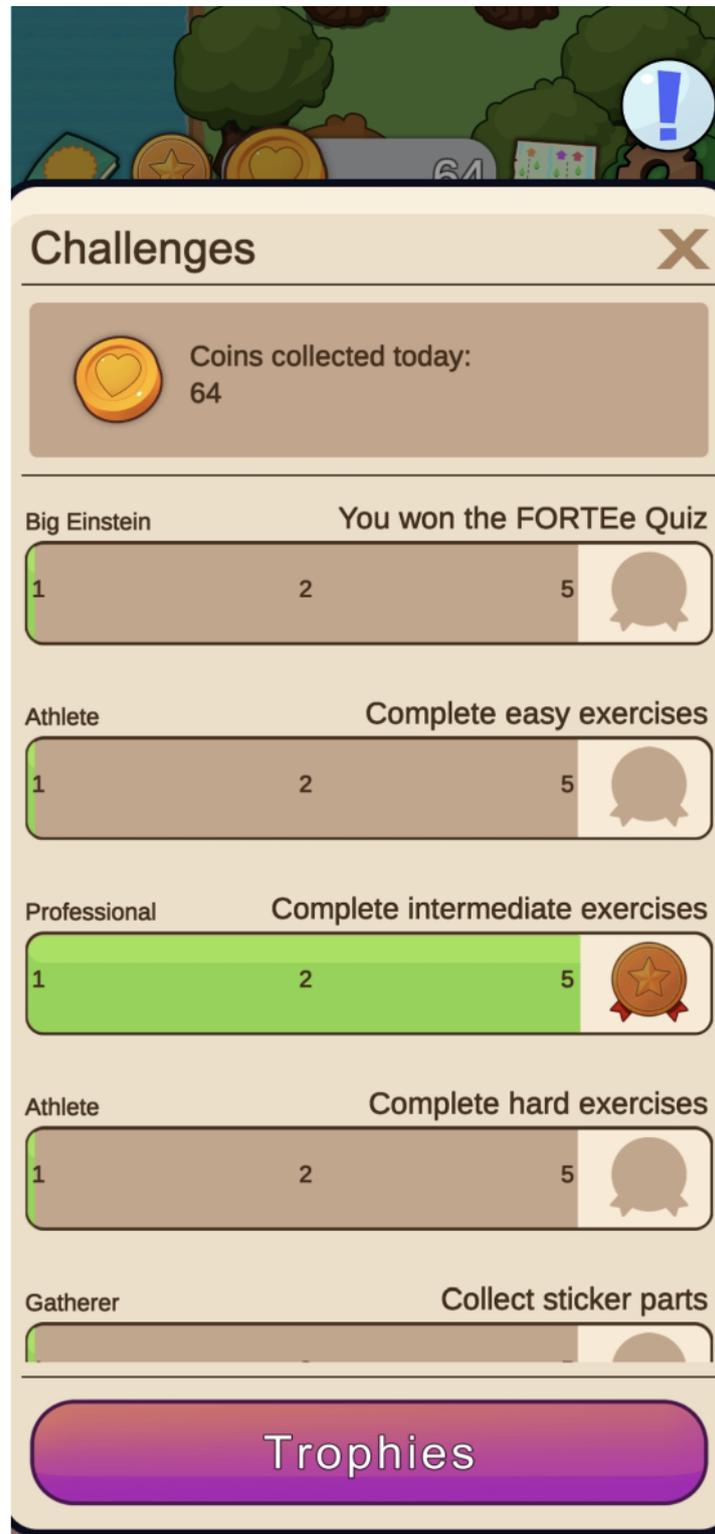
Patients are encouraged to use the app through a reward system. In addition to unlocking new content, patients will be rewarded with FORTEe coins, which can be used to personalize their avatar and customize their avatar's house. Various challenges also motivate exploration of the app and performance of exercise sessions. Completing challenges is rewarded with virtual medals

and trophies that can be viewed in the trophy collection, to further enhance the extrinsic motivation of the users (Figure 7).

In the context of supervised exercise sessions at the FORTEe treatment centers, exercise professionals have the option of providing patients with a code as part of the rewarding system. These codes can be used in the app to unlock puzzle pieces, which, when assembled, yield engaging exercise-related images.

This feature aims to enhance patient engagement, thereby making the experience more appealing and gratifying. promoting active participation in the supervised sessions, while

Figure 7. Various FORTEe challenges with the intention to motivate users to interact with the app and engage with the rich health content and the reward system.



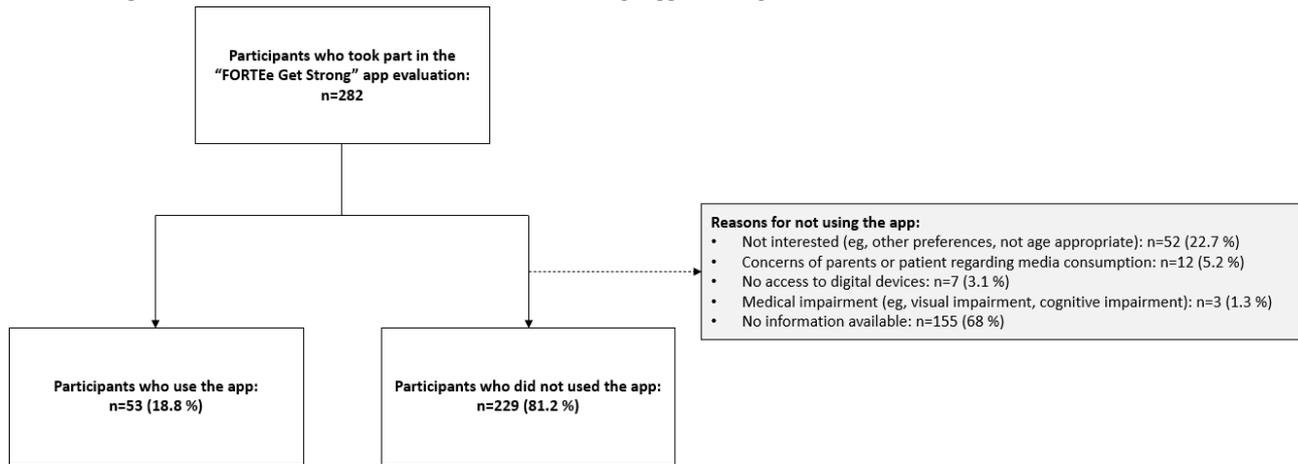
Evaluation of the App

Overview

A total of 282 participants took part in the evaluation regarding the “FORTEe Get Strong” app; thereof, 229 participants did

not use the app. The reasons for not using the app are presented in [Figure 8](#).

Figure 8. Participant flow in the evaluation of the “FORTEe Get Strong” app, showing user status and reasons for nonuse.



A total of 53 participants indicated having used the app and further provided detailed feedback. The number of responses varies between items because not all participants answered every question. Table 1 presents the characteristics of all participants who took part in the evaluation, as well as of those who used the app.

Group comparison analyses revealed no significant differences in app usage across age groups ($\chi^2_3=5.1$; $P=.16$; Cramer $V=0.135$), sex ($\chi^2_1=3.4$; $P=.06$; $\phi=0.110$), or diagnosis ($P=.54$; Cramer $V=0.168$).

Table . Characteristics of all evaluated participants and of those who used the app^a.

	All evaluated participants (n=282)	App users (n=53)
Age (y)		
Mean (SD)	11.8 (4.3)	11.4 (3.8)
Median (Range)	12 (4 - 21)	11 (5 - 19)
Age group, n (%)		
Pre-schoolers (4 - 5 y)	25 (9)	2 (4)
School children (6 - 11 y)	104 (37)	26 (49)
Adolescents (12 - 16 y)	109 (39)	18 (34)
Young adults (17 - 21 y)	44 (16)	7 (13)
Sex, n (%)		
Female	112 (40)	26 (49)
Male	170 (60)	27 (51)
Diagnosis, n (%)		
Leukemias	106 (38)	21 (40)
Lymphomas	59 (21)	12 (23)
CNS ^b tumors	25 (9)	7 (13)
Neuroblastoma	7 (2)	1 (2)
Renal tumors	4 (1)	1 (2)
Malignant bone tumors	40 (14)	7 (13)
Soft tissue tumors	16 (6)	1 (2)
Germ cell tumors	15 (5)	1 (2)
Others	10 (4)	2 (4)

^aPercentages may not total 100% due to rounding.

^bCNS: central nervous system.

Patient Experience of the “FORTEe Get Strong” App

Results of the Quantitative App Evaluation

Overall, the “FORTEe Get Strong” app was well-received by the patients surveyed, receiving an average rating of 4.14 (SD 1.04) points out of 5 points. A clear majority of 71% (36/51) of participants gave the app an overall positive rating (“I liked the app quite a lot” or “I liked the app very much”), while only 6% (3/51) gave a negative rating (“I hardly liked the app” or “I liked the app not at all”). In particular, the design (mean 4.37, SD 0.82 points; 42/51, 82% positive ratings; 1/51, 2% negative ratings), the user-friendliness (mean 4.14, SD 1.11 points; 37/50, 74% positive ratings; 4/50, 8% negative ratings), and the avatar (mean 4.14, SD 1.10 points; 36/51, 71% positive ratings; 3/51, 6% negative ratings) were well-received. The quiz received an average rating of 3.89 (SD 1.07) points, with 61% (23/38) positive and 8% (3/38) negative feedback.

Opinions on the exercises and the informational content were more varied. Although more than half of the participants rated

these features positively, nearly a quarter provided negative ratings. This represents a notably higher proportion of negative feedback than for the other app components. The average rating of the exercises was 3.52 (SD 1.44) points, with 52% (24/46) of participants rating the exercises positively and 22% (10/46) negatively. The informational content received an average rating of 3.64 (SD 1.48) points, with 60% (28/47) of participants giving positive ratings and 23% (11/47) giving negative ratings. Kruskal-Wallis tests revealed no significant differences in the evaluations of the various specific app features examined among age groups (Table 2).

The quantitative analysis demonstrated that the app was generally well-received, particularly in terms of its design, user-friendliness, and avatar. However, the ratings for exercises and informational content exhibited greater variability, with more than half of the participants providing positive ratings but nearly a quarter expressing negative feedback. Table 2 and Figure 9 present a summary of the quantitative rating of the app components.

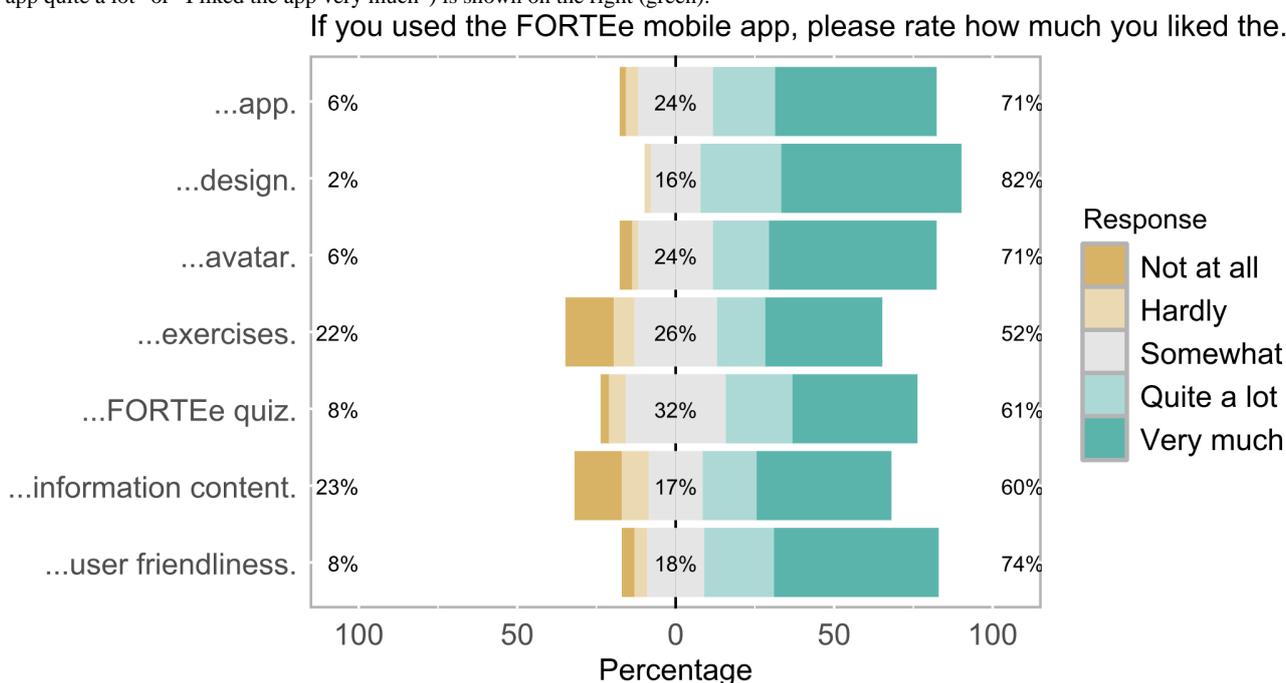
Table . Results of the quantitative app evaluation, including comparisons across age groups^{a,b}.

If you used the FORTEe mobile app, please rate how much you liked the...	n	Mean (SD)	95% CI	H (df)	P value
...app.	51	4.14 (1.04)	3.84 - 4.43	1.66 (3)	.65
...design.	51	4.37 (0.82)	4.14 - 4.60	1.42 (3)	.70
...avatar.	51	4.14 (1.10)	3.83 - 4.45	1.42 (3)	.70
...exercises.	46	3.52 (1.44)	3.09 - 3.95	1.81 (3)	.61
...FORTEe quiz.	38	3.89 (1.09)	3.54 - 4.25	3.27 (3)	.35
...informational content.	47	3.64 (1.48)	3.20 - 4.07	2.56 (3)	.46
...user friendliness.	50	4.14 (1.11)	3.83 - 4.45	0.99 (3)	.81

^aValues are presented as mean (SD) and 95% CIs.

^bGroup differences between age groups were analyzed using Kruskal-Wallis tests.

Figure 9. Results of the app evaluation. Participants rated various features of the “FORTEe Get Strong” app including the app in general, its design, the avatar, the exercises, the FORTEe quiz, the informational content as well as the user-friendliness rated on a scale from “Not at all” to “Very much.” The percentages at the bottom show the distribution of ratings relative to the central reference point (0%), with negative values extending to the left and positive values extending to the right. The percentage of negative ratings (“I liked the app not at all” or “I hardly liked the app”) is shown on the left (yellow), the percentage of neutral ratings (“I liked the app somewhat”) is shown in the middle (gray), and the percentage of positive ratings (“I liked the app quite a lot” or “I liked the app very much”) is shown on the right (green).



Results of the Qualitative App Evaluation

The qualitative analysis of the free text fields led to the identification of several central categories, such as (1) overall app evaluation, (2) evaluation of specific features (design, exercise content or app as a training tool, mini games, and informational content), (3) target group perception (age relevance), (4) app usability, and (5) suggestions for improvement, which summarize the experiences and opinions of the participants about the app. Thematic saturation was deemed achieved because responses repeatedly reflected the same set of themes, and no new categories emerged in the later stages of analysis. The main results are presented further in this study and in Table 3.

In total, 12 respondents provided additional positive feedback on the app in general, describing it as “very cool and fun” and said that the “app is nice during hospitalization,” highlighting the opportunity to “use it offline.” Moreover, 5 participants gave negative feedback, stating that the app had become uninteresting over time. With regard to the exercise content, patients (n=2) indicated that the exercises accessible via the app were beneficial. However, 5 patients stated that the range of exercises was less sophisticated than those offered in supervised

settings and that they preferred to engage in exercise sessions with an exercise professional. In addition, 5 patients provided feedback on the design of the app, with 4 patients specifically mentioning the design of the avatar. However, 1 patient gave negative feedback on the design of the app. Furthermore, 2 patients provided comments on the mini-games, offering exclusively positive feedback. The feedback on the informational content of the app was mixed. One participant commented that the information about other centers was “nice to learn about,” while 3 described the information as “too long” or stated that it was “more suitable for adults.” Of the total, 7 patients provided feedback on the suitability of the app for their age group. The most common feedback from adolescents and young adults was that the app was perceived as more useful and engaging for younger children, or that the patients themselves felt that they were too old for the app. Moreover, 2 patients provided feedback indicating that the game control did not function as effectively as desired, and that “the app is not very intuitive to use.” A total of 5 patients provided feedback on potential improvements to the app, giving both design-related suggestions and requests for additional features. Such suggestions included the addition of a chess game or a virtual fitness studio.

Table . Qualitative evaluation of the design and content of the “FORTEe Get Strong” app with main categories and subcategories to assess patient perspectives. Of the 53 patients included in the evaluation, 49 provided responses that could be assigned to the defined subcategories. Multiple responses per patient were possible, while some patients did not provide any feedback.

Main category and subcategory	Example quotes	Patients, n
Overall App evaluation		
General		
Positive	<ul style="list-style-type: none"> “The app is very cool and fun.” “The app is nice during hospitalization.” “Especially during radios (no internet) I was able to use it offline” 	12
Negative	<ul style="list-style-type: none"> “It feels like it’s not finished yet.” “The app got boring after a while.” 	5
Evaluation of specific features		
Design		
Positive	<ul style="list-style-type: none"> “I liked the avatar.” “The avatar is cute.” 	4
Negative	<ul style="list-style-type: none"> “I didn’t like the design of the app.” 	1
Exercise content or app as a training tool		
Positive	<ul style="list-style-type: none"> “I liked the exercises/workouts.” “The exercises are helpful.” 	2
Negative	<ul style="list-style-type: none"> “The app’s exercise offer is less sophisticated compared to other (supervised and supported) exercise offers.” “It is much more interesting to exercise with exercise professionals in real.” “The exercises are too easy.” 	5
Mini games		
Positive	<ul style="list-style-type: none"> “I like the games, they are funny.” “The quizzes are nice.” 	2
Informational content		
Positive	<ul style="list-style-type: none"> “It is nice to learn about other centers.” 	1
Negative	<ul style="list-style-type: none"> “The information content of the app is more suitable for adults.” “The information about the other centers is too long.” 	3
Target group perception		
Age relevance	<ul style="list-style-type: none"> “I am 20 years old and it is not funny for me. But nice idea for the youngest.” “Only a very small age group is suitable for the app.” “I like the games; they are funny, but more appropriate for younger children.” 	7
App usability	<ul style="list-style-type: none"> “The avatar moves too slowly. The game control doesn’t work well.” “The app is not very intuitive to use” 	2
Suggestions for improvement	<ul style="list-style-type: none"> “I would like the game design to be different, for example the first-person view of the player from the inside (more real).” “You must add a chess game.” 	5

Mixed Methods Integration

The integration of quantitative and qualitative data provided a comprehensive understanding of users' experiences with the "FORTEe Get Strong" app. Quantitative findings indicated an overall positive evaluation. Qualitative findings from open-ended responses supported and expanded these results. Participants frequently described the app as "fun," "cool," and particularly useful during hospitalization, confirming the generally positive quantitative ratings. However, several users expressed different concerns, providing context for the variability observed in the quantitative ratings of some features. Taken together, the integration of quantitative and qualitative data revealed converging findings regarding the app's strong usability and appeal, while qualitative insights further clarified user preferences and contextualized the quantitative variability observed for certain app features.

Discussion

Summary

This mixed methods study aimed to (1) introduce and describe the development of the "FORTEe Get Strong" app, and (2) evaluate its usability, acceptability, and user experience among children, adolescents, and young adults with cancer participating in the multicenter FORTEe trial. The "FORTEe Get Strong" app was developed to promote physical activity and health awareness among children and adolescents with cancer, by including behavior change theory and gamification principles to foster motivation, engagement, and long-term participation in physical activity. By integrating quantitative patterns with qualitative explanations, the embedded mixed methods approach was used to provide a comprehensive understanding of user experience. The evaluation included data from 53 children, adolescents, and young adults with cancer who used the app within the FORTEe trial. The quantitative assessment revealed that the app was generally well-received, with high ratings for the app in general, the design, the user-friendliness, and the avatar. Ratings for exercise content and informational materials were more variable. Group comparison analyses revealed no significant differences in app usage according to age group, sex, or diagnosis. Similarly, there were no significant differences in evaluations of certain app functions among the 4 age groups. Qualitative content analysis of open-ended responses produced several themes, such as "overall app evaluation," "evaluation of specific features" (including "design," "exercise content or app as a training tool," "mini games," and "informational content"), and "target group perception" dealing with "age appropriateness," "app usability," and "suggestions for improvement." Participants described the app as "fun" and "cool," and found it particularly useful during hospitalization. However, feedback also highlighted limitations, including a perceived lack of supervision, limited adaptability of exercises, and concerns about age appropriateness. Overall, the findings of this mixed methods evaluation indicate that the app is usable, engaging, and well-received by children, adolescents, and young adults with cancer, while also highlighting clear directions for future refinement to enhance its motivational and educational value.

Synthesis With Existing Knowledge

The App Development: Theoretical Framework

From a theoretical point of view, the "FORTEe Get Strong" app integrates key elements of SCT [20], particularly observational learning and self-efficacy. Exercise videos demonstrating the correct performance of exercises allow users to learn by watching, while adjustable difficulty levels build confidence in their abilities. The app further applies SCT principles through reinforcement via gamified rewards, outcome expectations conveyed through educational content, and reciprocal determinism by providing an adaptive, feedback-driven environment that promotes autonomy, long-term engagement, and long-term behavior change [20]. However, a critical aspect of Bandura's framework [20], social modeling through peer or caregiver interaction [22], is lacking, primarily due to data protection restrictions. This hinders the app's ability to promote social reinforcement, a key element in facilitating behavioral change [20,22]. The app's alignment with SDT [21] is evident in its promotion of autonomy and competence. Users are given autonomy to choose their activities, customize their avatars, and set their own pace, giving them a sense of control and personal achievement. The structured and progressive nature of the exercise options, reinforced by clear instructions and a reward system, serves to enhance competence. However, the limited adaptability of the exercises in the app (only 3 difficulty levels) may restrict opportunities for individualized progression and thus only partially address the SDT need for competence [21,24]. Similarly, the lack of continuous supervision and real-time social interaction limits the extent to which the app can support the SDT need for relatedness [21,24]. While features that connect the supervised training as part of the FORTEe exercise intervention with digital rewards received in the app provide some connection, the lack of real-time social interaction with peers or caregivers limits the sense of connection and community, which is critical to developing long-term intrinsic motivation [21,24]. In comparison with the recommendations outlined by Ghosh et al [18] for mHealth apps targeting adolescents, the "FORTEe Get Strong" app effectively incorporates gamification, evidence-based educational content, and personalized exercise programs. However, it does not adequately address the need for social connectedness and individualized progression, which are highly valued by young people, caregivers, and health care providers. Moreover, the absence of ongoing supervision further limits opportunities for personalized feedback and adaptive guidance, both of which could strengthen users' sense of competence and relatedness. The app's limitations regarding opportunities for real-time interaction, personalized feedback, and peer engagement are primarily due to strict data protection regulations and the multicenter study design. This trade-off exemplifies the tension between maintaining high standards of data protection and implementing features that could potentially support behavioral change. While the app successfully upholds high data protection standards, these constraints, stemming from the requirements of a multicenter study and strict privacy regulations, made real-time supervision and more individualized adaptation infeasible. Future interventions could explore the integration of privacy-preserving methods for social modeling,

adaptive feedback, and individualized progression to better address all 3 SDT needs of autonomy, competence, and relatedness [21,24].

Evaluation of the App by Patient Representatives

Of the 282 participants who completed the survey, 53 reported using the “FORTEe Get Strong” app and provided detailed feedback. While the overall number of app users was relatively small, understanding the reasons for nonuse provides important context. A considerable number of patients indicated disinterest in the app as a digital tool, preferring supervised exercise sessions at the hospital or independent training at home. This highlights the challenge of addressing diverse user needs and preferences within a single digital health solution. Some participants also perceived the app as not age-appropriate. However, this perception was not entirely reflected in the usage data. Although 37% (104/282) of all respondents were school-aged children between 6 and 11 years old, nearly half of the app users (26/53, 49%) were in this age group. Descriptively, app usage appeared higher among younger participants; however, the chi-square test revealed no significant difference by age. Similarly, despite 60% (170/282) of all respondents being male, app users were evenly distributed between females and males, and no significant difference in app usage regarding sex was found. Compared with previous studies linking older age and female sex to higher engagement with mHealth apps among adolescents [34,35], our data revealed no significant effects of age or sex. Nevertheless, the perception of the app as being more suitable for younger children may have reduced uptake among older adolescents, despite the absence of statistically significant age effects. Some parents also expressed concerns about increased media consumption resulting in their children not using the app. Those concerns are understandable, as excessive media consumption in children and adolescents can be associated with negative effects on sleep, attention, and learning, as well as higher incidence of obesity and depression [36]. In order to overcome these challenges, app developers must find the right balance by designing tools that limit passive screen engagement and encourage active participation. Transparent communication with caregivers about the purpose of the app as a facilitator of healthier habits, rather than a replacement for physical activity, remains critical. By promoting trust and addressing parental concerns, the app can serve as a valuable resource to support the health and well-being of children, adolescents, and young adults with cancer.

Given the wide range of cancer entities, a small number of patients were unable to use the app due to medical reasons, such as vision impairments or cognitive deficits. Despite efforts to develop a tool tailored to the needs of children, adolescents, and young adults with cancer, addressing the diverse diagnoses and associated limitations remains challenging. Future adaptations, such as audio feedback or caregiver guidance, could enhance accessibility and inclusion for this subgroup of patients, ensuring broader usability and engagement. A considerable number of participants did not provide further information on their reasons for not using the app. Given that the FORTEe trial is a multicomponent intervention encompassing exercise sessions, fitness assessments, questionnaires, video training, and other digital tools, the app was offered as an optional

component with voluntary participation. This may have led some participants to perceive the app as supplementary rather than essential, which could have contributed to its lower engagement and limited feedback from nonusers.

Although the number of users was relatively small, their feedback still provides valuable insights into the effectiveness and acceptance of the “FORTEe Get Strong” app among children, adolescents, and young adults with cancer. Designed specifically for this population, the app received overall positive feedback for its design, user-friendliness, and interactive features such as the avatar and quiz. These elements likely played an important role in making the app engaging and appealing to young patients—an essential factor when addressing the challenges of motivating children to participate in physical activity and engage in other health-related topics [16,18]. Suggestions for additional features, such as a chess game or a virtual fitness studio, further highlight the patients’ preferences for more interactivity and variety. These findings provide valuable guidance for future app development, ensuring that the app becomes a more comprehensive and dynamic tool to support children, adolescents, and young adults with cancer.

However, certain areas of the app, particularly the exercise content and informational features, also received some less favorable ratings. This can be explained by the feedback from young users, who expressed a preference for supervised face-to-face training and the opportunity for more interaction. For example, several users noted that the exercises were less individually tailored compared with those provided during supervised training sessions, underlying the importance of supervision and personalization. Future versions of the app could address this gap by including a wider range of exercises with more adjustable difficulty levels to meet the diverse needs of users. Furthermore, the qualitative feedback about the app’s suitability across different age groups indicated that the target age range was defined quite broadly, ranging from 4 to 21 years. While some patients appreciated the app, particularly during hospitalization, some adolescents and young adults found it less engaging and felt it was better suited for younger children. However, Kruskal-Wallis tests of the Likert-scale ratings of specific app features showed no significant differences across the 4 defined age groups, suggesting that age does not systematically affect app ratings in this sample. As a result, the development of future digital health technologies should be tailored to the different developmental stages of users to better address their individual needs. Furthermore, the involvement of users throughout the development process can further ensure that an app effectively addresses the interests and needs of its target audience. Therefore, patient representatives were consulted on as many aspects of the app as possible, including specific design aspects such as the main theme and avatar design to make the app appealing to the target group. However, even when user preferences are taken into account, not everyone’s taste can be catered for, which may explain the feedback on age appropriateness.

Limitations

A notable limitation of this study is that the evaluation was based on a relatively small sample (n=53) using and providing

feedback on the app at the time of the assessment. As this evaluation was embedded within the larger FORTEe trial, the subgroup of app users was smaller than the trial's overall sample, and no dedicated power calculation was performed for the substudy. Consequently, the study may have limited power to detect small differences between groups. As previously discussed, the small sample size in this investigation may be partly due to the comprehensive nature of the FORTEe trial.

Previous studies have shown that children are less likely to engage with apps that have a strong educational component and tend to rate them lower than scientific experts [37], which could have resulted in limited app use. Additionally, due to data protection regulations, the app usage duration was not recorded, resulting in limited insights into the frequency and consistency of engagement, as well as the interpretation of usage patterns and the relationship between intensity of use and perceived effectiveness. Similarly, adherence to the prescribed use of the app could not be objectively monitored; however, this might be critical to the app's potential effectiveness. Therefore, future studies should consider incorporating self-reported adherence measures to better understand engagement behavior while maintaining compliance with data protection requirements.

Furthermore, although the app was conceptually grounded in SCT and SDT, the evaluation did not include explicit measures of theory-driven constructs, such as self-efficacy, outcome expectations, autonomy, and competence support. Consequently, it is unclear to what extent these mechanisms were activated. Future research should incorporate targeted assessments to empirically evaluate these theoretical processes.

Additionally, despite being designed for a broad age range (4 - 21 y), some adolescents and young adults found the child-friendly design less engaging. Future iterations may therefore benefit from offering more age-differentiated design options to better accommodate diverse preferences across developmental stages.

Although efforts were made to minimize potential social desirability bias by encouraging participants to complete the questionnaire independently and limiting parental or staff assistance to clarification only, some degree of influence, particularly among younger children, cannot be ruled out. In future studies, this bias could be reduced further by systematically recording the presence of parents during completion of the questionnaire (eg, through a binary question about parental presence) and by standardizing the setting in which responses are collected.

Furthermore, while conducting the interviews in the native language of both the exercise professional and the participants likely improved the clarity and depth, this subsequent translation into English for documentation may have introduced information loss or translation errors. Additionally, some questions may have been difficult for younger children to fully comprehend, even in their native language, leading to potential misinterpretations or incomplete answers. These factors could affect the reliability and accuracy of the collected data.

Conclusion

This study demonstrates that the "FORTEe Get Strong" smartphone app can support children, adolescents, and young adults with cancer in maintaining physical activity and health education. This study provides novel insights into user experience of children, adolescents, and young adults with cancer with a gamified exercise and health education app during intensive treatment, a phase that has largely been underrepresented in pediatric oncology mHealth research. Unlike previous digital interventions that focused primarily on survivorship, this mixed methods evaluation provides a more comprehensive understanding of how children, adolescents, and young adults interact with digital tools during intensive treatment. The findings contribute to the field by identifying key facilitators and barriers to engagement, including exercise variety, age-appropriateness, and the need for privacy-compliant peer interaction, thereby informing the design of future pediatric oncology mHealth interventions. From a clinical perspective, digital tools can help bridge care gaps when exercise professionals are unavailable by providing low-threshold access to exercise and health education, when in-person support from exercise professionals is limited, for example, during inpatient stays or periods of treatment-related isolation. These findings indicate that gamified digital interventions can facilitate continuity of exercise engagement, complementing traditional care across different settings. Overall, digital tools like the "FORTEe Get Strong" app have the potential to become an integral part of the comprehensive, multidisciplinary approach by empowering young patients with cancer to stay active, engaged, and supported throughout their recovery and beyond. Moreover, the insights gained from this study may be applicable to other pediatric patient populations with chronic conditions, suggesting broader relevance for the development of gamified digital interventions. Future studies should evaluate adherence, behavioral outcomes, and the app's effectiveness in improving physical activity in larger samples to obtain a broader perspective on the implementation and impact of the "FORTEe Get Strong" app in the clinical setting of children, adolescents, and young adults with cancer.

Acknowledgments

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the GAIDeT taxonomy (2025), translation was delegated to GenAI tools under full human supervision. The GenAI tool used was DeepL (DeepL SE, Cologne, Germany). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

We would like to thank all the members of the FORTEe consortium for their support of this article. We especially thank the following consortium partners for their contribution, Christian Ruckes and Gabriele Schwarzländer for their support in data

curation; Peter Wright, Eila Watson, Hanne Bækgaard Larsen, Alejandro Lucia, and Lidija Kitanovski for providing resources such as study material and access to patients and supervision; and Jennifer Krick, Franziska Olivier, Tommaso Pietro Moriggi, Linda Peli, Eleonora Corti, Emanuele Villa, Olivia Pérol, Valentine Bertollin, Alicia Abadie, Martin Bjørn Hansen, Eva Santa Cruz-Ramos, Eva Garrido, Ronja Beller, William Zardo, Marco Chisari, Domen Ravnik, Meta Rován, and Neža Salobir for their great contribution to the FORTEe investigations. We would also like to thank Marco Polak and Andrew Pomasanski for their valuable contributions to the technical development of the “FORTEe Get Strong” app. Nurogames GmbH was responsible for the app’s development but was not involved in the data collection, analysis, or interpretation. We would further like to thank Heidi Diel and Jennifer Krick for their assistance in reviewing and refining the manuscript. Additionally, we thank the Network ActiveOncoKids that receives funding from the German Cancer Aid and Hernâni Oliveira from Youth Cancer Europe as well as all the FORTEe partner sites for their contribution to the content and evaluation of the app.

²²The members of the FORTEe Consortium are as follows:

Christian Ruckes, Dr rer physiol, University Medical Center of the Johannes Gutenberg-University Mainz, Interdisciplinary Center for Clinical Trials (IZKS), Mainz, Germany

Jennifer Krick, MSc, University Medical Center of the Johannes Gutenberg-University Mainz, Childhood Cancer Center Mainz, Mainz, Germany

Heidi Diel, Dr med, University Medical Center of the Johannes Gutenberg-University Mainz, Childhood Cancer Center Mainz, Mainz, Germany

Franziska Olivier, MSc, Heidelberg University Hospital and National Center for Tumor Diseases, a partnership between DKFZ and University Medical Center Heidelberg, Department of Medical Oncology, Working Group Exercise Oncology, 69120 Heidelberg, Germany

Gabriele Schwarzländer, MSc, Heidelberg University Hospital and National Center for Tumor Diseases, a partnership between DKFZ and University Medical Center Heidelberg, Department of Medical Oncology, Working Group Exercise Oncology, 69120 Heidelberg, Germany

Olivia Pérol, PhD, (1) Centre de Lutte Contre le Cancer Léon Bérard, Prevention Cancer Environment Department, Lyon, France; (2) INSERM U1296 Radiations: Defense, Health, Environment

Valentine Bertollin, MSc, Centre de Lutte Contre le Cancer Léon Bérard, Prevention Cancer Environment Department, Lyon, France

Alicia Abadie, Centre de Lutte Contre le Cancer Léon Bérard, Prevention Cancer Environment Department, Lyon, France

Peter Wright, PhD, Oxford Brookes University, School of Sport, Nutrition and Allied Health Professions, Oxford, UK

Eila Watson, PhD, Oxford Brookes University, Institute of Applied Health Research, Oxford, UK

Tommaso Pietro Moriggi MSc Fondazione Monza e Brianza per Il Bambino e La Sua Mamma, Centro Maria Letizia Verga, Monza, Italy

Emanuele Villa, MSc, Fondazione Monza e Brianza per Il Bambino e La Sua Mamma, Centro Maria Letizia Verga, Monza, Italy

Eleonora Corti, MSc, Fondazione Monza e Brianza per Il Bambino e La Sua Mamma, Centro Maria Letizia Verga, Monza, Italy

Linda Peli, MSc, Fondazione Monza e Brianza per Il Bambino e La Sua Mamma, Centro Maria Letizia Verga, Monza, Italy

Hernâni Zão Oliveira, PhD, Youth Cancer Europe, Youth Cancer Europe, Cluj-Napoca, Romania

Martin Bjørn Hansen, University Hospital Copenhagen, Rigshospitalet, Department of Pediatrics and Adolescent Medicine, Copenhagen, Denmark

Hanne Bækgaard Larsen PhD University Hospital Copenhagen, Rigshospitalet, Department of Clinical Medicine, Copenhagen, Denmark

Alejandro Lucia, MD, PhD, Department of Sports Sciences, Faculty of Medicine, Health and Sports, Universidad Europea de Madrid, Madrid, Spain

Eva Garrido, Hospital 12 de Octubre Research Institute, 'imas12, Instituto de Investigación, 28041 Madrid, Spain

Eva Santa Cruz Ramos, Hospital 12 de Octubre Research Institute, 'imas12,, Physical Activity and Health Laboratory, 28041 Madrid, Spain

Ronja Beller, Dr rer medic, Hematology and Oncology, Clinic for Pediatrics III, University Hospital Essen, West German Cancer Center, Essen, Germany

Marco Chisari, Pediatric Unit, Fondazione IRCCS Istituto nazionale dei Tumori, Milano, Italy

William Zardo Pediatric Unit, Fondazione IRCCS Istituto nazionale dei Tumori, Milano, Italy

Meta Rovan, University Medical Center Ljubljana, Division of Pediatrics, Department of Haematooncology, Ljubljana, Slovenia

Domen Ravnik, University Medical Center Ljubljana, Division of Pediatrics, Department of Haematooncology, Ljubljana, Slovenia

Lidija Kitanovski, MD, PhD, University Medical Center Ljubljana, Department for Pediatric Hemato-Oncology, Ljubljana, Slovenia

Neza Salobir, MD, University Medical Center Ljubljana, Department for Pediatric Hemato-Oncology, Ljubljana, Slovenia

Marco Polak, Nurogames GmbH, Cologne (DE)

Andrew Pomasanski, Nurogames GmbH, Cologne (DE)

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 945153. This publication reflects only the authors' views and the European Union is not liable for any use that may be made of the information contained therein. Research by CF-L is funded by the Spanish Ministry of Science and Innovation and (Fondo de Investigaciones Sanitarias, ISCIII) and Fondos FEDER (grant PI23/00396). Research by CF-L is also funded by ISCIII (FORT23/00023) and by Wereld Kanker Onderzoek Fonds (WKOF) as part of the World Cancer Research Fund International grant program (IG_FULL_2021_007). The funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated and analyzed during this study contain personal information that cannot be made publicly available due to data protection regulations, but are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: MK, LW, MAN, JF, FL, KR, GG

Data curation: MK, LW, MAN, ED, JW

Formal analysis: MK, LW

Funding acquisition: JF, MAN

Investigation: MK, LW, NB, HM, AS-S, RM, MKF, ESCR, CF-L, MS

Methodology: MK, LW, MAN, ED, GG, KR

Project administration: JF, MAN, ED,

Resources: FL, ACB, HM, JW, RM, WB, KR, MKF, CF-L, MG, GG, FS, BK, JF

Supervision: JF, MAN, ED, NK, BH, FL

Visualization: MK, LW, GG

Writing – original draft: MK, LW, MAN, ED

Writing – review & editing: MK, LW, FL, AB, HM, AS-S, JW, NB, NK, BH, NWP, RM, WB, KR, MKF, ESC-R, CF-L, MG, GG, FS, BK, MS, ED, MAN, JF.

MK and LW contributed equally to this work and share first authorship. MAN and JF share last authorship.

All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire / interview guidance of the FORTEe trial (English version).

[\[PDF File, 404 KB - games_v14i1e75653_app1.pdf\]](#)

Checklist 1

Mixed Methods Reporting in Rehabilitation and Health Sciences checklist.

[\[PDF File, 789 KB - games_v14i1e75653_app2.pdf\]](#)

References

1. ECIR - European Cancer Inequalities Registry: Childhood cancer. European Commission. 2023. URL: <https://cancer-inequalities.jrc.ec.europa.eu/childhood-cancer> [accessed 2025-03-14]
2. Cochran ED, Jacobson JC, Nehrubabu M, Qiao J, McCreery S, Chung DH. Social determinants of outcomes disparity among pediatric patients with solid tumor. *J Am Coll Surg* 2024 Apr 1;238(4):463-478. [doi: [10.1097/XCS.0000000000001010](https://doi.org/10.1097/XCS.0000000000001010)] [Medline: [38258890](https://pubmed.ncbi.nlm.nih.gov/38258890/)]
3. Erdmann F, Frederiksen LE, Bonaventure A, et al. Childhood cancer: survival, treatment modalities, late effects and improvements over time. *Cancer Epidemiol* 2021 Apr;71(Pt B):101733. [doi: [10.1016/j.canep.2020.101733](https://doi.org/10.1016/j.canep.2020.101733)] [Medline: [32461035](https://pubmed.ncbi.nlm.nih.gov/32461035/)]
4. Braam KI, van der Torre P, Takken T, Veening MA, van Dulmen-den Broeder E, Kaspers GJL. Physical exercise training interventions for children and young adults during and after treatment for childhood cancer. *Cochrane Database Syst Rev* 2016 Mar 31;3(3):CD008796. [doi: [10.1002/14651858.CD008796.pub3](https://doi.org/10.1002/14651858.CD008796.pub3)] [Medline: [27030386](https://pubmed.ncbi.nlm.nih.gov/27030386/)]
5. Stössel S, Neu MA, Wingerter A, et al. Benefits of exercise training for children and adolescents undergoing cancer treatment: results from the randomized controlled MUCKI trial. *Front Pediatr* 2020;8:243. [doi: [10.3389/fped.2020.00243](https://doi.org/10.3389/fped.2020.00243)] [Medline: [32582585](https://pubmed.ncbi.nlm.nih.gov/32582585/)]
6. Fiuzza-Luces C, Padilla JR, Soares-Miranda L, et al. Exercise intervention in pediatric patients with solid tumors: the physical activity in pediatric cancer trial. *Med Sci Sports Exerc* 2017 Feb;49(2):223-230. [doi: [10.1249/MSS.0000000000001094](https://doi.org/10.1249/MSS.0000000000001094)] [Medline: [27631396](https://pubmed.ncbi.nlm.nih.gov/27631396/)]
7. Saultier P, Vallet C, Sotteau F, et al. A randomized trial of physical activity in children and adolescents with cancer. *Cancers (Basel)* 2021 Jan 2;13(1):121. [doi: [10.3390/cancers13010121](https://doi.org/10.3390/cancers13010121)] [Medline: [33401713](https://pubmed.ncbi.nlm.nih.gov/33401713/)]
8. Coombs A, Schilperoord H, Sargent B. The effect of exercise and motor interventions on physical activity and motor outcomes during and after medical intervention for children and adolescents with acute lymphoblastic leukemia: a systematic review. *Crit Rev Oncol Hematol* 2020 Aug;152:103004. [doi: [10.1016/j.critrevonc.2020.103004](https://doi.org/10.1016/j.critrevonc.2020.103004)] [Medline: [32580035](https://pubmed.ncbi.nlm.nih.gov/32580035/)]
9. Morales JS, Valenzuela PL, Velázquez-Díaz D, et al. Exercise and childhood cancer-a historical review. *Cancers (Basel)* 2021 Dec 24;14(1):82. [doi: [10.3390/cancers14010082](https://doi.org/10.3390/cancers14010082)] [Medline: [35008246](https://pubmed.ncbi.nlm.nih.gov/35008246/)]
10. Braam KI, van Dijk-Lokkart EM, van Dongen JM, et al. Cost-effectiveness of a combined physical exercise and psychosocial training intervention for children with cancer: results from the quality of life in motion study. *Eur J Cancer Care* 2017 Nov;26(6):e12586. [doi: [10.1111/ecc.12586](https://doi.org/10.1111/ecc.12586)]
11. Kauhanen L, Järvelä L, Lähteenmäki PM, et al. Active video games to promote physical activity in children with cancer: a randomized clinical trial with follow-up. *BMC Pediatr* 2014 Apr 5;14:94. [doi: [10.1186/1471-2431-14-94](https://doi.org/10.1186/1471-2431-14-94)] [Medline: [24708773](https://pubmed.ncbi.nlm.nih.gov/24708773/)]
12. Mönninghoff A, Kramer JN, Hess AJ, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2021 Apr 30;23(4):e26699. [doi: [10.2196/26699](https://doi.org/10.2196/26699)] [Medline: [33811021](https://pubmed.ncbi.nlm.nih.gov/33811021/)]
13. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. *Transl Behav Med* 2014 Dec;4(4):363-371. [doi: [10.1007/s13142-014-0293-9](https://doi.org/10.1007/s13142-014-0293-9)] [Medline: [25584085](https://pubmed.ncbi.nlm.nih.gov/25584085/)]
14. Han M, Lee E. Effectiveness of mobile health application use to improve health behavior changes: a systematic review of randomized controlled trials. *Healthc Inform Res* 2018 Jul;24(3):207-226. [doi: [10.4258/hir.2018.24.3.207](https://doi.org/10.4258/hir.2018.24.3.207)] [Medline: [30109154](https://pubmed.ncbi.nlm.nih.gov/30109154/)]
15. Gu Y, Guan Y, Meng Z. Health providers' perceptions and experiences of using mHealth for chronic noncommunicable diseases: qualitative systematic review and meta-synthesis. *J Med Internet Res* 2023 Sep 12;25:e45437. [doi: [10.2196/45437](https://doi.org/10.2196/45437)] [Medline: [37698902](https://pubmed.ncbi.nlm.nih.gov/37698902/)]
16. Baumann H, Fiedler J, Wunsch K, Woll A, Wollesen B. mHealth interventions to reduce physical inactivity and sedentary behavior in children and adolescents: systematic review and meta-analysis of randomized controlled trials. *JMIR Mhealth Uhealth* 2022 May 11;10(5):e35920. [doi: [10.2196/35920](https://doi.org/10.2196/35920)] [Medline: [35544294](https://pubmed.ncbi.nlm.nih.gov/35544294/)]
17. Schoeppe S, Alley S, Rebar AL, et al. Apps to improve diet, physical activity and sedentary behaviour in children and adolescents: a review of quality, features and behaviour change techniques. *Int J Behav Nutr Phys Act* 2017 Jun 24;14(1):83. [doi: [10.1186/s12966-017-0538-3](https://doi.org/10.1186/s12966-017-0538-3)] [Medline: [28646889](https://pubmed.ncbi.nlm.nih.gov/28646889/)]
18. Ghosh P, Proffitt R, Bosworth KT, et al. mHealth app features that facilitate adolescent use for lifestyle management, and are endorsed by caregivers and health care providers. *Mhealth* 2024;10(21):21. [doi: [10.21037/mhealth-24-3](https://doi.org/10.21037/mhealth-24-3)] [Medline: [39114461](https://pubmed.ncbi.nlm.nih.gov/39114461/)]
19. Neu MA, Dreismickenbecker E, Lanfranconi F, et al. Get strong to fight childhood cancer - an exercise intervention for children and adolescents undergoing anti-cancer treatment (FORTEe): rationale and design of a randomized controlled exercise trial. *BMC Cancer* 2025 Aug 7;25(1):1275. [doi: [10.1186/s12885-025-14489-y](https://doi.org/10.1186/s12885-025-14489-y)] [Medline: [40775303](https://pubmed.ncbi.nlm.nih.gov/40775303/)]
20. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*: Prentice-Hall, Inc; 1986.
21. 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](https://doi.org/10.1037//0003-066x.55.1.68)] [Medline: [11392867](https://pubmed.ncbi.nlm.nih.gov/11392867/)]
22. Beauchamp MR, Crawford KL, Jackson B. Social cognitive theory and physical activity: mechanisms of behavior change, critique, and legacy. *Psychol Sport Exerc* 2019 May;42:110-117. [doi: [10.1016/j.psychsport.2018.11.009](https://doi.org/10.1016/j.psychsport.2018.11.009)]

23. Fuemmeler BF, Holzwarth E, Sheng Y, et al. Mila Blooms: a mobile phone application and behavioral intervention for promoting physical activity and a healthy diet among adolescent survivors of childhood cancer. *Games Health J* 2020 Aug 1;9(4):279-289. [doi: [10.1089/g4h.2019.0060](https://doi.org/10.1089/g4h.2019.0060)] [Medline: [32391734](https://pubmed.ncbi.nlm.nih.gov/32391734/)]
24. 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 Jun 22;9(1):78. [doi: [10.1186/1479-5868-9-78](https://doi.org/10.1186/1479-5868-9-78)] [Medline: [22726453](https://pubmed.ncbi.nlm.nih.gov/22726453/)]
25. Edwards EA, Lumsden J, Rivas C, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. *BMJ Open* 2016 Oct;6(10):e012447. [doi: [10.1136/bmjopen-2016-012447](https://doi.org/10.1136/bmjopen-2016-012447)] [Medline: [27707829](https://pubmed.ncbi.nlm.nih.gov/27707829/)]
26. Götte M, Gauß G, Dirksen U, et al. Multidisciplinary Network ActiveOncoKids guidelines for providing movement and exercise in pediatric oncology: consensus - based recommendations. *Pediatr Blood Cancer* 2022 Nov;69(11):e29953 [FREE Full text] [doi: [10.1002/pbc.29953](https://doi.org/10.1002/pbc.29953)] [Medline: [36073842](https://pubmed.ncbi.nlm.nih.gov/36073842/)]
27. Tovin MM, Wormley ME. Systematic development of standards for mixed methods reporting in rehabilitation health sciences research. *Phys Ther* 2023 Nov 4;103(11):pzad084. [doi: [10.1093/ptj/pzad084](https://doi.org/10.1093/ptj/pzad084)] [Medline: [37672215](https://pubmed.ncbi.nlm.nih.gov/37672215/)]
28. Mellor D, McCabe M, Ricciardelli L, Ball K. Body image importance and body dissatisfaction among Indigenous Australian adolescents. *Body Image* 2004 Sep;1(3):289-297. [doi: [10.1016/j.bodyim.2004.05.003](https://doi.org/10.1016/j.bodyim.2004.05.003)] [Medline: [18089160](https://pubmed.ncbi.nlm.nih.gov/18089160/)]
29. Mellor D, Moore KA. The use of Likert scales with children. *J Pediatr Psychol* 2014 Apr 1;39(3):369-379. [doi: [10.1093/jpepsy/jst079](https://doi.org/10.1093/jpepsy/jst079)] [Medline: [24163438](https://pubmed.ncbi.nlm.nih.gov/24163438/)]
30. van Laerhoven H, van der Zaag - Loonen H, Derckx B. A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatr* 2004 Jun;93(6):830-835 [FREE Full text] [doi: [10.1111/j.1651-2227.2004.tb03026.x](https://doi.org/10.1111/j.1651-2227.2004.tb03026.x)]
31. Wright KD, Asmundson GJG. Health anxiety in children: development and psychometric properties of the Childhood Illness Attitude Scales. *Cogn Behav Ther* 2003 Nov;32(4):194-202. [doi: [10.1080/16506070310014691](https://doi.org/10.1080/16506070310014691)] [Medline: [16291551](https://pubmed.ncbi.nlm.nih.gov/16291551/)]
32. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence. *Lancet Child Adolesc Health* 2018 Mar;2(3):223-228. [doi: [10.1016/S2352-4642\(18\)30022-1](https://doi.org/10.1016/S2352-4642(18)30022-1)] [Medline: [30169257](https://pubmed.ncbi.nlm.nih.gov/30169257/)]
33. Mayring P. Qualitative content analysis: theoretical foundation, basic procedures and software solution. In: *Approaches to Qualitative Research in Mathematics Education*: Springer; 2014. [doi: [10.1007/978-94-017-9181-6_13](https://doi.org/10.1007/978-94-017-9181-6_13)]
34. Gulec H, Smahel D. Individual and parental factors of adolescents' mHealth app use: nationally representative cross-sectional study. *JMIR Mhealth Uhealth* 2022 Dec 16;10(12):e40340. [doi: [10.2196/40340](https://doi.org/10.2196/40340)] [Medline: [36525286](https://pubmed.ncbi.nlm.nih.gov/36525286/)]
35. Mateo-Orcajada A, Vaquero-Cristóbal R, Abenza-Cano L. Gender and academic year as moderators of the efficacy of mobile app interventions to promote physical activity in adolescents: a randomized controlled trial. *Humanit Soc Sci Commun* 2023;10(1):1-15. [doi: [10.1057/s41599-023-02502-3](https://doi.org/10.1057/s41599-023-02502-3)]
36. Reid Chassiakos YL, Radesky J, Christakis D, Moreno MA, Cross C, COUNCIL ON COMMUNICATIONS AND MEDIA. Children and adolescents and digital media. *Pediatrics* 2016 Nov;138(5):e20162593. [doi: [10.1542/peds.2016-2593](https://doi.org/10.1542/peds.2016-2593)] [Medline: [27940795](https://pubmed.ncbi.nlm.nih.gov/27940795/)]
37. Wirth A, Mues A, Birtwistle E, Niklas F. Evaluating educational apps for preschoolers: differences and agreements between the assessments of experts, parents, and their children. *Comput Human Behav* 2024 Nov;160:108361. [doi: [10.1016/j.chb.2024.108361](https://doi.org/10.1016/j.chb.2024.108361)]

Abbreviations

mHealth: mobile health

SCT: social cognitive theory

SDT: self-determination theory

Edited by S Brini; submitted 16.Apr.2025; peer-reviewed by KC Wu, M Collier; accepted 05.Jan.2026; published 01.Apr.2026.

Please cite as:

Kuehn M, Wypyrsczyk L, Lanfranchi F, Balduzzi AC, Marriott H, Solera-Sanchez A, Wiskemann J, Bauer N, Karguth N, Heißerer B, Paul NW, Mongondry R, Bloch W, Rizvi K, Fridh MK, S C Ramos E, Fiuza-Luces C, Götte M, Gauß G, Spreafico F, Konda B, Stefanović M, Dreismückenbecker E, Neu MA, Faber J, FORTE Consortium

Development and Usability of the "FORTEe Get Strong" App to Promote Physical Activity and Health Awareness in Children and Adolescents With Cancer During Intensive Treatment Using an App-Based Approach: Mixed Methods Study

JMIR Serious Games 2026;14:e75653

URL: <https://games.jmir.org/2026/1/e75653>

doi: [10.2196/75653](https://doi.org/10.2196/75653)

© Mareike Kuehn, Lena Wypyrsczyk, Francesca Lanfranconi, Adriana Cristina Balduzzi, Hayley Marriott, Alba Solera-Sanchez, Joachim Wiskemann, Nikolai Bauer, Nina Karguth, Barbara Heierer, Norbert W Paul, Rodolf Mongondry, Wilhelm Bloch, Katie Rizvi, Martin Kaj Fridh, Eva S C Ramos, Carmen Fiuza-Luces, Miriam Gtte, Gabriele Gau, Filippo Spreafico, Barbara Konda, Milica Stefanovi, Elias Dreismickenbecker, Marie A Neu, Joerg Faber, FORTEe Consortium. Originally published in JMIR Serious Games (<https://games.jmir.org>), 1.Apr.2026. 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.

Usability Study of Augmented Reality Visualization Modalities on Localization Accuracy in the Head and Neck: Randomized Crossover Trial

Yao Li^{1,2}, MSc; Gijs Luijten^{3,4,5}, MSc; Christina Gsaxner^{3,5}, PhD; Kim Grunert^{1,2}, MD; Alexis Bader^{1,2,3,6}; Frank Hölzle¹, MD, DMD, PhD; Rainer Röhrig², MD, PhD; Matías de la Fuente⁷, PhD; Jan Egger^{3,4}, PhD; Kunpeng Xie^{1,2}, MSM; Behrus Hinrichs-Puladi^{1,2}, MD, DMD

¹Department of Oral and Maxillofacial Surgery, University Hospital RWTH Aachen, Pauwelsstrasse 30, Aachen, Germany

²Institute of Medical Informatics, University Hospital RWTH Aachen, Aachen, Germany

³Institute for Artificial Intelligence in Medicine (IKIM), Essen University Hospital, Essen, Germany

⁴Center for Virtual and Extended Reality in Medicine (ZvRM), Essen University Hospital, Essen, Germany

⁵Institute of Computer Graphics and Vision (ICG), Graz University of Technology, Graz, Austria

⁶Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

⁷Chair of Medical Engineering, RWTH Aachen University, Aachen, Germany

Corresponding Author:

Kunpeng Xie, MSM

Department of Oral and Maxillofacial Surgery, University Hospital RWTH Aachen, Pauwelsstrasse 30, Aachen, Germany

Abstract

Background: Augmented reality head-mounted displays could overcome the spatial dissociation between medical imaging and the surgical field, which may be particularly important in anatomically dense regions, such as the head and neck. Although many head-mounted displays offer markerless inside-out tracking at a fraction of the cost of navigation systems, their overlay accuracy with superimposition (SI) modality onto the surgical field remains limited. The virtual twin (VT), displaying holography adjacent to the surgical field, may offer a viable alternative. However, its performance is still unclear.

Objective: This study aimed to compare the accuracy and efficiency of the two visualization modalities, SI and VT, for anatomical localization in the head and neck region.

Methods: In a randomized crossover trial to compare two augmented reality visualization modalities (SI and VT), 38 participants used a HoloLens 2 to localize point, line-based, and volume-based anatomical structures on head phantoms. Their performance was evaluated with respect to accuracy, workload, time, and user experience.

Results: SI achieved significantly better point localization accuracy than VT both in absolute (mean 14.4, SD 4.2 mm vs mean 15.8, SD 5.5 mm; $P=.003$) and relative accuracy (mean 3.4, SD 2.2 mm vs mean 6.0, SD 5.0 mm; $P<.001$). In line-based structures, accuracy was comparable between SI (average surface distance [ASD], mean 23.4, SD 4.1 mm; Hausdorff distance [HD], mean 31.5, SD 7.8 mm) and VT (ASD=mean 23.0, SD 4.5 mm; $P=.51$; HD=mean 31.0, SD 7.5 mm; $P=.57$). However, SI showed significantly higher deviation than VT in volume-based structure (ASD=mean 37.1, SD 13.8 mm vs mean 34.1, SD 14.2 mm; $P=.01$; HD=mean 52.0, SD 16.8 mm vs mean 49.1, SD 15.8 mm; $P=.03$). Participants were faster with SI ($P=.02$), while workload NASA-TLX (National Aeronautics and Space Administration Task Load Index) scores did not demonstrate a significant difference ($P=.79$).

Conclusions: Given that SI did not clearly outperform VT under overlaid soft tissue and viewing challenges, VT remains a viable alternative in certain surgical scenarios where high accuracy is not required. Future research should focus on optimizing viewing angle guidance and the linkage between the anatomical target and the skin surface.

Trial Registration: German Clinical Trial Register DRKS00032835; <https://drks.de/search/en/trial/DRKS00032835>

(*JMIR Serious Games* 2026;14:e75962) doi:[10.2196/75962](https://doi.org/10.2196/75962)

KEYWORDS

mixed reality; computer-assisted surgery; visualization techniques; human-machine interface; preoperative planning

Introduction

The head and neck region contains a variety of complex anatomical structures, including numerous vital nerves, blood vessels, and organs [1]. Accurate localization of these anatomical structures is crucial in surgical practice to minimize deviation and improve outcomes [2]. Conventional medical imaging techniques, such as computed tomography (CT) and cone beam CT, as well as magnetic resonance imaging, are primarily used for diagnosis and preoperative planning [3,4]. Medical images require surgeons to mentally map medical images onto the patient's anatomy during the operation. This process demands a high level of cognitive effort, especially in the anatomically dense head and neck region, where misinterpretation could compromise the surgical accuracy and outcomes [5,6]. Surgical navigation systems (SNS) offer solutions by integrating image data into the surgical workflow. However, the limitations of the 3D display still leave the operator reliant on spatial imagination to understand complex anatomy. Furthermore, the broader adoption of SNS has been impeded by high expenses, the inherently sophisticated configurations like optical tracking cameras and reflective markers, and the possible additional radiation exposure to patients and staff [7,8]. As a result, there is still a lack of a cost-effective, intuitive, 3D interactive visualization approach that seamlessly displays the patient's medical images in the field.

Augmented reality (AR) could fill this gap by providing real-time holographic images directly within the surgical field mainly through head-mounted displays (HMDs) [7,9]. Moreover, many current AR HMDs can provide markerless inside-out tracking at a fraction of the cost of SNS and eliminate the need for additional markers [10,11]. Unlike SNS, which typically tracks the patient and instruments, this kind of HMD-based tracking focuses on aligning virtual content with the patient's anatomy to enable hologram overlay. However, the overlay or registration accuracy of many HMDs is still not as accurate as traditional SNSs with external optical tracking at the millimeter level [10]. This limitation becomes particularly critical for the superimposition (SI) visualization modality, where virtual anatomical structures need to be precisely placed on real anatomy, a process referred to as registration [12-14]. In addition, SI may introduce occlusion, as holograms can obstruct the surgeon's view of anatomy or instruments. These challenges raise concerns about the feasibility of SI as the optimal visualization modality for AR-assisted surgery, given the setup of currently available HMDs free of external tracking [15].

An alternative visualization modality is the virtual twin (VT), where the holographic representation is displayed adjacent to the physical anatomy instead of directly overlaid on the anatomy [15,16]. By avoiding overlay, VT reduces dependence on registration accuracy and eliminates occlusion.

However, the accuracy between two modalities under markerless HMD-based tracking remains unexplored. Yet, this could be important, since if SI with intrinsic markerless tracking does not show any advantage over VT, then VT would be the favored modality for certain surgical scenarios. Therefore, the aim of

this crossover randomized controlled trial (RCT) was to compare the accuracy and efficiency of the two visualization modalities, SI and VT, for anatomical localization in the head and neck region. Localization accuracy was assessed on phantom heads for clinically relevant targets, including nerve exit points, the inferior alveolar nerve, and the salivary glands. Task duration and subjective workload were evaluated as secondary endpoints.

Methods

Overview

In total, 38 participants with different professional backgrounds (dental and medical students, resident and specialist surgeons in oral and maxillofacial, oral, and plastic surgery) were recruited and performed drawings on polystyrene foam head phantoms (Model SAM, Friseurbedarf D. M. Rudolph) in a crossover RCT with SI and VT visualization modalities. The participants were asked to draw the structures on the head phantoms, wearing HoloLens 2 (HL2; Microsoft Corp). The primary endpoint was the localization accuracy of the anatomical points (0D), which encompass nerve exit points at the supraorbital, infraorbital, and mental foramina. Secondary endpoints included the delineation accuracy of the inferior alveolar nerve pathways (2D) and salivary glands (parotid and submandibular; 3D), cognitive workload, and user experience.

System Description and Implementation

The AR visualization software for the HL2 was developed in-house to display anatomical 3D models in relation to the physical anatomy of patients or phantoms. Within the application, switching between the two different visualization modalities for the 3D models was possible. In addition to the HL2 software, a pipeline processed the medical image data. This pipeline converted volumetric CT scans into 3D models optimized for interventional planning and efficient rendering on the HL2.

Based on these requirements, the planning pipeline was built to segment the structures into meshes in 3D Slicer (version 5.2.2; The Slicer Community). The structures comprising the skull, salivary glands, and nerve exit points were manually segmented from a publicly available head and neck CT dataset [17], while the inferior alveolar nerves were segmented from a nonpublic dataset from the Medical University of Graz. A head phantom mesh was scanned by the Artec Leo 3D scanner (Artec 3D) as the skin surface. Finally, all segmented anatomical structures were nonlinearly registered to the scanned skin surface.

Our AR application was developed using Unity (version 2022.3.6f1; Unity Technologies). The registration between the head phantoms and the virtual head was implemented using the Vuforia software development kit (version 10.16.5, Parametric Technology Corporation). Vuforia Engine is a cross-platform AR solution that offers a variety of tracking features, which was frequently used in research for AR registration in surgical scenarios [18-20]. The model targets (object tracking) were applied, which possibly used edge-based techniques (not revealed by Vuforia) to recognize and track objects in real-time [21]. First, the scanned head model was uploaded to the model

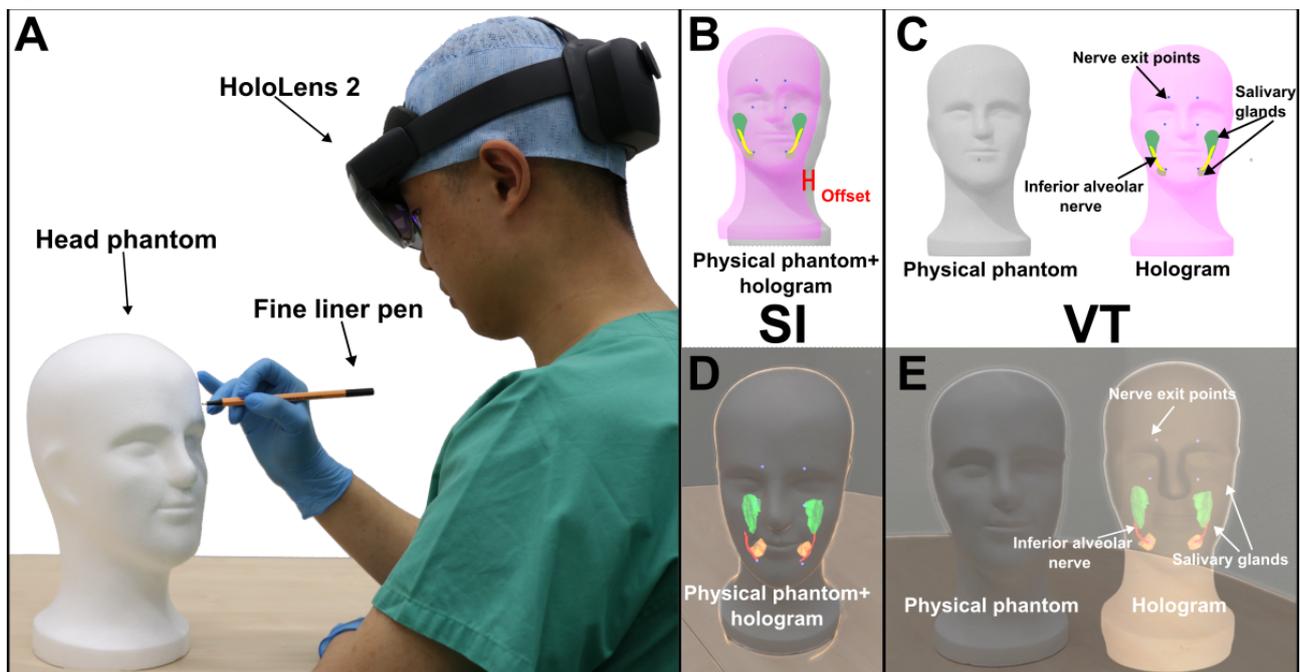
target generator tool and configured into a model target that could be integrated into Unity. After the software was deployed to the HL2, the Vuforia engine initiated tracking for target alignment. Once the participant is satisfied with the alignment, she or he could lock the tracking to anchor the virtual model in the environment. Similarly, in VT, Vuforia would track the phantom, and then the model would appear next to it; locking the tracking again would fix the model in place. The hand menu assisted users in controlling the visibility of various anatomical structures, including the skin and target structures. In addition, sliders were implemented to allow real-time adjustment of the transparency and brightness of these structures.

Trial

The participants were asked to fill out the initial questionnaire, which included demographic information (age, gender, educational stage or professional experience, professional field, and prior experience with AR and HL2). Randomization was generated by BHP using a randomized allocation rule to determine the starting modality (sequence) and the side of the face (right or left). The experiment assistant (KG) enrolled and

assigned participants to the sequence of intervention. Registration was done once at the beginning of each modality by the experiment assistant, who could lock or unlock the tracking for registration as needed. Subsequently, they wore HL2, ran the eye calibration, and received a brief introduction to the device and the user interface with the 2 modalities. During this short session, they familiarized themselves with the device and its functions. The entire familiarization process was completed in less than 3 minutes, although precise timing was not recorded. Participants were then instructed to delineate target anatomical structures on the head phantom surface using Point 88 fine liner pens (Stabilo; Figure 1). This task was performed on the assigned half of the face using the first modality, with time recorded via a stopwatch. Upon completion, participants filled out the Likert questionnaire and NASA-TLX (National Aeronautics and Space Administration Task Load Index) for that method and an open-ended questionnaire. The same procedure was then repeated on the other side of the face using the second modality, followed by the corresponding questionnaires. Finally, an open-ended questionnaire for preference was answered.

Figure 1. Illustration of two augmented reality visualization modalities using HoloLens 2. (A) Participant drawing anatomical structures (nerve exit points, inferior alveolar nerves, and salivary glands) on the polystyrene head phantom with HoloLens 2. (B) Schematic illustration of SI showing physical and holographic alignment with potential rigid offset and occlusion. (C) Schematic illustration of virtual twin showing how holograms are displayed free of misalignment and occlusion problem. (D) SI modality in HoloLens 2, where holograms were overlaid directly into the physical head phantom. (E) Virtual twin modality in HoloLens 2, where the holograms were displayed spatially adjacent to the physical head phantom. SI: superimposition; VT: virtual twin.

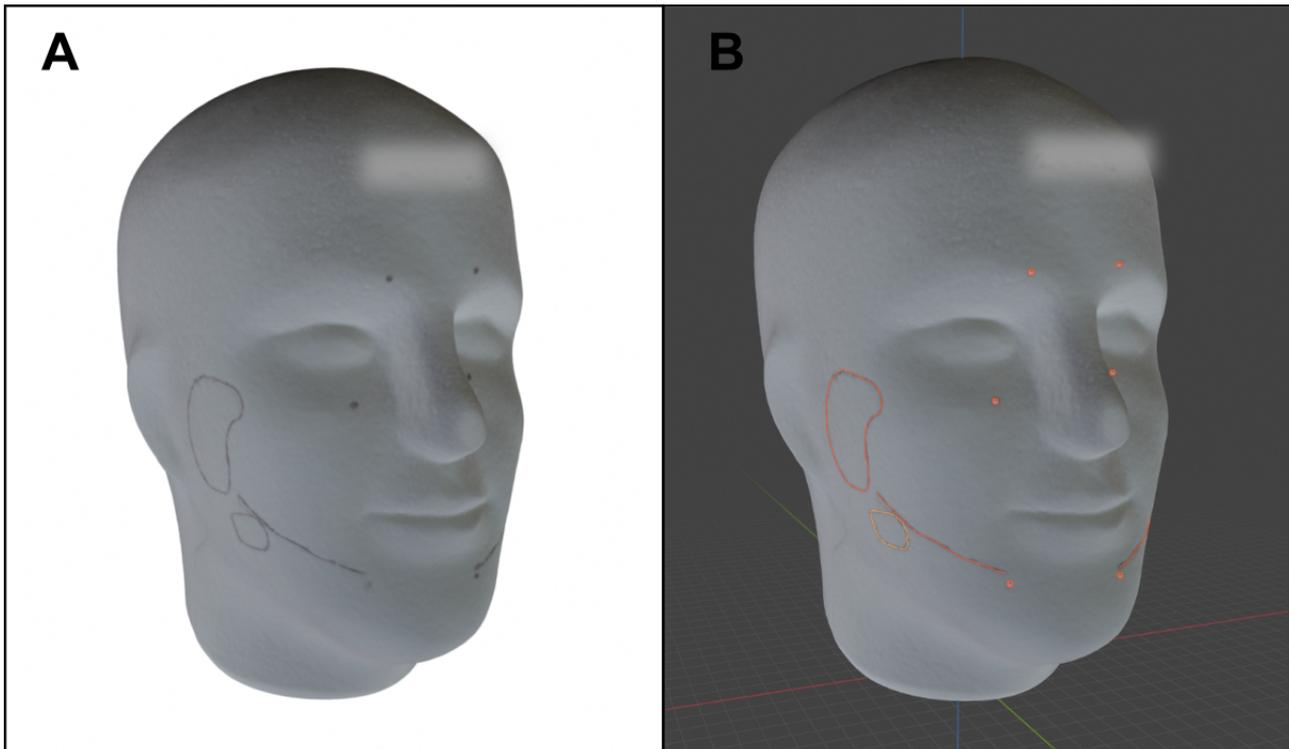


Evaluation

After the trial, all the polystyrene head phantoms were scanned with the Artec Leo 3D scanner (Figure 2). To enable comparison, all head phantoms with the participant's delineations were registered to the virtual planned head in a pipeline by a Python (version 3.10; Python Software Foundation) script. The two-stage pipeline was initiated with a global random sample consensus alignment, followed by a local refinement with point-to-plane iterative closest point, achieving <0.4 mm root mean square error. Two independent investigators (YL and

KG) evaluated the scanned heads using Blender (version 4.2; Blender Foundation). Both investigators were blinded to the applied visualization modality. To minimize a possible recall bias, KG, who served as the experiment assistant during data acquisition, underwent a washout period of 2 months before participating in the blinded evaluation. Nerve exit points were drawn by the stroke points and placed spheres. The nerve paths and salivary glands were drawn by the grease pencil tool along the curves on the head phantom surface, and the strokes were transformed into meshes in Blender (Figure 2).

Figure 2. (A) Scanned polystyrene head phantom with delineation. (B) In Blender, the scanned polystyrene head phantom is shown with inferior alveolar nerve and salivary glands annotations, and spheres marking the nerve exit points (orange).



Afterward, the points and curves were automatically compared using a Python script. The analysis for nerve exit points (OD) involved calculating Euclidean distance, which is the shortest distance in 3D space between the planned and drawn points, and we referred to this as absolute accuracy. Relative accuracy, defined as landmark-to-landmark localization accuracy, was compared by the Euclidean distance between the supraorbital-infraorbital and infraorbital-mental foramina on drawn versus planned landmarks. Since all anatomical targets were located on the underlying bone, yet localization was performed on the phantom's external surface, the concept of soft-tissue thickness was additionally introduced to capture the distance between the target structures and the skin. It was defined as the shortest distance from each anatomical point (OD) to the surface and as the mean of the vertex-to-surface distances for 2D nerve pathways and 3D salivary glands. Furthermore, the Hausdorff distance (HD) and the average surface distance (ASD) were used in order to assess the alignment and accuracy of the contours of the nerve paths (line, 2D) and salivary glands (volume, 3D). HD captures the maximum of the minimum distances between the two surfaces, providing insight into the worst-case alignment error, while the ASD quantifies the mean discrepancy, reflecting the overall degree of alignment.

The Likert questionnaire and NASA-TLX were quantitatively analyzed to assess usability and perceived workload. In addition, the feedback from open-ended questions was summarized by YL and reviewed by BHP.

Sample Size Calculation

The sample size calculation was conducted in R software (version 4.3.1; R Foundation for Statistical Computing). A minimum effect size of 5 mm was established as the threshold for an acceptable difference between the two modalities in

absolute accuracy. A 5 mm difference in absolute accuracy causes a surface discrepancy exceeding 5 mm due to the geometric relationship, making it clinically relevant and detectable by oral and maxillofacial surgeons, corresponding to the widely accepted minimum margin in head and neck oncologic surgery [22,23]. Based on the results of a pretrial with 4 participants, the mean absolute accuracy was 10.1 (SD 4.8) mm (SI) and 12.1 (SD 5.0) mm (VT) across all nerve exit points. A normal distribution of the pretest values (Shapiro-Wilk test; $P=.70$) resulted in a required number of cases of 34 for the unpaired t test. An additional 4 participants were included to compensate for nonevaluable datasets and for dropout or withdrawal of consent.

Statistical Analysis

Statistical analysis was also performed in R. A linear mixed-effects model (LMM) was applied using the *lmerTest* package [24]. This LMM assessed the absolute accuracy at the point structures, modalities (SI vs VT), the sequence (starting method), the group (dental and medical students, and surgeons), subcutaneous soft tissue thickness, and side (left or right) as fixed effects and the participants as a random effect. When analyzing the ASD and HD for line and volume-based structures, the same LMM framework was applied. Subcutaneous soft tissue thickness was specifically included to account for anatomical variation across different locations. However, it was not considered in the analysis of relative accuracy for point structures, which instead relied more on spatial reference to other anatomical landmarks.

The normality of the data distribution was assessed using the Shapiro-Wilk test. Duration and each Likert question between methods were compared using the Mann-Whitney U test. The NASA-TLX scales were compared by unpaired two-tailed t

test. For all tests mentioned, a P value of $<.05$ was considered significant.

Ethical Considerations

This study was approved by the local ethics committee of the University Hospital RWTH (Rheinisch-Westfälische Technische Hochschule) Aachen (EK 24 - 127; Chairman Prof Ralf Hausmann; April 3, 2024). The study was registered with a study protocol in advance in the German Clinical Trial Register (DRKS00032835) and followed the CONSORT (Consolidated Standards of Reporting Trials) 2010 guidelines ([Checklist 1](#)) and its extension designed and modified specifically for crossover studies, as illustrated by the flow diagram [25,26]. Informed consent was obtained from all participants involved in the study. To protect the privacy of the participants, all participants were anonymized, and no personally identifiable information was stored with the research data. It is to be noted that no financial compensation was provided to the participants involved in the present trial. Nevertheless, as a token of

appreciation, two vouchers with a total value of €15 (US \$17.5) were distributed through a raffle.

Results

Cohort

A total of 38 participants (16 females and 22 males) were successfully included in the study, comprising two groups, namely surgeons, and medical and dental students following the flow ([Figure 3](#)). Among the 18 surgeons, there were 12 residents and 6 specialists. This group included 9 oral and maxillofacial surgeons, 5 oral surgeons, and 4 plastic surgeons. In the student group, which consisted of 20 participants, 17 were dental students and 3 were medical students. The average age of participants was 26.8 (SD 5.1; range 20 - 43). The average clinical experience of surgeons was 4.0 (SD 4.3) years, and the average clinical experience of medical and dental students was 4.2 (SD 0.9) years; mean 8.3 (SD 1.6) semesters ([Table 1](#)).

Figure 3. CONSORT (Consolidated Standards of Reporting Trials) flow diagram illustrating the enrollment, allocation, crossover, follow-up, and analysis of participants in the study. SI: superimposition; VT: virtual twin.

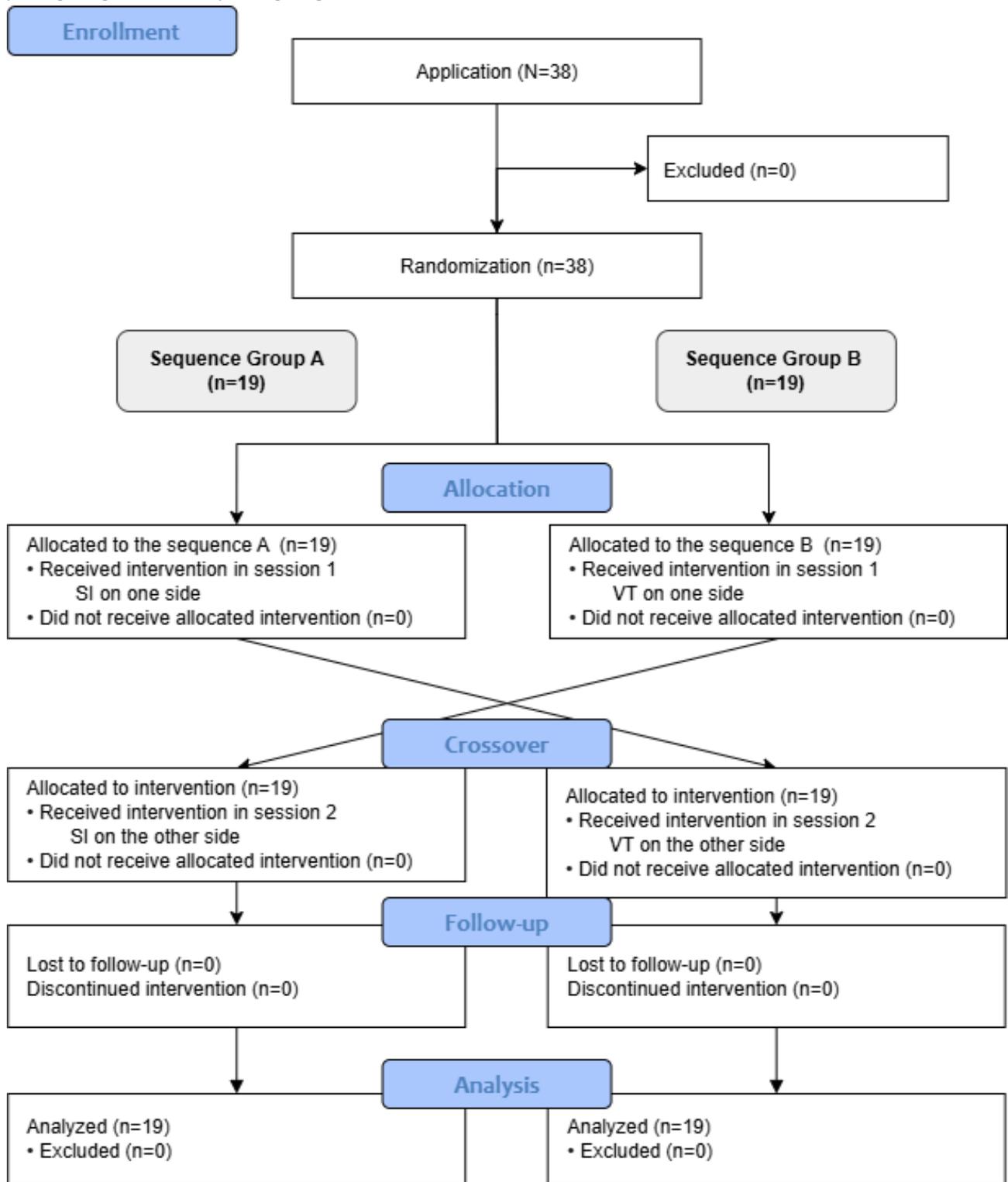


Table . Characteristics of the cohort.

Parameter	Surgeon (n=18)	Student (n=20)	Total (n=38)
Sex, n (%)			
Female	4 (22.2)	12 (60)	16 (42.1)
Male	14 (77.8)	8 (40)	22 (57.9)
Age (years)			
Mean (SD)	30.3 (5.2)	23.8 (2.5)	26.8 (5.1)
Range	23-43	20-30	20-43
Profession, n (%)			
Medical	— ^a	3 (15)	3 (7.9)
Dental student	—	17 (85)	17 (44.7)
Oral surgery	5 (27.8)	—	5 (13.2)
Oral and maxillofacial surgery	9 (50)	—	9 (23.7)
Plastic surgery	4 (22.2)	—	4 (10.5)
Clinical study/work experience (years)			
Mean (SD)	4.0 (4.3)	4.2 (0.9)	4.1 (3.0)
Range	0.0-15.0	3.0-6.0	0.0-15.0
Previous experience with AR (Likert score, 1–5) ^b			
Mean (SD)	2.4 (0.8)	2.1 (1.0)	2.2 (0.9)
Range	1.0-4.0	1.0-5.0	1.0-5.0
Previous experience with HL (Likert score, 1–5) ^c			
Mean (SD)	1.7 (0.8)	1.1 (0.3)	1.4 (0.6)
Range	1.0-3.0	1.0-2.0	1.0-3.0

^aNot applicable.

^bAR: augmented reality; Likert scores from 1="never heard of" to 5="expert."

^cHL: HoloLens; Likert scores from 1="never used" to 5="I use it several times a week."

Localization Accuracy

In the 38 scanned head phantoms, all the required structures were successfully delineated, except for 1 pair of nerve exit points at infraorbital foramina and 1 pair at supraorbital foramina, which were missed by a single participant. The absolute accuracy of the nerve exit points (0D) was significantly higher in SI (mean 14.4, SD 4.2 mm) than VT (mean 15.8, SD 5.5 mm), with a mean difference of 1.4 (95% CI 0.5 - 2.3; LMM; $P=.003$) mm. The absolute accuracy was correlated with the soft tissue thickness. For each 1 mm soft tissue thickness, the accuracy decreased by 1.4 mm ($P<.001$), while no significant difference was found in sequence ($P=.84$) and group ($P=.40$) as fixed effects in the LMM. The average participant bias was 0.8 (SD 0.8) mm. The mean absolute error of the LMM residuals was 1.8 (SD 2.9) mm for SI and 2.5 (SD 3.5) mm for VT, respectively. The relative accuracy of the points was significantly higher for SI (mean 3.4, SD 2.2 mm) than VT (mean 6.0, SD 5.0 mm) by 2.6 (95% CI 1.3 - 3.8 mm; LMM; $P<.001$; Figure 4). In Figure 4, each violin plot (colored) includes a boxplot (white), with a red dot indicating the mean

value. The black points represent the outliers. The dashed line marked the average subcutaneous soft tissue thickness over the nerve exit points.

The localization accuracy of the inferior alveolar nerve pathways (2D) assessed with ASD and HD was comparable between SI (ASD/HD=mean 23.4, SD 4.1 mm/mean 31.5, SD 7.8 mm) and VT (ASD/HD=mean 23.0, SD 4.5 mm/mean 31.0, SD 7.5 mm), with no significant difference (ASD/HD=mean difference 0.4 mm, 95% CI -1.0 to 2.0 mm; LMM; $P=.51$ /mean difference 0.6 mm, 95% CI -1.6 to 2.9 mm; LMM; $P=.57$). Regarding the salivary glands (3D), the localization accuracy measured with ASD/HD (mean 34.1, SD 14.2 mm/mean 49.1, SD 15.8 mm) for VT was significantly more accurate than SI (ASD/HD=mean 37.1, SD 13.8 mm/mean 52.0, SD 16.8 mm) by ASD 3.0 (95% CI 0.7 - 5.4 mm; LMM; $P=.01$) mm and HD 2.9 (95% CI 0.2 - 5.8 mm; LMM; $P=.03$) mm (Figure 5). In Figure 5, each violin plot (colored) includes a boxplot (white), with a red dot indicating the mean value. The black points represent the outliers. The dashed line marked the average subcutaneous soft tissue thickness over the inferior alveolar nerves and salivary glands.

Figure 4. Comparison of localization accuracy (y-axis) at nerve exit points (OD) between superimposition (purple) and virtual twin (green; x-axis). (A) Euclidean distance for absolute accuracy. (B) Absolute residual error from the linear mixed-effects model. (C) Euclidean distance for relative accuracy. (D) Relationship between Euclidean distance for absolute accuracy (y-axis) and subcutaneous soft tissue thickness (x-axis). The solid blue line depicts the fitted linear mixed-effects model regression. ED: Euclidean distance; LMM: linear mixed-effects model; SI: superimposition; VT: virtual twin.

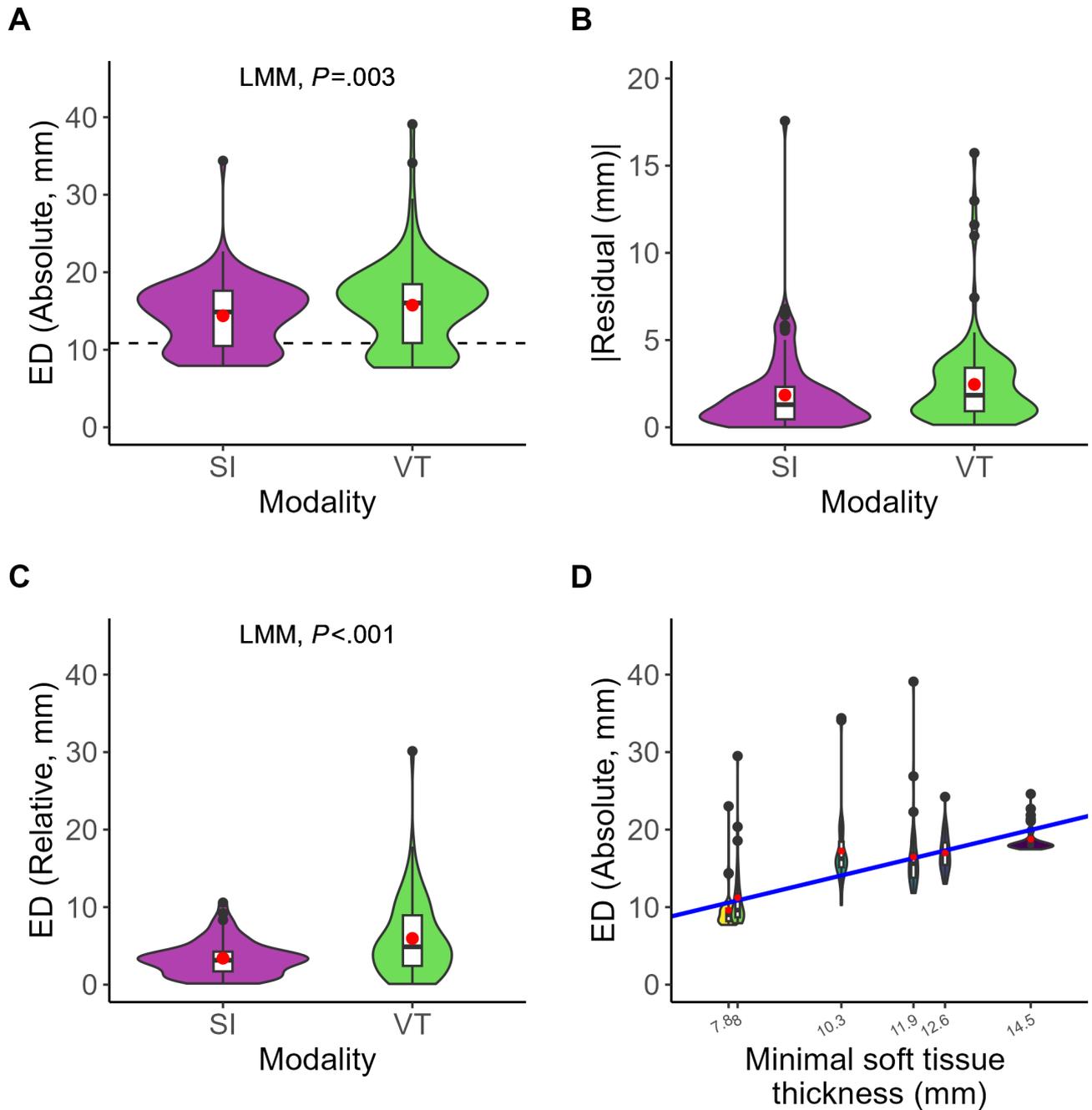
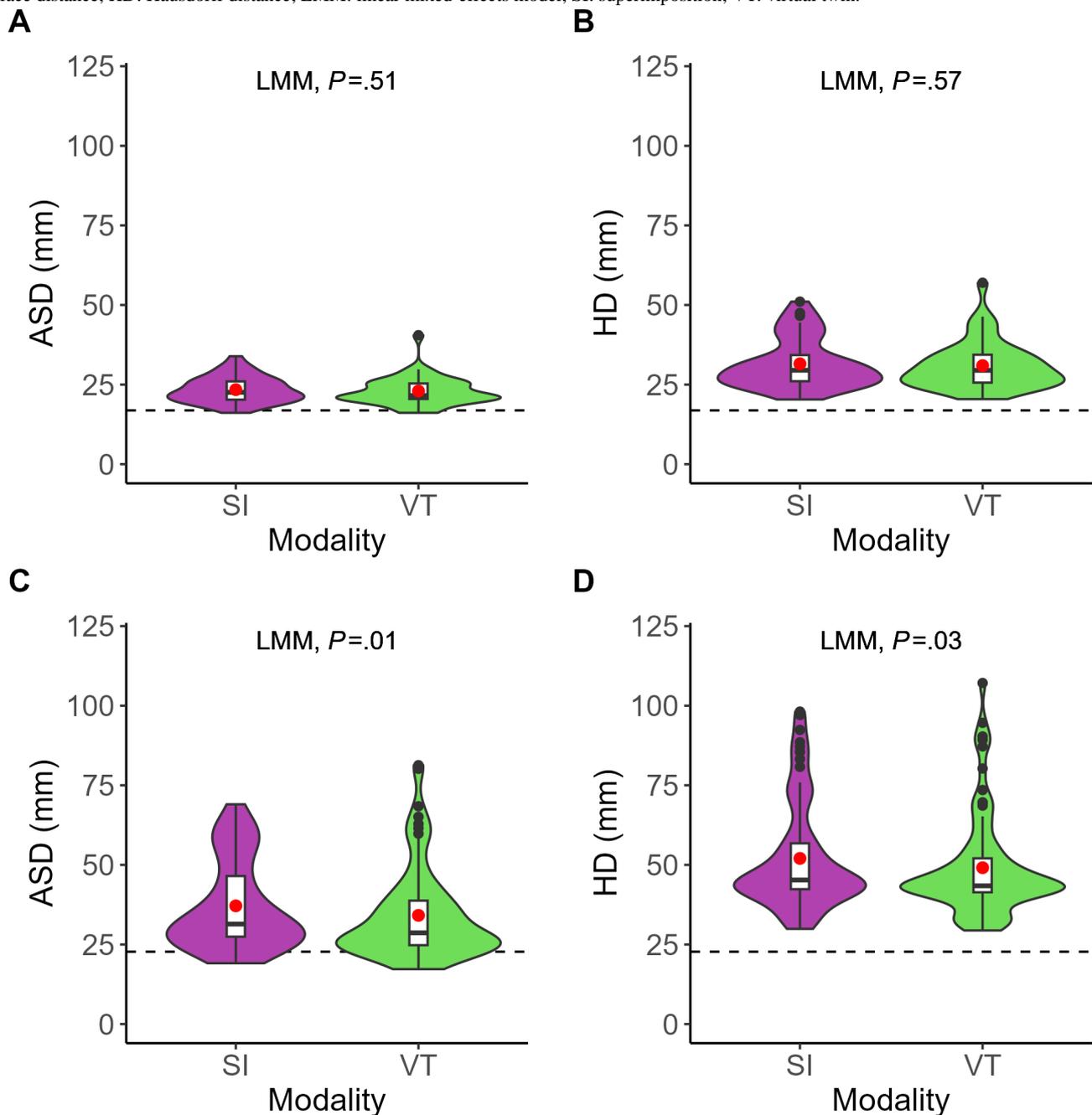


Figure 5. Comparison of localization accuracy (y-axis) for inferior alveolar nerve pathways (2D) and salivary glands (3D) between superimposition (purple) and virtual twin (green; x-axis). (A) Average surface distance for inferior alveolar nerve pathways (2D). (B) Hausdorff distance for inferior alveolar nerve pathways (2D). (C) Average surface distance for salivary glands (3D). (D) Hausdorff distance for salivary glands (3D). ASD: average surface distance; HD: Hausdorff distance; LMM: linear mixed-effects model; SI: superimposition; VT: virtual twin.

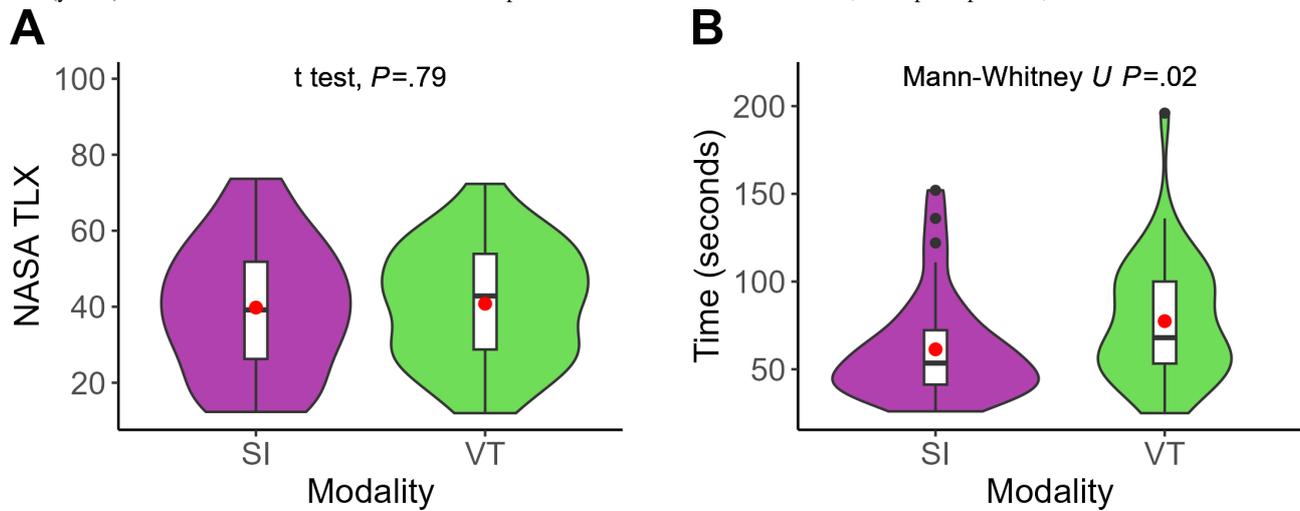


Workload and Time

The SI method (mean 61.3, SD 29.6 seconds) was significantly faster than the VT method (mean 77.4, SD 34.5 seconds) by 16.1 (95% CI 2.0 - 29.0; Mann-Whitney U test; $P=.02$) seconds. The NASA-TLX score for the SI method (mean 39.8, SD 17.3)

and VT method (mean 40.8, SD 15.2) was comparable, with no significant difference (mean difference 1.0, 95% CI -4.2 to 6.2; t test; $P=.79$; Figure 6). In Figure 6, each violin plot (colored) includes a boxplot (white), with a red dot indicating the mean value. The black points represent the outliers.

Figure 6. Subjective ratings and task completion time between superimposition (purple) and virtual twin (green) visualizations (x-axis). (A) Subjective workload assessed using NASA-TLX (National Aeronautics and Space Administration Task Load Index) scores (y-axis). (B) Task completion time in seconds (y-axis). NASA TLX: National Aeronautics and Space Administration Task Load Index; SI: superimposition; VT: virtual twin.



Questionnaires

The Likert-type questions (scale 1-4; 1=strong disagreement; 4=strong agreement) showed no significant difference (Mann-Whitney U test) between the two modalities (Table 2). The participants perceived no clear advantage in accurate localization of target structures between SI and VT (mean 3.0, SD 0.9 vs mean 3.0, SD 0.6 points; $P=.61$; mean 2.9, SD 0.8 vs mean 2.8, SD 0.7 lines; $P=.37$; mean 2.7, SD 0.8 vs mean 2.7, SD 0.7 volume; $P=.95$). Participants also reported similar levels of confidence (mean 2.7, SD 0.7 vs mean 2.7, SD 0.6;

$P=.84$), distraction (mean 2.2, SD 1.0 vs mean 1.8, SD 0.9; $P=.05$), provided assistance (mean 2.9, SD 0.8 vs mean 3.1, SD 0.6; $P=.46$), practicality (mean 2.4, SD 0.9 vs mean 2.8, SD 0.8; $P=.09$), perceived feasibility in interventions (mean 2.8, SD 1.1 vs mean 2.7, SD 0.9; $P=.57$), safety enhancement (mean 2.5, SD 0.9 vs mean 2.7, SD 0.9; $P=.19$), and overall satisfaction (mean 2.8, SD 0.9 vs mean 2.9, SD 0.8; $P=.66$). In addition, positive and negative detailed feedback was provided for both visualization modalities (Table 3). It is noteworthy that 19 participants expressed a preference for VT, 18 participants for SI, and 1 participant expressed equal preference for both.

Table . Likert questionnaire.

Likert questions	SI ^a , mean (SD)	VT ^b , mean (SD)	Total, mean (SD)	<i>P</i> value
I was able to accurately mark the nerve exit points using the (SI or VT) visualization.	3.0 (0.9)	3.0 (0.6)	3.0 (0.8)	.61
I was able to accurately mark the nerve pathways using the (SI or VT) visualization.	2.9 (0.8)	2.8 (0.7)	2.8 (0.7)	.37
I was able to accurately mark the salivary glands using the (SI or VT) visualization.	2.7 (0.8)	2.7 (0.7)	2.7 (0.7)	.95
I was sure where the anatomical structures were located and where to mark them.	2.7 (0.7)	2.7 (0.6)	2.7 (0.7)	.84
I found the using (SI or VT) visualization distracting while marking.	2.2 (1.0)	1.8 (0.9)	2.0 (1.0)	.05
The using (SI or VT) visualization facilitated the localization of anatomical structures in the face.	2.9 (0.8)	3.1 (0.6)	3.0 (0.7)	.46
I found the using (SI or VT) visualization to be practical for use.	2.4 (0.9)	2.8 (0.8)	2.6 (0.8)	.09
I could imagine performing interventions with AR support using (SI or VT) visualization.	2.8 (1.1)	2.7 (0.9)	2.7 (1.0)	.57
I believe that AR ^c support through (SI or VT) visualization enhances patient safety.	2.5 (0.9)	2.7 (0.9)	2.6 (0.9)	.19
I was generally satisfied with the AR support through the using (SI or VT) visualization.	2.8 (0.9)	2.9 (0.8)	2.8 (0.8)	.66

^aSI: superimposition.

^bVT: virtual twin.

^cAR: augmented reality.

Table . Summarized open questions.

Visualization modalities	Positive	Negative
Superimposition	<ul style="list-style-type: none"> • 3D, intuitive^{a,b} • A novelty experience^a • Accurate^{a,b} • Beginner-friendly^{a,b} • Clear and detailed^a • Contrasting colors enhance structural differentiation^b • Could be observed in all directions^b • Easy localization of structures^{a,b} • Easy to use^{a,b} • Feeling of safety^b • Free of time delay^a • Good guidance and spatial relationship^{a,b} • Inner structures could be easily seen in all directions^a • Potential to simplify the process^a • Simple design^a • Time-saving^{a,b} 	<ul style="list-style-type: none"> • Depth is perceived differently in different angles^{a,b} • Difficult to map 3D structures to 3D surface^{a,b} • Hard to identify the position of the structures^a • Hard to recognize the tip of the pen and place to draw^{a,b} • Have to lock the registration and move the head phantom to fine-tune it to the hologram^a • Inaccurate overlay, holograms are partially overlaid to the physical head phantom^{a,b} • Need familiarization time^a • Need to move the head phantom to overlay^b • Not practical^a • Relatively lacks sharpness^a • Restriction of viewpoint, cannot rotate the head phantom to observe after locking the registration^a • Some structures have merged^{a,b} • The guidance makes the user neglect the critical anatomical landmarks, causing imprecise localization^a • The hologram is blurred, and the double image is tiring^a • The position of the head phantom and the participant should be kept constant^a

Visualization modalities	Positive	Negative
Virtual twin	<ul style="list-style-type: none"> • 3D visualization, intuitive^a • Accurate^a • Assistive setup^{a,b} • Clear visualization of the anatomical structures' location^{a,b} • Direct views^a • Easy to use^{a,b} • Good guidance^{a,b} • Guidance to the targets depends on the distance to the landmarks^a • Head could be moved to draw^a • Intuitive^a • Less irritating than SI^b • Like working with a textbook on the side^a • No registration problem, 3D model hardly disturbs as an aid^a • Only exit points are good to paint and recognizable^a • Opportunity to apply to other structures^b • With improved guidance, anatomical structures can be localized more effectively, referring to reference only when necessary^{a,b} • Without overlapping, both the pen and the drawing position are clearly visible^{a,b} 	<ul style="list-style-type: none"> • Better able to rotate or zoom in^a • Better to reposition or move the model without moving yourself^a • Confusing and inaccurate^a • Deficiency of necessary landmarks^b • Hard to estimate where to draw^a • Image lacking sharpness^a • Lack of 3D guidance^a • Limited transferability to the real head^b • Little added value compared to drawing according to anatomical landmarks^a • Localization cannot be tracked as precisely as with SI^b • Longer time for eyes to adapt to^a • Need to turn around the head phantom and hard to find the correct position^a • Not practical for clinical use^a • Possible spatial discrepancy, inaccurate drawing, impractical^{a,b} • Required more cognitive effort compared to direct projection^b • Rotation of the virtual head is restricted^a • Slower in time^{a,b} • Spatial depth is hard to estimate^b • Switching attention back and forth between the head and hologram is confusing^a • The smooth white head phantom offers few points of reference, hard to transfer the anatomical structures^b

^aDental and medical student.

^bOral and maxillofacial, plastic, and oral surgeons.

Discussion

Principal Findings

We systematically evaluated the localization accuracy between two visualization modalities: SI with markerless inside-out tracking and VT for different types of anatomical structures in the head and neck region. The primary endpoint (absolute accuracy of OD structure) revealed that SI was significantly more accurate than VT by 1.4 mm ($P=.003$). In terms of relative accuracy of OD point structures, SI also outperformed VT by a margin of 2.6 mm ($P<.001$). VT showed comparable accuracy for 2D structures and notably superior accuracy (ASD, $P=.01$; HD, $P=.03$) for 3D structures, although it required an additional 16 seconds on average ($P=.02$). Likert questions revealed comparable results between two modalities. Feedback from open-ended questions (Table 3) highlighted SI for ease of understanding, intuitiveness, and time efficiency, yet noted persistent challenges with depth perception, visual occlusion, and virtual-real misalignment. Conversely, VT was perceived as simpler, clearer, and free of occlusion and misalignment issues, despite lacking direct positional cues on real head phantoms and requiring frequent attention shifts between physical and virtual models. Overall, user preferences were

evenly split, reflecting comparable experiences despite each modality's distinct strengths and limitations.

Respective Strengths and Weaknesses

In contrast to VT without tracking, the accuracy of SI depends on the inside-out tracking of the HMD used and can be attributed to 3 main factors, namely the registration accuracy of the tracking (Vuforia), the spatial mapping performance (HL2), and the visual occlusion [27]. Previous studies illustrated Vuforia software development kit's registration in the HL2 highly depended on the richness of the shape and texture of the tracked target and ranged from less than 2 mm to more than 10 mm for translational error [19,28], which can propagate into an angular deviation of the task-specific cutting plane up to 14.7° [20]. Furthermore, Vuforia tracking is sensitive to environmental light intensity, distance to the target object, and the extent of the surface covered [29]. In addition, HL2 used visual inertial-simultaneous localization and mapping (VI-SLAM) to continuously map the environment and update its position and orientation within a global coordinate system, anchoring virtual content to real-world features [30,31]. However, VI-SLAM's accuracy can be affected by factors, such as pose prediction latency, user motion, environment, and sensor fusion, such as poor integration between the red, green, blue camera and inertial

measurement unit [30,31]. This VI-SLAM error accumulated along the way, reaching 5 mm per 628 mm traveled in the clinical environment [31,32]. Moreover, the jitter latency caused by such sensor fusion could further compromise user experience, increase cognitive load, and induce fatigue [30,33]. Last but not least, visual occlusion, where virtual objects can obstruct or distort the view of the physical counterpart, further compromises the accuracy of SI. Many participants reported difficulty in identifying the position of the pen, drawn line, and occluded virtual content, which was also observed in another study [34]. This occlusion problem could lead to severe damage during surgery by overlooking anatomical structures and events [35,36]. All these factors together may contribute to the overall accuracy achieved by SI.

On the other hand, VT showed comparable accuracy (inferior in 0D, comparable in 2D, and superior in 3D structures) to SI with markerless inside-out tracking of the HMD, but without the aforementioned problems of SI. This was largely due to VT's design, which bypassed the need for precise virtual-real overlay or accurate anchoring by displaying the virtual model next to the real head phantom. Nevertheless, VT as a visualization modality free of misalignment, unaffected by occlusion, and less sensitive to spatial mapping instability could substitute SI in macro localization tasks. Since VT lacked direct positional cues to guide localization, it likely depended on the surgeon's ability to estimate distance, where surgeons performed an average error of 1.4 (SD 1.2) mm in 5 mm and 2.0 (SD 1.9) mm in 1 cm estimation in a research [37]. The distances between the nerve exit points in our study were approximately 4 cm between supraorbital and infraorbital foramina and 7 cm between infraorbital and mental foramina. If we assume the estimation error was in a linear model, this corresponded to a mean error of 5.6-9.2 mm, which aligned with VT's average relative accuracy (mean 6.0, SD 5.0 mm), inferior to SI (mean 3.4, SD 2.2) mm). Therefore, one could argue that SI is only meaningful if its accuracy exceeds the limits of human distance estimation.

Comparison to Prior Work

In scenarios where precise localization is required, such as orbital fracture reconstruction or trajectory drilling, optical tracking remains the most accurate method to date [15]. Consequently, numerous studies have adopted optical tracking to optimize registration of SI. For instance, Tu et al [38] achieved entry point accuracy of mean 2.8 (SD 1.3) mm and angular accuracy of mean 3.0° (SD 1.2°), optimizing registration accuracy to mean 2.0 (SD 0.7) mm through optical tracking. Similarly, Iqbal et al [39] combined the HL2 built-in camera with an external optical tracking camera, further reducing translation and rotation errors to 2.1 mm and 1.5°, respectively. In contrast, VT with external optical tracking could also visualize both the virtual instrument and the target anatomy but adjacent to the patient in real-time. This framework achieved higher accuracy than the aforementioned SI systems and comparable accuracy to SNS, with translational deviations of mean 0.9 (SD 0.4) mm and mean 1.0 (SD 0.5) mm at entry and end points, respectively, and a rotational deviation of mean 1.1° (SD 0.6°) [15], within the clinically feasible range (~2 mm) [12]. The noticeable difference between VT by 0.9 (SD 0.4)

mm and SI by 2.1 mm with a similar optical tracking framework likely resulted from the aforementioned factors, such as registration errors, VI-SLAM instability, jitter, and visual occlusion. This raises the question of whether SI with optical tracking should be considered the optimal AR visualization modality for surgical scenarios, particularly given that VT achieved similar accuracy under similar tracking conditions without encountering these limitations.

However, all these values assume that localization accuracy is measured in anatomically exposed structures, where perfect localization could theoretically reach 0 mm. However, this ignores a crucial aspect of real-world scenarios: anatomical structures are typically covered by tissue, which prevents direct access and inherently limits localization accuracy. In our study, this was particularly relevant due to the soft and bone tissue overlying the anatomical target structures (ie, nerve exit points, inferior alveolar nerves, and salivary glands). According to the literature, the average soft tissue thickness of the head, face, and neck is 9.4 (SD 6.2; range 2.4-28.1 mm) mm in women and 10.5 (SD 7.2; range 2.7-32.4) mm in men [40]. Thus, the measured localization accuracies observed with SI and VT for nerve exit points cannot be directly compared to classical navigation scenarios with fully exposed anatomical targets, as they are composed of 4 main influencing factors

Localization Accuracy_{ij}=β₁·Overlying Tissue Thickness_i Anatomical Constraint+

β₂·Modality_{ij} Visualization Modality (SI vs. VT)+u_j Subject Bias+ε_{ij} Residual Error (AR Noise)

First, the tissue thickness overlaid the target structure. This resulted in a decrease in localization accuracy of 1.4 mm per 1 mm of overlying tissue thickness (LMM, unstandardized coefficient β₁=1.4; *P*<.001). Then there is the influence of the VT modality, which added an additional error of 1.4 mm compared to SI (LMM, β₂=1.4; *P*=.003), and the average participant-specific bias, which was 0.8 mm (average magnitude of random intercept for individuals in the LMM). Finally, the residuals described the general pattern of localization error, with a mean absolute error of 1.8 (SD 2.9) mm for SI and 2.5 (SD 3.5) mm for VT.

For line- or volume-based structures, this correlation could not be reliably captured by the model. This was probably because localization accuracy depended not only on the viewing direction but also on dynamic changes in the perceived target margin along that direction. This contrasts with the single-point structure, which is invariably depicted as a point in all directions. As a result, the localization accuracy for line- and volume-based structures is biased by viewing direction and margin variability, in addition to the tissue thickness. As in Van Gestel et al [41], where a brain tumor was dynamically projected onto the skin along a vector from its center to the instrument tip, the participant's line of sight in our study played a comparable role to the instrument tip. As the viewing angle shifted, the visible margin of the gland changed in real time, introducing variability in the drawn curves and affecting both ASD and HD. Even for targets on the skin, like in wound area estimation using photography, variation in camera angle could introduce 10% error [42]. Although we could not directly quantify viewing

angles and changing margins, aligning the participant's gaze with the vector from the structure's centroid to its nearest skin projection may help minimize delineation errors related to such bias.

Clinical Implications

First, VT appears particularly advantageous for tasks requiring coarse localization and stable spatial orientation. VT provides a reliable anatomical context and could help mitigate cognitive errors, such as confusion of lateral sides or anatomical levels. These errors often arise in apparent symmetrical regions, especially in the absence of clear preoperative marking or adequate visual guidance. For example, in thoracolumbar spine surgery, reliance solely on intraoperative fluoroscopy may be insufficient to reliably distinguish vertebral levels, especially in the presence of anatomic variants, inadequate intraoperative imaging fields, and unreliable surface landmarks, with 50% - 67% of surgeons reporting such errors [43,44]. VT could orient the surgeon by allowing the user to align CT-based virtual models with the patient's posture, enabling clear visualization of the spine and reducing wrong-level or wrong-side misorientation. Second, in maxillofacial reconstruction, the VT technique offers significant value by displaying planned bone segments and prebent fixation plates alongside the operative field. This side-by-side visualization enables real-time comparison and intraoperative adjustment of plate bending, reducing the need for repeated fitting at the surgical site as standard techniques do [45], and thereby lowering the risk of infection. Compared with preoperative 3D printing, such a technique could also minimize fabrication time and offer greater flexibility for intraoperative adjustments. In the following free flap reconstruction procedures, VT offers robustness in environments prone to bleeding, swelling, or tissue deformation, where SI overlays can drift or become unreliable. By anchoring the virtual model generated from virtual surgical planning adjacent to the surgical site, VT provides a stable frame of reference with consistent skeletal landmarks, even when soft tissues shift [46]. Third, VT is well-suited to fractures and postoncologic defects of the orbit and midface requiring symmetry (eg, zygomatic arch, orbital floor, and medial wall) [47]. By rendering the contralateral mirrored anatomy, target orbital volume, planned implant contour, and craniofacial buttresses adjacent to the field, the surgeon could continuously compare the intraoperative reduction with the surgical plan.

While VT may help reduce orientation errors, SI demonstrates its strength in scenarios that demand high-precision localization. For example, in mandibular reconstruction surgery using the anterolateral femoral flap, accurate localization of the perforator vessels is crucial to flap viability and surgical success. One study found the SI with remote-controlled overlay (mean 3.5, SD 2.8 mm) achieved significant superior localization accuracy in anterolateral femoral perforator vessels than ultrasonic color Doppler (mean 9.6, SD 5.8 mm; $P < .001$) [48]. Our findings showed that SI had clear advantages in point-based localization tasks. This feature is particularly important in procedures such as sentinel lymph node biopsy, where accurately identifying nodes just a few millimeters beneath the surface is crucial for surgical success. Duan et al [49] reported that AR SI with motion compensation achieved sub-3 mm localization error in

melanoma sentinel lymph node biopsy. Moreover, SI demonstrated significantly superior relative accuracy ($P < .001$). This is because, despite the offset, SI preserved the spatial relationships between landmarks. The scenario that benefits from this strength is when relative distances between anatomical points must be accurately estimated, especially when a landmark has already been explored and exposed. For example, in head and neck tumor surgeries, surgeons often use the tragal pointer as a surgical landmark to identify the facial nerve trunk and the maxillary artery during procedures, such as parotidectomy, mandibular osteotomy, and temporomandibular joint arthroplasty [50]. In addition, in skull base surgery, surgeons often rely on stable bony landmarks, such as the occipital condyle or mastoid process, to sequentially locate cranial nerve exit points, including the jugular foramen and hypoglossal canal [51].

SI with markerless inside-out tracking and VT could be combined across different stages of the tasks. First, VT provides general spatial awareness, such as adapting to specific patient positioning, orienting with comprehensive medical imaging, or selecting approximate entry points. Once a key anatomical landmark is exposed, SI could rapidly guide surgeons to adjacent structures by using relative spatial relationships, minimizing the need for repeated attention switching [52,53]. If SI causes visual obstruction, cognitive overload, or registration instability and inaccuracy, SI can be temporarily deactivated, allowing VT to take over as a stable spatial reference. This hybrid modality enables adaptive assistance, providing surgeons with tailored support at different procedure stages based on clinical needs.

Our findings showed that user preferences were almost evenly split between SI and VT, underscoring the limitations of relying on either visualization method in isolation. Rather than competing alternatives, SI and VT could be viewed as complementary tools that respond to different scenarios. While SI enables precise overlay of subcutaneous landmarks, VT provides more reliable orientation under deformation or registration drift. These complementary features suggest that future AR systems should integrate both approaches within a single workflow.

Limitations

This study has some limitations. First, the polystyrene foam head phantoms used in the experiment lacked realistic features, such as skin texture, natural color, and anatomical details, which are critical for accurate identification of anatomical landmarks in the real clinical scenarios. However, using these phantoms allowed for reproducible evaluation of the performance of two modalities. Second, the homogeneous and rigid phantom surface may have favored SI by registration. Unlike real surgical environments, phantoms lack deformable soft tissues, surgical draping, fluids, and light reflections, all of which can substantially increase registration and tracking errors for SI [54,55]. In contrast, VT does not require accurate overlay; thus, it was not hindered by those problems. These considerations suggest that the relative advantage of SI observed in phantom experiments may be attenuated in vivo, whereas VT could perform more robustly in real surgical settings. Third, in real clinical scenarios, the phantom's components, such as the

mandible, could not replicate the mobility of human anatomy. This mobility may pose a significant challenge to markerless inside-out registration and further accurate anatomical localization for SI. In contrast, mobile parts in VT may be a potential solution. To address these challenges, cadaver studies or studies with high-fidelity phantoms replicating the mobility of anatomical structures should be conducted to validate the clinical applicability and generalizability of the findings. Fourth, since the difference in absolute accuracy in the sample size calculation was less than 5 mm, the study may have been underpowered to detect the influence of some fixed effects. Subsequent studies should consider increasing the sample size to enhance statistical power and generalizability. Finally, current findings are constrained to the facial region, where underlying bone structures provide a stable spatial reference. It would be valuable to investigate the performance of two modalities in other regions of the body like the abdomen, where soft tissue may bring additional challenges.

Future Directions

In addition to further validation with cadaveric studies or high-fidelity phantoms, future work should also address technical factors that directly influence localization accuracy. In particular, subcutaneous soft tissue thickness, variations in viewing perspective, and the resulting margin variability were shown to pose consistent challenges for both AR modalities. To mitigate these effects, new visualization approaches need to be developed to reduce the effects of viewing perspective and account for the effects of the overlying tissue, regardless of the visualization modality. First, the user's viewing angle could be guided in AR. One possible strategy would be to create a virtual cylindrical tunnel of 2 circles between the target structure and the skin surface, orienting the user to view in a planned

direction. Second, the AR visualization should establish a clear connection between the overlying tissue and the target structures, for example, for nerve exit points, a line connecting the points and their planned skin projection, clearly identifying the planned margin and mitigating the inaccuracy introduced by the overlying tissue.

While these approaches address specific visualization challenges, the next step lies in advancing toward a hybrid, context-aware AR system. With advances in registration accuracy, hardware performance, and integration of AI technologies, such a system could autonomously detect procedural phases, surgical context, and anatomical exposure. Based on this contextual understanding, it could dynamically switch between VT and SI modes, providing global spatial orientation and reference by VT and precise overlays for local structure localization by SI. This intelligent modality would reduce cognitive load and enable phase-specific surgical guidance.

Conclusion

This study systematically compared SI with markerless inside-out tracking and VT for surgical localization tasks in the head and neck region. SI demonstrated superior localization accuracy in 0D structures, whereas VT revealed robust spatial orientation, comparable accuracy in 2D, and superior accuracy in 3D structures. These complementary strengths suggested that VT represents a viable alternative for macro localization, while SI may be preferable for fine-grained, sequential landmark tasks. Rather than assuming SI to be universally applicable across all surgical contexts, our findings emphasize the need for context-adaptive AR strategies that can dynamically leverage the strengths of both modalities.

Acknowledgments

We sincerely thank all participating students and surgeons for their valuable contributions to this study. Open access funding provided by the Open Access Publishing Fund of RWTH (Rheinisch-Westfälische Technische Hochschule) Aachen University.

Funding

This work was supported by the REACT-EU project KITE (grant number EFRE-0801977, Plattform für KI-Translation Essen (<https://kite.ikim.nrw/>), FWF enFaced 2.0 (grant number KLI-1044 (<https://enfaced2.ikim.nrw/>), Clinician Scientist Program of the Faculty of Medicine RWTH Aachen University (BHP). CG was funded by the Advanced Research Opportunities Program (AROP) of RWTH (Rheinisch-Westfälische Technische Hochschule) Aachen University.

Data Availability

The datasets generated or analyzed during this study, as well as the Python scripts used for data processing and analysis (stored on GitHub [58]), are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: BHP, CG
Methodology: CG, YL
Software: CG, GL, YL
Validation: YL, KG, KX
Formal analysis: YL, BHP, KX
Investigation: YL, KG, GL, KX
Resources: BHP, JE, GL, FH, RR

Data curation: YL, KG, GL, KX

Writing–original draft: YL

Writing–review and editing: BHP, YL, KX, GL, CG, KG, AB, FH, RR, MdlF, JE

Visualization: YL, BHP

Supervision: BHP, JE

Project administration: BHP

Funding acquisition: CG, BHP

Conflicts of Interest

BHP is an associate editor of the *Journal of Medical Internet Research*. All other authors declare no other conflicts of interest.

Checklist 1

CONSORT checklist.

[[PDF File, 68 KB - games_v14i1e75962_app1.pdf](#)]

References

- Christison-Lagay E. Complications in head and neck surgery. *Semin Pediatr Surg* 2016 Dec;25(6):338-346. [doi: [10.1053/j.sempedsurg.2016.10.007](https://doi.org/10.1053/j.sempedsurg.2016.10.007)] [Medline: [27989359](https://pubmed.ncbi.nlm.nih.gov/27989359/)]
- Nadershah M, Salama A. Removal of parotid, submandibular, and sublingual glands. *Oral Maxillofac Surg Clin North Am* 2012 May;24(2):295-305. [doi: [10.1016/j.coms.2012.01.005](https://doi.org/10.1016/j.coms.2012.01.005)] [Medline: [22364886](https://pubmed.ncbi.nlm.nih.gov/22364886/)]
- Bachar G, Siewerdsen JH, Daly MJ, Jaffray DA, Irish JC. Image quality and localization accuracy in C-arm tomosynthesis-guided head and neck surgery. *Med Phys* 2007 Dec;34(12):4664-4677. [doi: [10.1118/1.2799492](https://doi.org/10.1118/1.2799492)] [Medline: [18196794](https://pubmed.ncbi.nlm.nih.gov/18196794/)]
- Wippold FJ. Head and neck imaging: the role of CT and MRI. *J Magn Reson Imaging* 2007 Mar;25(3):453-465. [doi: [10.1002/jmri.20838](https://doi.org/10.1002/jmri.20838)] [Medline: [17279529](https://pubmed.ncbi.nlm.nih.gov/17279529/)]
- Ishibashi M, Fujii S, Kawamoto K, et al. The ability to identify the intraparotid facial nerve for locating parotid gland lesions in comparison to other indirect landmark methods: evaluation by 3.0 T MR imaging with surface coils. *Neuroradiology* 2010 Nov;52(11):1037-1045. [doi: [10.1007/s00234-010-0718-1](https://doi.org/10.1007/s00234-010-0718-1)] [Medline: [20505928](https://pubmed.ncbi.nlm.nih.gov/20505928/)]
- Glas HH, Kraeima J, van Ooijen PMA, Spijkervet FKL, Yu L, Witjes MJH. Augmented reality visualization for image-guided surgery: a validation study using a three-dimensional printed phantom. *J Oral Maxillofac Surg* 2021 Sep;79(9):1943. [doi: [10.1016/j.joms.2021.04.001](https://doi.org/10.1016/j.joms.2021.04.001)] [Medline: [34033801](https://pubmed.ncbi.nlm.nih.gov/34033801/)]
- Remschmidt B, Rieder M, Gsaxner C, Gaessler J, Payer M, Wallner J. Augmented reality-guided apicoectomy based on maxillofacial CBCT scans. *Diagnostics (Basel)* 2023 Sep 25;13(19):3037. [doi: [10.3390/diagnostics13193037](https://doi.org/10.3390/diagnostics13193037)] [Medline: [37835780](https://pubmed.ncbi.nlm.nih.gov/37835780/)]
- Landaeta-Quinones CG, Hernandez N, Zarroug NK. Computer-assisted surgery: applications in dentistry and oral and maxillofacial surgery. *Dent Clin North Am* 2018 Jul;62(3):403-420. [doi: [10.1016/j.cden.2018.03.009](https://doi.org/10.1016/j.cden.2018.03.009)] [Medline: [29903558](https://pubmed.ncbi.nlm.nih.gov/29903558/)]
- Karnatz N, Möllmann HL, Wilkat M, Parviz A, Rana M. Advances and innovations in ablative head and neck oncologic surgery using mixed reality technologies in personalized medicine. *J Clin Med* 2022 Aug 16;11(16):4767. [doi: [10.3390/jcm11164767](https://doi.org/10.3390/jcm11164767)] [Medline: [36013006](https://pubmed.ncbi.nlm.nih.gov/36013006/)]
- Gsaxner C, Pepe A, Wallner J, Schmalstieg D, Egger J. Markerless image-to-face registration for untethered augmented reality in head and neck surgery. : Springer; 2019 Presented at: Medical Image Computing and Computer Assisted Intervention (MICCAI 2019); Oct 13-17, 2019; Cham, Switzerland p. 236-244. [doi: [10.1007/978-3-030-32254-0](https://doi.org/10.1007/978-3-030-32254-0)]
- Pepe A, Trotta GF, Mohr-Ziak P, et al. A marker-less registration approach for mixed reality-aided maxillofacial surgery: a pilot evaluation. *J Digit Imaging* 2019 Dec;32(6):1008-1018. [doi: [10.1007/s10278-019-00272-6](https://doi.org/10.1007/s10278-019-00272-6)] [Medline: [31485953](https://pubmed.ncbi.nlm.nih.gov/31485953/)]
- Andrews CM, Henry AB, Soriano IM, Southworth MK, Silva JR. Registration techniques for clinical applications of three-dimensional augmented reality devices. *IEEE J Transl Eng Health Med* 2021;9:4900214. [doi: [10.1109/JTEHM.2020.3045642](https://doi.org/10.1109/JTEHM.2020.3045642)] [Medline: [33489483](https://pubmed.ncbi.nlm.nih.gov/33489483/)]
- Pérez-Pachón L, Poyade M, Lowe T, Gröning F. Image overlay surgery based on augmented reality: a systematic review. *Adv Exp Med Biol* 2020;1260:175-195. [doi: [10.1007/978-3-030-47483-6_10](https://doi.org/10.1007/978-3-030-47483-6_10)] [Medline: [33211313](https://pubmed.ncbi.nlm.nih.gov/33211313/)]
- Rose AS, Kim H, Fuchs H, Frahm JM. Development of augmented-reality applications in otolaryngology-head and neck surgery. *Laryngoscope* 2019 Oct;129(S3):S1-S11. [doi: [10.1002/lary.28098](https://doi.org/10.1002/lary.28098)] [Medline: [31260127](https://pubmed.ncbi.nlm.nih.gov/31260127/)]
- Li Y, Drobinsky S, Becker P, et al. Accuracy and efficiency of drilling trajectories with augmented reality versus conventional navigation randomized crossover trial. *NPJ Digit Med* 2024 Nov 10;7(1):316. [doi: [10.1038/s41746-024-01314-2](https://doi.org/10.1038/s41746-024-01314-2)] [Medline: [39523443](https://pubmed.ncbi.nlm.nih.gov/39523443/)]
- Wolf J, Luchmann D, Lohmeyer Q, Farshad M, Fürnstahl P, Meboldt M. How different augmented reality visualizations for drilling affect trajectory deviation, visual attention, and user experience. *Int J Comput Assist Radiol Surg* 2023 Aug;18(8):1363-1371. [doi: [10.1007/s11548-022-02819-5](https://doi.org/10.1007/s11548-022-02819-5)] [Medline: [36808552](https://pubmed.ncbi.nlm.nih.gov/36808552/)]

17. Grossberg A, Mohamed A, Elhalawani H, et al. HNSCC. The Cancer Imaging Archive. 2020. URL: <https://stage.cancerimagingarchive.net/collection/hnsc/> [accessed 2026-02-02]
18. Thabit A, Benmahdjoub M, van Veelen MLC, Niessen WJ, Wolvius EB, van Walsum T. Augmented reality navigation for minimally invasive craniostylosis surgery: a phantom study. *Int J CARS* 2022 Aug;17(8):1453-1460. [doi: [10.1007/s11548-022-02634-y](https://doi.org/10.1007/s11548-022-02634-y)] [Medline: [35507209](https://pubmed.ncbi.nlm.nih.gov/35507209/)]
19. Frantz T, Jansen B, Duerinck J, Vandemeulebroucke J. Augmenting Microsoft's HoloLens with vuforia tracking for neuronavigation. *Healthc Technol Lett* 2018 Oct;5(5):221-225. [doi: [10.1049/htl.2018.5079](https://doi.org/10.1049/htl.2018.5079)] [Medline: [30464854](https://pubmed.ncbi.nlm.nih.gov/30464854/)]
20. Chui ECS, Mak KKK, Ng RHT, et al. Application of image recognition-based tracker-less augmented reality navigation system in a series of sawbone trials. *Arthroplasty* 2024 Aug 2;6(1):39. [doi: [10.1186/s42836-024-00263-1](https://doi.org/10.1186/s42836-024-00263-1)] [Medline: [39090719](https://pubmed.ncbi.nlm.nih.gov/39090719/)]
21. Vorobiov V. On 3D pose estimation for XR. classic approaches vs neural networks [Bachelor's thesis]. : Ukrainian Catholic University; 2021 URL: <https://er.ucu.edu.ua/items/3696ed7e-49d4-45c4-bba4-be06fd4320eb> [accessed 2025-12-03]
22. Palla B, Callahan N. Does the use of computer-assisted surgery affect the margin status in resections of ameloblastoma? *J Oral Maxillofac Surg* 2021 Jul;79(7):1467-1473. [doi: [10.1016/j.joms.2020.09.007](https://doi.org/10.1016/j.joms.2020.09.007)] [Medline: [33031772](https://pubmed.ncbi.nlm.nih.gov/33031772/)]
23. Fowler J, Campanile Y, Warner A, et al. Surgical margins of the oral cavity: is 5 mm really necessary? *J Otolaryngol Head Neck Surg* 2022 Oct 4;51(1):38. [doi: [10.1186/s40463-022-00584-8](https://doi.org/10.1186/s40463-022-00584-8)] [Medline: [36195903](https://pubmed.ncbi.nlm.nih.gov/36195903/)]
24. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Soft* 2017;82(13):1-26. [doi: [10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13)]
25. Dwan K, Li T, Altman DG, Elbourne D. CONSORT 2010 statement: extension to randomised crossover trials. *BMJ* 2019 Jul 31;366:14378. [doi: [10.1136/bmj.14378](https://doi.org/10.1136/bmj.14378)] [Medline: [31366597](https://pubmed.ncbi.nlm.nih.gov/31366597/)]
26. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010 Mar 23;340:c332. [doi: [10.1136/bmj.c332](https://doi.org/10.1136/bmj.c332)] [Medline: [20332509](https://pubmed.ncbi.nlm.nih.gov/20332509/)]
27. Doughty M, Ghugre NR, Wright GA. Augmenting performance: a systematic review of optical see-through head-mounted displays in surgery. *J Imaging* 2022 Jul 20;8(7):203. [doi: [10.3390/jimaging8070203](https://doi.org/10.3390/jimaging8070203)] [Medline: [35877647](https://pubmed.ncbi.nlm.nih.gov/35877647/)]
28. Allen DR, Peters TM, Chen ECS. A co-calibration framework for the accuracy assessment of vision-based tracking systems. In: Linte CA, Siewerdsen JH, editors. Presented at: SPIE Medical Imaging 2022: Image-Guided Procedures, Robotic Interventions, and Modeling; Feb 20-24, 2022; San Diego, United States p. 598-608. [doi: [10.1117/12.2606815](https://doi.org/10.1117/12.2606815)]
29. Sendari S, Firmansah A. Performance analysis of augmented reality based on Vuforia using 3D marker detection. Presented at: 2020 4th International Conference on Vocational Education and Training (ICOVET); Sep 19, 2020; Malang, Indonesia p. 294-298 URL: <https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=9229480> [accessed 2025-12-03] [doi: [10.1109/ICOVET50258.2020.9230276](https://doi.org/10.1109/ICOVET50258.2020.9230276)]
30. Wilmott JP, Erkelens IM, Murdison TS, Rio KW. Perceptibility of jitter in augmented reality head-mounted displays. Presented at: 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); Oct 17-21, 2022; Singapore, Singapore p. 470-478 URL: <https://ieeexplore.ieee.org/document/9995228> [accessed 2025-12-03] [doi: [10.1109/ISMAR55827.2022.00063](https://doi.org/10.1109/ISMAR55827.2022.00063)]
31. Feigl T, Porada A, Steiner S, Löffler C, Mutschler C, Philippsen M. Localization limitations of ARCore, ARKit, and HoloLens in dynamic large-scale industry environments. Presented at: 15th International Conference on Computer Graphics Theory and Applications; Feb 27-29, 2020; Valletta, Malta p. 307-318. [doi: [10.5220/0008989903070318](https://doi.org/10.5220/0008989903070318)]
32. Matyash I, Kutzner R, Neumuth T, Rockstroh M. Accuracy measurement of HoloLens2 IMUs in medical environments. *Curr Dir Biomed Eng* 2021 Oct 1;7(2):633-636. [doi: [10.1515/cdbme-2021-2161](https://doi.org/10.1515/cdbme-2021-2161)]
33. Gsaxner C, Li J, Pepe A, Schmalstieg D, Egger J. Inside-out instrument tracking for surgical navigation in augmented reality. Presented at: VRST '21; Dec 8-10, 2021; Osaka, Japan p. 1-11 URL: <https://dl.acm.org/doi/proceedings/10.1145/3489849> [accessed 2025-12-03] [doi: [10.1145/3489849.3489863](https://doi.org/10.1145/3489849.3489863)]
34. Koyachi M, Sugahara K, Tachizawa K, et al. Mixed-reality and computer-aided design/computer-aided manufacturing technology for mandibular reconstruction: a case description. *Quant Imaging Med Surg* 2023 Jun 1;13(6):4050-4056. [doi: [10.21037/qims-22-1118](https://doi.org/10.21037/qims-22-1118)] [Medline: [37284085](https://pubmed.ncbi.nlm.nih.gov/37284085/)]
35. Dixon BJ, Daly MJ, Chan H, Vescan AD, Witterick IJ, Irish JC. Surgeons blinded by enhanced navigation: the effect of augmented reality on attention. *Surg Endosc* 2013 Feb;27(2):454-461. [doi: [10.1007/s00464-012-2457-3](https://doi.org/10.1007/s00464-012-2457-3)] [Medline: [22833264](https://pubmed.ncbi.nlm.nih.gov/22833264/)]
36. Dixon BJ, Daly MJ, Chan HHL, Vescan A, Witterick IJ, Irish JC. Inattentive blindness increased with augmented reality surgical navigation. *Am J Rhinol Allergy* 2014;28(5):433-437. [doi: [10.2500/ajra.2014.28.4067](https://doi.org/10.2500/ajra.2014.28.4067)] [Medline: [25198032](https://pubmed.ncbi.nlm.nih.gov/25198032/)]
37. Conway RG, O'Neill N, Brown J, Kavic S. An educated guess - distance estimation by surgeons. *Surg Open Sci* 2020 Jul;2(3):113-116. [doi: [10.1016/j.sopen.2020.04.001](https://doi.org/10.1016/j.sopen.2020.04.001)] [Medline: [33981983](https://pubmed.ncbi.nlm.nih.gov/33981983/)]
38. Tu P, Wang H, Joskowicz L, Chen X. A multi-view interactive virtual-physical registration method for mixed reality based surgical navigation in pelvic and acetabular fracture fixation. *Int J Comput Assist Radiol Surg* 2023 Sep;18(9):1715-1724. [doi: [10.1007/s11548-023-02884-4](https://doi.org/10.1007/s11548-023-02884-4)] [Medline: [37031310](https://pubmed.ncbi.nlm.nih.gov/37031310/)]
39. Iqbal H, Rodriguez y Baena F. Semi-automatic infrared calibration for augmented reality systems in surgery. Presented at: Proceedings of the 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); Oct 23-27, 2022; Kyoto, Japan p. 4957-4964. [doi: [10.1109/IROS47612.2022.9982215](https://doi.org/10.1109/IROS47612.2022.9982215)]

40. Yang W, He RK, Chen T, Haining W, Goossens R, Huysmans T. Soft tissue thickness estimation for head, face, and neck from CT data for product design purposes. In: Rebelo F, editor. *Ergonomics In Design: AHFE International*; 2022:558-565. [doi: [10.54941/ahfe1001983](https://doi.org/10.54941/ahfe1001983)]
41. Van Gestel F, Frantz T, Buyck F, et al. Neuro-oncological augmented reality planning for intracranial tumor resection. *Front Neurol* 2023;14:1104571. [doi: [10.3389/fneur.2023.1104571](https://doi.org/10.3389/fneur.2023.1104571)] [Medline: [36998774](https://pubmed.ncbi.nlm.nih.gov/36998774/)]
42. Chang AC, Dearman B, Greenwood JE. A comparison of wound area measurement techniques: visitrak versus photography. *Eplasty* 2011 Apr 18;11:e18. [Medline: [21559060](https://pubmed.ncbi.nlm.nih.gov/21559060/)]
43. Epstein N. A perspective on wrong level, wrong side, and wrong site spine surgery. *Surg Neurol Int* 2021;12:286. [doi: [10.25259/SNI_402_2021](https://doi.org/10.25259/SNI_402_2021)] [Medline: [34221617](https://pubmed.ncbi.nlm.nih.gov/34221617/)]
44. Shah M, Halalmeh DR, Sandio A, Tubbs RS, Moisi MD. Anatomical variations that can lead to spine surgery at the wrong level: part III lumbosacral. *Cureus* 2020 Jul 28;12(7):e9433. [doi: [10.7759/cureus.9433](https://doi.org/10.7759/cureus.9433)] [Medline: [32864257](https://pubmed.ncbi.nlm.nih.gov/32864257/)]
45. Wilde F, Winter K, Kletsch K, Lorenz K, Schramm A. Mandible reconstruction using patient-specific pre-bent reconstruction plates: comparison of standard and transfer key methods. *Int J Comput Assist Radiol Surg* 2015 Feb;10(2):129-140. [doi: [10.1007/s11548-014-1065-1](https://doi.org/10.1007/s11548-014-1065-1)] [Medline: [24810111](https://pubmed.ncbi.nlm.nih.gov/24810111/)]
46. Metzler P, Geiger EJ, Alcon A, Ma X, Steinbacher DM. Three-dimensional virtual surgery accuracy for free fibula mandibular reconstruction: planned versus actual results. *J Oral Maxillofac Surg* 2014 Dec;72(12):2601-2612. [doi: [10.1016/j.joms.2014.07.024](https://doi.org/10.1016/j.joms.2014.07.024)] [Medline: [25315311](https://pubmed.ncbi.nlm.nih.gov/25315311/)]
47. Khaqani MS, Tavosi F, Gholami M, Eftekharian HR, Khojastepour L. Analysis of facial symmetry after zygomatic bone fracture management. *J Oral Maxillofac Surg* 2018 Mar;76(3):595-604. [doi: [10.1016/j.joms.2017.10.005](https://doi.org/10.1016/j.joms.2017.10.005)] [Medline: [29121488](https://pubmed.ncbi.nlm.nih.gov/29121488/)]
48. Wei B, Lu G, Bai Z, et al. Augmented reality in preoperative anterolateral thigh flap perforators positioning: a pilot diagnostic study. *Oral Oncol* 2025 Mar;162:107189. [doi: [10.1016/j.oraloncology.2025.107189](https://doi.org/10.1016/j.oraloncology.2025.107189)] [Medline: [39954606](https://pubmed.ncbi.nlm.nih.gov/39954606/)]
49. Duan H, Yang Y, Niu WL, et al. Localization of sentinel lymph nodes using augmented-reality system: a cadaveric feasibility study. *Eur J Nucl Med Mol Imaging* 2025 Aug;52(10):3643-3652. [doi: [10.1007/s00259-025-07216-z](https://doi.org/10.1007/s00259-025-07216-z)] [Medline: [40123008](https://pubmed.ncbi.nlm.nih.gov/40123008/)]
50. Muhleman MA, Wartmann CT, Hage R, et al. A review of the tragal pointer: anatomy and its importance as a landmark in surgical procedures. *Folia Morphol (Warsz)* 2012 May;71(2):59-64. [Medline: [22648581](https://pubmed.ncbi.nlm.nih.gov/22648581/)]
51. Thintharua P, Chentanez V. Morphological analysis and morphometry of the occipital condyle and its relationship to the foramen magnum, jugular foramen, and hypoglossal canal: implications for craniovertebral junction surgery. *Anat Cell Biol* 2023 Mar 31;56(1):61-68. [doi: [10.5115/acb.22.105](https://doi.org/10.5115/acb.22.105)] [Medline: [36635090](https://pubmed.ncbi.nlm.nih.gov/36635090/)]
52. Liebmann F, Roner S, von Atzigen M, et al. Pedicle screw navigation using surface digitization on the Microsoft HoloLens. *Int J Comput Assist Radiol Surg* 2019 Jul;14(7):1157-1165. [doi: [10.1007/s11548-019-01973-7](https://doi.org/10.1007/s11548-019-01973-7)] [Medline: [30993519](https://pubmed.ncbi.nlm.nih.gov/30993519/)]
53. Pietruski P, Majak M, Świątek-Najwer E, et al. Supporting mandibular resection with intraoperative navigation utilizing augmented reality technology - a proof of concept study. *J Craniomaxillofac Surg* 2019 Jun;47(6):854-859. [doi: [10.1016/j.jcms.2019.03.004](https://doi.org/10.1016/j.jcms.2019.03.004)] [Medline: [30914226](https://pubmed.ncbi.nlm.nih.gov/30914226/)]
54. Schlueter-Brust K, Henckel J, Katinakis F, et al. Augmented-reality-assisted K-wire placement for glenoid component positioning in reversed shoulder arthroplasty: a proof-of-concept study. *J Pers Med* 2021 Aug 10;11(8):777. [doi: [10.3390/jpm11080777](https://doi.org/10.3390/jpm11080777)] [Medline: [34442421](https://pubmed.ncbi.nlm.nih.gov/34442421/)]
55. Suter D, Hodel S, Liebmann F, Fürnstahl P, Farshad M. Factors affecting augmented reality head-mounted device performance in real OR. *Eur Spine J* 2023 Oct;32(10):3425-3433. [doi: [10.1007/s00586-023-07826-x](https://doi.org/10.1007/s00586-023-07826-x)] [Medline: [37552327](https://pubmed.ncbi.nlm.nih.gov/37552327/)]
56. AR-VIZ [accessible by request]. GitHub. URL: <https://github.com/OMFSdigital/AR-VIZ> [accessed 2025-12-24]

Abbreviations

- AR:** augmented reality
- ASD:** average surface distance
- CONSORT:** Consolidated Standards of Reporting Trials
- CT:** computed tomography
- HD:** Hausdorff distance
- HL2:** HoloLens 2
- LMM:** linear mixed-effects model
- NASA-TLX:** National Aeronautics and Space Administration Task Load Index
- RWTH:** Rheinisch-Westfälische Technische Hochschule
- SNS:** surgical navigation systems
- VI-SLAM:** visual inertial-simultaneous localization and mapping
- VT:** virtual twin

Edited by A Coristine; submitted 22.Apr.2025; peer-reviewed by E Berglund, Y Yang; revised version received 21.Oct.2025; accepted 21.Oct.2025; published 13.Jan.2026.

Please cite as:

*Li Y, Luijten G, Gsaxner C, Grunert K, Bader A, Hölzle F, Röhrig R, de la Fuente M, Egger J, Xie K, Hinrichs-Puladi B
Usability Study of Augmented Reality Visualization Modalities on Localization Accuracy in the Head and Neck: Randomized Crossover Trial*

JMIR Serious Games 2026;14:e75962

URL: <https://games.jmir.org/2026/1/e75962>

doi: [10.2196/75962](https://doi.org/10.2196/75962)

© Yao Li, Gijs Luijten, Christina Gsaxner, Kim Grunert, Alexis Bader, Frank Hölzle, Rainer Röhrig, Matías de la Fuente, Jan Egger, Kunpeng Xie, Behrus Hinrichs-Puladi. Originally published in JMIR Serious Games (<https://games.jmir.org>), 13.Jan.2026. 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.

Development of a Novel Dietary Assessment Method Using Gamification Concepts: Exploratory and Application Study

Yulin Zhang¹, MSc; Tianwu Lin², MSc; Yuefeng Tan¹, MSc; Xiaojin Shi¹, MSc; Yucheng Yang¹, MSc; Haibing Yang¹, MSc; Xiaona Na¹, PhD; Junhan Zhang¹, MSc; Ai Zhao¹, PhD

¹Vanke School of Public Health, Tsinghua University, Jinchunyuan Building, Tsinghua University, 30th Shuang Qing Road, Hadian District, Beijing, China

²Shenzhen Yantian Maternal and Child Health Hospital, Shenzhen, Guangdong, China

Corresponding Author:

Ai Zhao, PhD

Vanke School of Public Health, Tsinghua University, Jinchunyuan Building, Tsinghua University, 30th Shuang Qing Road, Hadian District, Beijing, China

Abstract

Background: Childhood and adolescent malnutrition, encompassing undernutrition and overnutrition, poses significant global health challenges, necessitating comprehensive dietary assessment tools. Existing dietary assessment methods, such as 24-hour dietary recalls (24HR), often fail to capture eating behaviors and food preferences.

Objective: This study aims to compare a newly developed gamified dietary assessment tool (GDA) with a traditional 24HR and to explore its applicability in assessing dietary behavior tendencies among children and adolescents.

Methods: A 2-phase study was designed, including an exploratory and an application study. The exploratory study included 30 school-aged participants, comparing the GDA with the 3-day 24HR. Nutrient and food intakes were analyzed using Pearson or Spearman correlation coefficients and Bland-Altman plots. The application study, conducted among 1541 adolescents (11 - 18 y), assessed dietary intake and eating behavior tendencies. Differences in dietary intake across age, gender, socioeconomic status, and weight status were analyzed using the Kruskal-Wallis rank sum test. Multiple linear regression models were used to examine the associations of dietary intakes with emotional eating and with dining environment tendencies, respectively.

Results: In the exploratory study (n=30), the GDA demonstrated moderate agreement with 24HR for energy ($r=0.46$, $P=.01$) and carbohydrates ($r=0.50$, $P=.005$). Bland-Altman plots indicated good agreement for energy and carbohydrate intake between methods (mean differences around 0). For fat intake, although the mean difference was close to 0, the correlation was not statistically significant. In contrast, the GDA substantially overestimated protein intake (mean difference around 25 g). In the application study (n=1541), higher emotional eating scores were associated with higher snack consumption ($\beta=0.438$, 95% CI 0.035-0.840), and with lower protein ($\beta=-0.159$, 95% CI -0.267 to -0.052), fruit ($\beta=-0.464$, 95% CI -0.854 to -0.073), and nut consumption ($\beta=-0.183$, 95% CI -0.304 to -0.062). Participants who chose solitary screen eating consumed significantly more carbohydrates than those who selected “eat with peers” ($\beta=4.2$, 95% CI 1.2-7.1).

Conclusions: This study demonstrates that the GDA effectively captures both dietary intake and contextual eating behaviors in young populations, providing data distinct from traditional methods, such as 24HR. As a complementary assessment approach, it offers valuable insights into food preferences and eating patterns through its interactive design.

(JMIR Serious Games 2026;14:e72387) doi:[10.2196/72387](https://doi.org/10.2196/72387)

KEYWORDS

dietary assessment; gamification; dietary intake; food preference; eating behavior tendencies

Introduction

Children and adolescents have been facing a growing burden of malnutrition worldwide [1]. It is estimated that more than 390 million children and adolescents around the world are overweight, of whom 41% are obese [2]. At the same time, in low- and middle-income countries, approximately 16% of adolescents experience stunting and thinness [3], while micronutrient deficiencies have increasingly become a major

threat to their development [4]. Malnutrition during childhood and adolescence can lead to substantial long-term health consequences in adulthood. Children in the highest BMI quartile could have more than twice the rate of death from endogenous causes compared to those in the lowest quartile [5], and obesity in adolescence also predicts a higher risk of type 2 diabetes [6] and cardiovascular disease [7] in adulthood. Also, early undernutrition has long-lasting effects on energy expenditure, lipid metabolism, insulin resistance, and mental health, which

may even be passed to the next generation [8]. Addressing these issues requires a better understanding of dietary patterns and the factors that influence food choices among children and adolescents. Traditional and emerging dietary assessment methods are both utilized to evaluate nutritional intake. Traditional methods, including retrospective methods (eg, 24-hour dietary recalls [24HR]) and prospective methods (eg, weighed food records), have long been used in research and clinical practice to capture individuals' food consumption [9]. Meanwhile, emerging technologies, including mobile apps [10], wearable devices [11], and virtual reality [12,13], are increasingly used to assist in dietary surveys. For retrospective methods, the main challenge is their reliance on self-reporting, which can be prone to bias, especially in younger populations. Young children have limited ability to recall, estimate, and cooperate in dietary investigations [14], and their diets and eating habits tend to be more variable compared to those of adults [15]. Prospective methods are often challenging to implement due to high labor costs and also rely on respondents' self-reports. Their food choices may be influenced by their tendency to report socially desirable food consumption [16], while novel methods, such as mobile apps, often underestimate the intake of energy or specific nutrients [10,17]. Additionally, all these methods were unable to capture the broader contextual dietary factors, such as eating environment and food preferences, that influence dietary behaviors in adolescents. Particularly, children's diets are typically shaped by their families or school environment [18], which may obscure their food preferences, including adverse tendencies toward unhealthy eating habits. It is reported that children would select different foods when they were away from their parents compared to when they were watched [19]. When food is freely available and parents are absent, they always tend to choose more restricted foods, often considered unhealthy, over unrestricted options, which may contribute more to their health and development [20].

To address these limitations, a combination of gamification and dietary assessment could be a potential solution. Gamification, defined as the use of game design elements in nongame contexts [21], is an emerging interactive digital strategy that has gained widespread application and research across various fields [22,23]. In the domains of health, gamification primarily targets children for chronic disease management [24], physical activity [25], mental health [26], and nutrition education [27,28]. To the best of our knowledge, there is a lack of studies applying gamification in dietary surveys and assessments, especially for children and adolescents.

Therefore, we developed a new gamified dietary assessment tool (GDA) that engages children and adolescents in a simulated autonomous eating scenario, aiming to capture their authentic food preferences and eating behavior tendencies. An exploratory comparison study was conducted to compare the GDA with 24HR, followed by an application study to explore how the GDA captures dietary behavior tendencies beyond the scope of traditional assessment methods. Through the integration of gamification, we seek to determine whether this innovative method can provide a more comprehensive understanding of young individuals' nutrition status and true dietary tendencies.

Methods

Study Design and Population

The study comprised 2 parts. The first is an exploratory study, aiming to compare the hypothetical food choices assessed by the GDA with the actual dietary intake estimated through a traditional 24HR. The second part is an application study, using the GDA to simulate the eating process of children and adolescents to assess their dietary intake and tendencies of eating behaviors.

The exploratory study was a cross-sectional study assessing the agreement of the interviewer-led 3-day 24HR and the GDA. The research took place at an elementary school in Changzhou, Jiangsu province, China. A total of 30 school-aged children were recruited voluntarily. The exclusion criteria for participants were (1) patients with metabolic diseases or other conditions requiring a special diet and (2) patients with eating disorders or other disorders that significantly impact eating behaviors.

The application study was conducted among school-aged children and adolescents at a middle school in Lanzhou, Gansu province, China. Cluster sampling was used to select 40 classes, which contained 2241 students. The inclusion and exclusion criteria were the same as the exploratory study. To ensure data quality, we further excluded (1) the ones with abnormal energy intake (energy intake exceeding 4000 kcal or below 500 kcal), (2) participants who did not complete the Emotional Eating Scale, and (3) those who failed the attention check questions.

Ethical Considerations

The study was approved by the ethics committee of Institutional Review Board of Tsinghua University (Project No: 20220102). All participants and their parents were required to give consent before participating in the study, through which they were informed about the procedures and other related information of the study. No compensation was provided to participants. All human participant data have been anonymized or deidentified to protect participants' privacy, with no personal identifiers retained in records or analyses.

Data Collection

Data collection for 24HR and GDA in the exploratory study took place in May 2023. The first 24HR interview was conducted on day 1 by trained researchers, followed by the self-administered GDA survey on the same day. The remaining 2 24HR interviews were conducted on days 2 and 3, respectively.

In the application study, basic information was collected by trained researchers using a standard protocol. BMI-for-age z score (BMI z score) was then calculated according to the growth reference standards for children and adolescents aged 5 to 19 years provided by the World Health Organization [29]. The severity of emotional eating behaviors was assessed by emotional eating scores (EESs) based on the Emotional Eating Scale [30]. This scale is based on 1 dimension of the Intuitive Eating Scale-2, specifically the "Eating for Physical Rather than Emotional Reasons" dimension, which includes 8 items, 4 of which are reverse-scored. As no version specifically validated

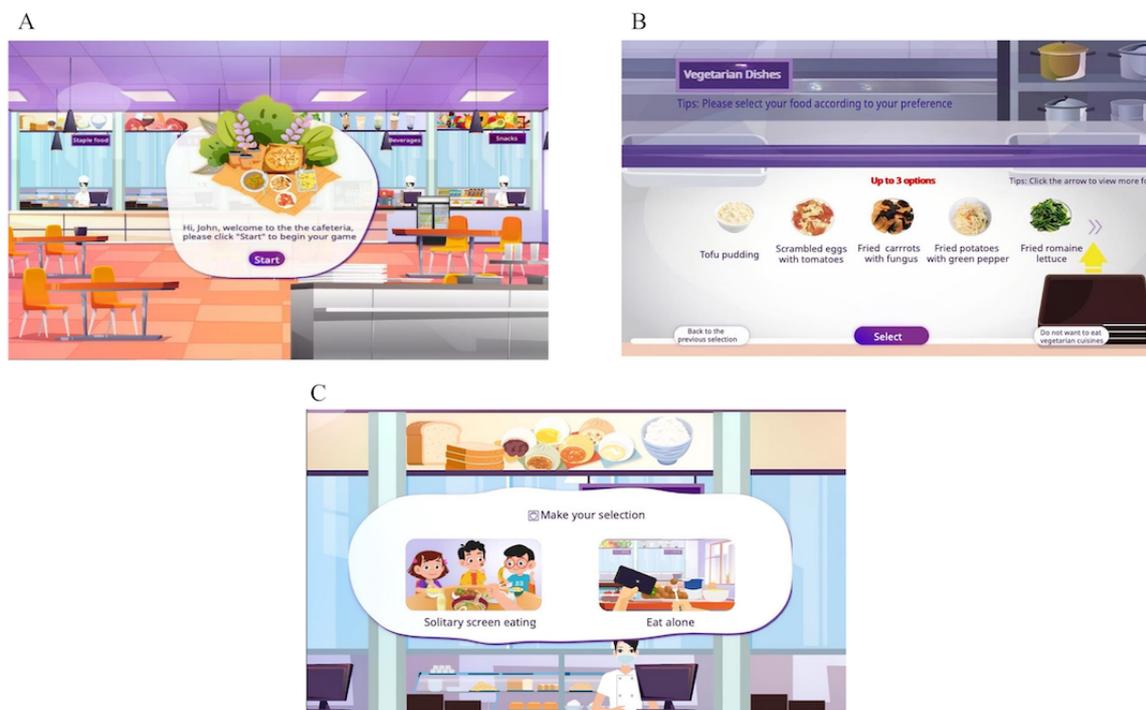
for Chinese children and adolescents is currently available, a widely applied general version of the EES was adopted. Participants were instructed, “For each item, please check the answer that best characterizes your eating attitudes or behaviors.” The item-response scale was a 5-point scale, where 1 indicates strongly disagree, 2 indicates disagree, 3 indicates neutral, 4 indicates agree, and 5 indicates strongly agree. Reverse-scored items were rated inversely, with 1 indicating strongly agree to 5 indicating strongly disagree. EES ranged from 5 to 40, with higher scores indicating a greater likelihood of engaging in emotional eating behaviors as a way to cope with negative emotions. The EES was self-administered by participants via the digital platform under the supervision and guidance of trained investigators. Dietary intake data and dietary behavior data were collected with the GDA.

Gamified Dietary Assessment Tool

The dietary assessment tool was designed with gamification principles to improve user engagement and accuracy. Core gamification elements included (1) a drag-and-drop interaction mechanic for food selection, (2) a dynamic progress bar and encouraging prompts for real-time feedback, and (3) a visually engaging interface with animations and child-friendly graphics to reduce cognitive burden and enhance immersion. These features aimed to transform traditional dietary recall into a more intuitive and motivating experience for children and adolescents. The GDA consisted of 3 primary modules (Figure 1A-C): the guided tutorial module, the food selection module, and the dining scenario selection module. In the guided tutorial module, the GDA offered prompts to instruct participants on how to

complete the survey. Participants then proceeded to the second module to choose foods for breakfast, lunch, and dinner in sequence under a virtual buffet situation. Children were asked to make their independent decisions on what types of dishes or beverages and what quantities they wished to consume by dragging and dropping food portions from the buffet table. There was a total of 66 dishes or beverages, which consisted of 81 food items belonging to 10 main food groups, including cereals, tubers, vegetables, fruits, livestock and poultry, aquatic products, eggs, dairy products, legumes, nuts, as well as 2 groups of processed food (including snacks and beverages). Specifically, alcoholic beverages were included in the GDA food list to reflect real-life dietary exposure and to allow the identification of potentially harmful alcohol intake tendencies, although alcohol intake data were not analyzed in this study. Dishes were designed by nutritional experts according to the dietary culture of China and their nutritional contents. The details of dish and beverage selection are shown in Multimedia Appendix 1. After completing the food selection, participants were supposed to select the dining scenarios, with 2 options available: “solitary screen eating” and “eat with peers.” Participants who chose “eat with peers” were presented with an animation featuring several children engaging in cheerful conversation while eating. In contrast, those who selected “solitary screen eating” were presented with an animation depicting a child eating while watching cartoons on a mobile phone. The main outputs of the GDA encompassed eating behavior tendencies, estimated food intake, and nutrient intake based on the selected food items in the GDA.

Figure 1. Three primary modules designed in the gamified dietary assessment tool (GDA): (A) module of guided tutorial, (B) module of food selection, and (C) module of dining scenario selection.



Food, Energy, and Nutrient Intake Assessment

In an exploratory study, the quantitative measurement of food intake by 24HR was assisted by Photographic Atlas of Food

Portions for Accurate Quantification of Dietary Intakes in China [31] and standard utensil models, which allowed participants to estimate portion sizes by comparing their intake against

standardized images and models. Thereafter, energy and nutrient intake were calculated according to the China Food Composition Tables, Sixth Standard Edition [32]. The 3-day average intakes of energy and nutrients were used in statistical analysis.

For the GDA, all mixed dishes were decomposed into their constituent ingredients. Energy and nutrient intakes in both exploratory and application studies were calculated by summing the nutrient values of each component based on China Food Composition Tables, Sixth Standard Edition [32].

Data Analysis

The Shapiro-Wilk test for normality indicated that the energy intake and the nutrient intake in the exploratory study were not normally distributed. Consequently, descriptive statistics for energy and nutrient intake were reported as medians with interquartile ranges. The logarithmic transformation of the variables was performed, as dietary data on energy, nutrient, and food intake typically exhibit a right-skewed distribution that can be effectively normalized by this transformation, and Pearson correlation analysis was used to examine the association between the 2 dietary assessment methods regarding these intake measures. If the log-transformed data still did not follow a normal distribution, Spearman rank correlation analysis was then used to assess the correlation between the 2 methods. Additionally, Bland-Altman plots were constructed to visually assess the agreement between the 24HR and the GDA. This method involves plotting the mean value of the 2 measurements on the x-axis and the difference between them on the y-axis, with the mean difference and 95% limits of agreement displayed to evaluate the consistency between the 2 methods.

In the application study, descriptive statistics for energy intake, nutrient intake, food intake, and macronutrient energy ratio were also presented as medians with interquartile ranges. A comparison of demographic characteristics was also conducted between the included population and those who failed the attention check question using the Wilcoxon test and Fisher exact test. The distribution of macronutrient energy ratios was described and compared with the Dietary Reference Intakes for China [33]. The Kruskal-Wallis rank sum test was used to compare the energy intake, nutrient intake, food intake, and macronutrient energy ratio among participants with different sociodemographic characteristics, while Fisher exact test was

used to compare the distributions of macronutrient energy ratio across different categorical groups due to small numbers in certain distribution groups. We used multiple linear regression models to examine the association between EES and nutritional outcomes. Age, sex, and socioeconomic status were included as covariates. Energy intake was divided by 10 due to the large values of energy intake. Stratified analysis was further conducted according to BMI z scores. BMI z scores less than -2 were defined as “thin,” scores -2 or higher and 1 or less as “normal,” and scores greater than 1 as “overweight and obese” [34]. The β -values and 95% CI for each model were calculated to assess the effects of emotional eating behavior tendencies on nutritional outcomes. Three models were applied: model 1 made no adjustments; model 2 adjusted for age, sex, and socioeconomic status; and model 3 further adjusted for energy intake. Additionally, dynamic scatter plots were generated using the “gganimate” package in R, illustrating the association between EES and macronutrient intake across different BMI z score groups to visualize how emotional eating affects energy and macronutrient intake across weight statuses. Multiple linear regression models were also used to evaluate the association between dining environment tendencies and nutritional outcomes.

Statistical analysis was conducted using the R software (version 4.3.0; Posit PBC), with all tests being 2-sided and a significance level set at .05.

Results

Exploratory Study

A total of 30 school-aged children were recruited, including 9 (30%) male participants and 21 (70%) female participants. Participants ranged in age from 9 to 11 years, with an average age of 10.7 years (SD 0.55). Nutrient data from the GDA and 24HR were not normally distributed. Consequently, a logarithmic transformation was applied to the energy and nutrient intake values. Pearson correlation analysis of the transformed data showed significant positive correlations between the 2 methods for energy intake, carbohydrates, cholesterol, and a few micronutrients (vitamin A and vitamin C). No significant correlations were observed for the other nutrients, as presented in Table 1.

Table . Correlation between intakes of energy, nutrients, and main food groups in 24-hour dietary recalls (24HR) and the gamified dietary assessment tool (GDA).

Variables	GDA, median (IQR)	24HR, median (IQR)	Correlation coefficient	<i>P</i> value
Energy and nutrients				
Energy (kcal)	1596.8 (1370.3-1895.3)	1510.8 (1335.0-1841.8)	0.46	.01
Protein (g)	81.4 (63.3-94.0)	48.7 (42.4-57.0)	0.28	.13
Fat (g)	73.0 (46.9-87.4)	67.8 (58.0-80.5)	0.31	.10
Carbohydrate (g)	170.5 (129.3-209.1)	188.1 (167.8-220.5)	0.50	.005
Cholesterol (mg)	624.0 (392.4-746.3)	328.9 (219.3-448.6)	0.38	.04
VA (µgRAE)	377.2 (224.7-525.3)	343.9 (259.7-486.5)	0.41	.03
VB1 (mg)	0.5 (0.3-0.6)	0.6 (0.4-0.7)	-0.17	.37
VB2 (mg)	1.0 (0.8-1.2)	0.6 (0.5-0.7)	0.23	.21
Niacin (mg)	15.4 (9.9-20.7)	9.4 (6.9-10.8)	-0.02	.91
VC (mg)	63.5 (41.2-84.0)	30.3 (41.5-54.7)	0.52	.003
VD (µg) ^a	3.3 (2.4-4.3)	0 (0-0)	0.03	.86
VE (mg)	13.2 (11.5-16.8)	30.8 (25.8-37.4)	0.01	.94
Ca (mg)	525.0 (396.0-631.2)	461.7 (327.6-561.9)	0.07	.71
P (mg)	977.7 (756.0-1134.1)	684.7 (538.5-782.3)	0.01	.95
K (mg)	1987.2 (1585.8-2255.6)	1287.3 (1032.5-1740.6)	0.31	.10
Na (mg)	2446.9 (1797.3-3061.4)	5054.7 (4541.2-5710.1)	0.07	.72
Mg (mg)	241.0 (189.8-269.4)	193.2 (152.4-243.9)	0.24	.19
Fe (mg)	15.2 (12.0-17.9)	11.3 (9.1-16.1)	-0.01	.96
Zn (mg)	10.4 (7.6-61.5)	6.7 (6.0-8.3)	0.26	.16
Se (µg)	51.8 (35.3-76.9)	25.3 (20.4-32.5)	0.36	.05
Cu (mg)	1.3 (0.9-9.5)	1.0 (0.8-1.3)	0.17	.38
Mn (mg)	2.2 (1.9-2.5)	2.0 (1.5-2.5)	0.08	.69
Main food groups				
Cereals (g)	194.2 (160.0-267.1)	205.0 (168.9-257.0)	0.44	.01
Tubers (g)	58.5 (0-58.5)	6.7 (0-17.5)	0.11	.55
Vegetables (g)	181.1 (129.1-241.1)	73.3 (41.7-125.8)	0.34	.07
Fruits (g)	60.0 (30.6-61.2)	55.5 (0-132.2)	0.41	.02
Livestock and poultry (g)	206.3 (137.6-247.3)	80.8 (53.1-118.7)	0.18	.34
Aquatic products (g)	90.0 (0-150.0)	0 (0-19.2)	0.25	.18
Eggs (g)	41.0 (36.8-73.8)	37.5 (20.0-60.0)	0.15	.43
Dairy products (g)	124.5 (55.4-176.7)	141.7 (83.3-226.7)	0.53	.003
Legumes (g)	0.0 (0-58.4)	0 (0, 7.5)	-0.28	.13
Nuts (g) ^a	0.0 (0.0-30.0)	0.0 (0.0-0.0)	-0.26	.17

^aThe intake of vitamin D and nuts showed minimal variability in 24-hour dietary recalls, which limits the interpretability of correlation coefficients for these variables.

Food intake data were nonnormally distributed, requiring logarithmic transformation, but the transformed data still remained nonnormally distributed. Thus, Spearman rank correlation analysis was conducted. Significant positive correlations were observed between 24HR and the GDA for the

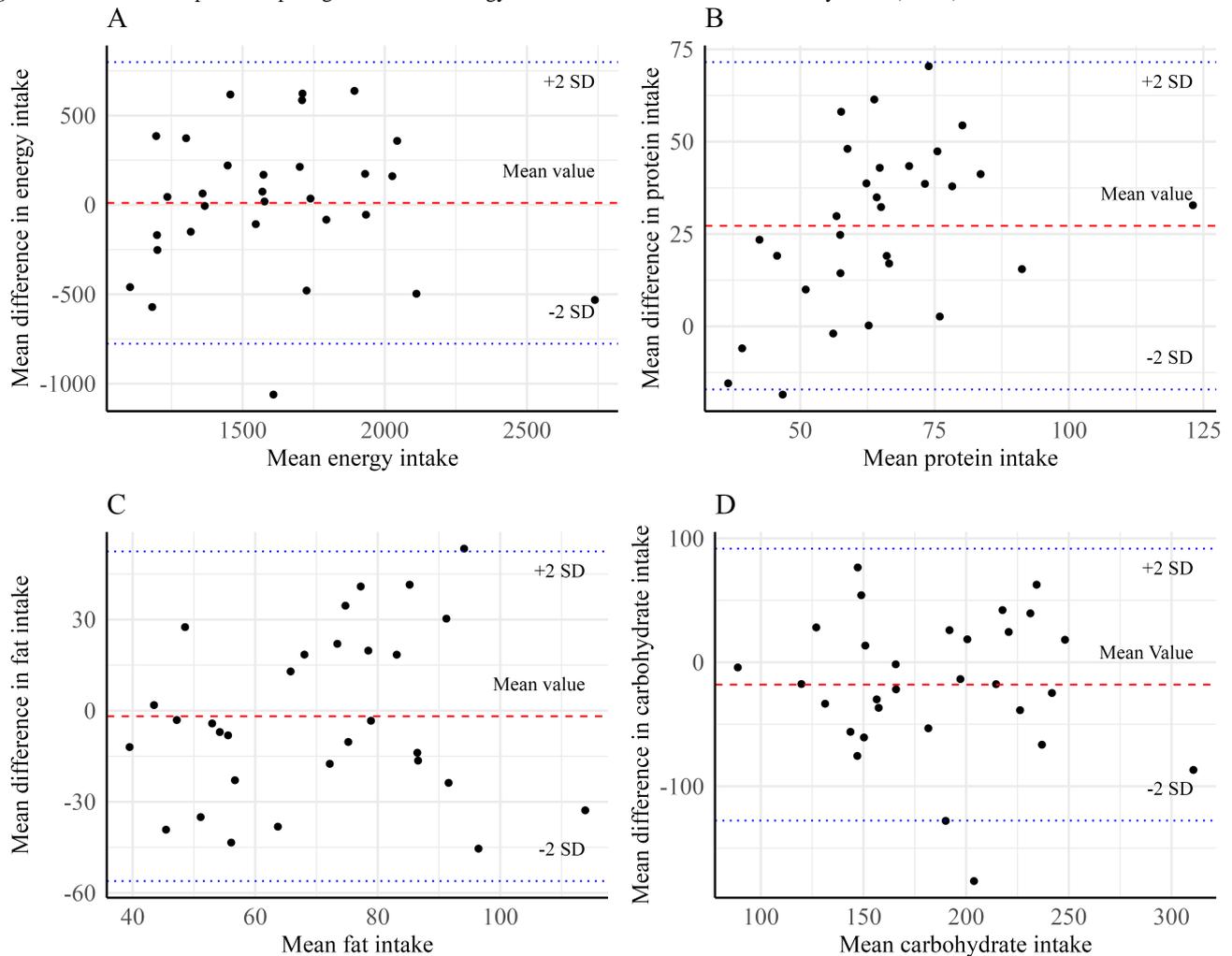
intake of cereals, fruits, and dairy products, while no significant correlations were found for other food categories (Table 1).

The Bland-Altman plot showed that the mean differences in energy (Figure 2A), fat (Figure 2C), and carbohydrate (Figure 2D) intake between the 24HR and the GDA were centered around zero, with all observations evenly distributed within the

95% CI on both sides of the mean difference, which indicated good agreement between the 2 methods. However, the mean difference for protein intake (Figure 2B) was approximately 25

g, suggesting that the GDA may indicate a higher protein intake compared with 24HR.

Figure 2. Bland-Altman plots comparing the intake of energy and macronutrients in 24-hour dietary recall (24HR).

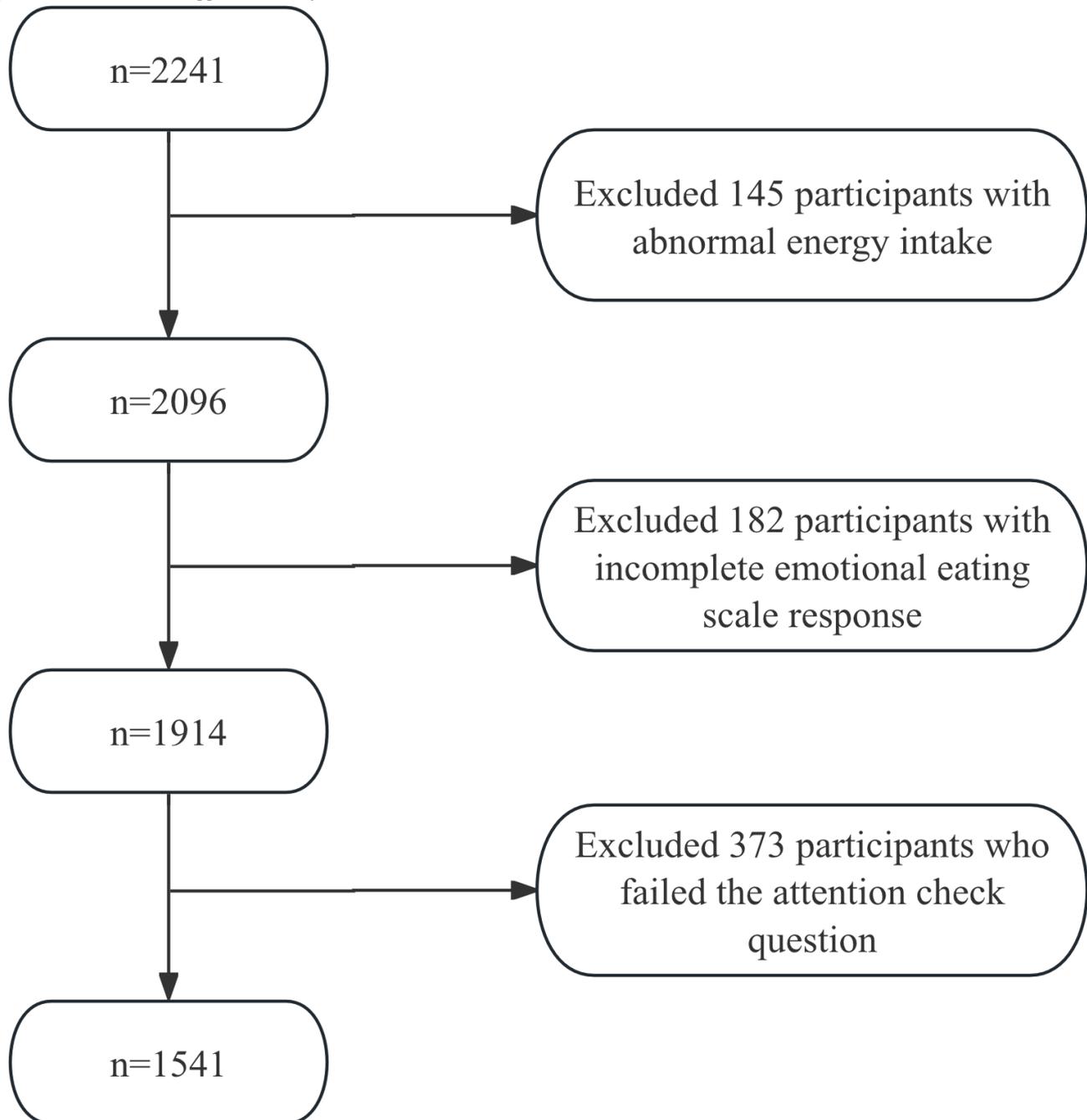


Application Study

In the application study, a total of 2241 participants were initially recruited, resulting in 1541 valid samples (Figure 3), comprising 949 (61.6%) male participants and 592 (38.4%)

female participants with an age range of 11 to 18 years and a mean age of 15.2 (SD 1.6) years. Participants excluded due to failed attention checks were significantly younger than those included in the final sample ($P < .001$).

Figure 3. Flowchart of the application study.



The median GDA-estimated energy, protein, fat, and carbohydrate intakes were 1877.3 (IQR 1410.0-2221.8) kcal, 96.4 (IQR 69.5-113.9) g, 74.0 (IQR 52.5-96.8) g, and 202.0 (IQR 156.3-238.7) g, respectively. Participants aged 11 to 14 years showed a significantly greater tendency to intake more carbohydrate compared to those aged 15 to 18 years ($P < .001$). Male participants demonstrated a higher tendency to intake more energy ($P = .003$), carbohydrate ($P = .04$), protein ($P = .01$), and fat ($P = .01$) than female participants. However, no significant differences were observed in GDA-estimated energy and macronutrient intake across different weight status levels or socioeconomic status levels. The median GDA-estimated intake of various food groups among the participants was 210.1 (IQR 159.7-270.0) g for cereals, 58.5 (IQR 0-58.5) g for tubers, 201.0 (IQR 126.7-279.3) g for vegetables, 30.6 (IQR 0-61.2) g for

fruits, 250.5 (IQR 156.2-340.0) g for livestock and poultry, 170.0 (IQR 90.0-250.0) g for aquatic products, 39.0 (IQR 12.9-67.1) g for eggs, 45.7 (IQR 11.1-148.1) g for dairy products, 0 (IQR 0-58.4) g for legumes, and 0 (IQR 0-0) g for nuts, 240.0 (IQR 120.0-360.0) mL for beverages, and 78.0 (IQR 48.0-106.8) g for snacks. Significant differences were found in GDA-estimated intake of cereals ($P = .02$), fruits ($P = .003$), and legumes ($P = .02$) across age groups. Sex differences were also observed, with male participants tending to consume significantly more cereals ($P < .001$), vegetables ($P < .001$), livestock and poultry ($P = .02$), and legumes ($P = .009$) than female participants, while female participants had a higher GDA-estimated intake of fruits ($P < .001$) and dairy products ($P < .001$) than male participants. No significant differences were observed in other food groups across different weight status

levels or socioeconomic status levels ([Multimedia Appendix 2](#)).

The median ratios of GDA-estimated protein energy contribution, fat energy contribution, and carbohydrate energy contribution were 20.4% (IQR 18.4%-22.1%), 35.8% (IQR 31.4%-41.1%), and 44.0% (IQR 39.7%-49.0%), respectively. The 11- to 14-year group had a higher carbohydrate energy ratio ($P<.001$) and a lower fat energy ratio ($P<.001$) compared with the 15- to 18-year group. Sex difference was also observed, with female participants having a higher carbohydrate energy ratio ($P=.002$). A total of 42.9% ($n=661$) of the participants had a protein energy ratio within the acceptable macronutrient distribution range (AMDR) (10% - 20%). 17.9% ($n=276$) of the participants had a fat energy ratio within the AMDR (20% - 30%). 20.1% ($n=310$) of the participants had a

carbohydrate energy ratio within the AMDR (50% - 65%). Only the distribution of carbohydrate energy intake ratios ($P=.02$) showed significant differences across gender ([Multimedia Appendix 3](#)).

Furthermore, we explored the association between emotional eating behavior and nutritional outcomes among participants using the GDA. The median (IQR) EES was 22.3 (19-25). After adjusting for age, sex, and socioeconomic status, multiple linear regression analysis showed that higher EESs were significantly associated with a greater tendency to select more carbohydrates, while no significant associations were found for protein or fat selections. After further adjusting for energy intake tendencies, higher EESs were associated with a lower protein intake tendency, with no significant associations observed for fat or carbohydrate intake tendencies ([Table 2](#)).

Table . The association between gamified dietary assessment tool (GDA)-estimated intakes of energy, macronutrients, food groups, and emotional eating scores.

Variables	β (95% CI)		
	Model 1 ^a	Model 2	Model 3
Energy/10 (kcal)	0.202 (-0.276 to 0.680)	0.348 (-0.136 to 0.832)	— ^b
Protein (g)	-0.052 (-0.307 to 0.202)	0.009 (-0.249 to 0.267)	-0.159 (-0.267 to -0.052)
Fat (g)	0.006 (-0.266 to 0.278)	0.064 (-0.212 to 0.339)	-0.106 (-0.249 to 0.038)
Carbohydrate (g)	0.382 (-0.121 to 0.884)	0.518 (0.009 to 1.027)	0.199 (-0.050 to 0.448)
Cereals (g)	-0.297 (-1.037 to 0.443)	0.060 (-0.683 to 0.803)	-0.135 (-0.827 to 0.557)
Tubers (g)	-0.047 (-0.425 to 0.331)	-0.055 (-0.439 to 0.329)	-0.123 (-0.496 to 0.249)
Vegetables (g)	-0.466 (-1.373 to 0.442)	-0.135 (-1.052 to 0.783)	-0.398 (-1.240 to 0.445)
Fruits (g)	-0.284 (-0.676 to 0.107)	-0.429 (-0.822 to -0.035)	-0.464 (-0.854 to -0.073)
Livestock and poultry (g)	-0.013 (-1.108 to 1.081)	0.196 (-0.910 to 1.302)	-0.405 (-1.131 to 0.322)
Aquatic products (g)	-0.110 (-1.044 to 0.824)	-0.004 (-0.952 to 0.944)	-0.239 (-1.130 to 0.652)
Eggs (g)	-0.199 (-0.483 to 0.084)	-0.163 (-0.450 to 0.124)	-0.237 (-0.505 to 0.031)
Dairy products (g)	0.527 (-0.310 to 1.364)	0.250 (-0.597 to 1.096)	0.077 (-0.736 to 0.890)
Legumes (g)	-0.031 (-0.461 to 0.400)	0.099 (-0.338 to 0.535)	-0.027 (-0.398 to 0.452)
Nuts (g)	-0.169 (-0.289 to -0.049)	-0.171 (-0.294 to -0.049)	-0.183 (-0.304 to -0.062)
Beverages (mL)	1.270 (-0.015 to 2.637)	1.390 (0.087 to 2.683)	1.040 (-0.166 to 2.250)
Snacks (g)	0.461 (0.040 to 0.882)	0.540 (0.113 to 0.966)	0.438 (0.035 to 0.840)

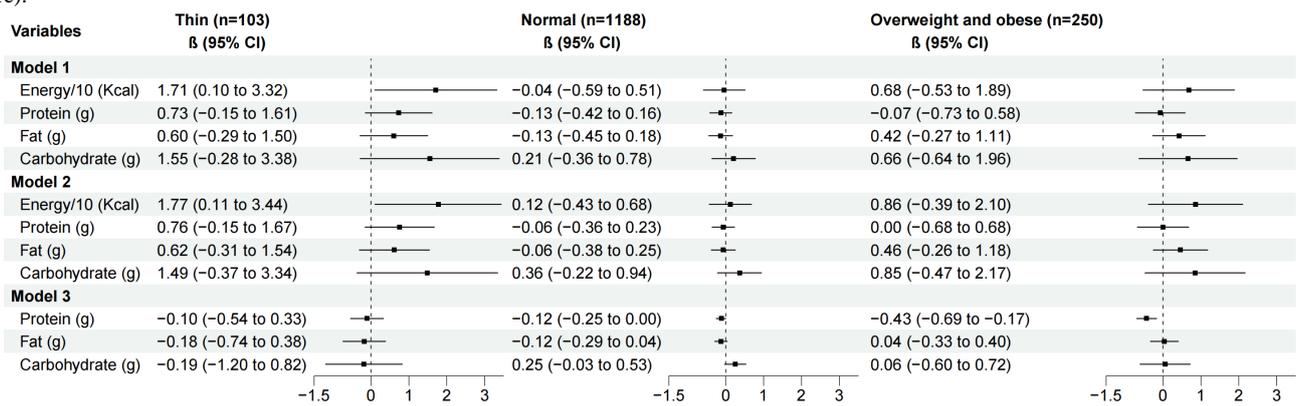
^aModel 1 made no adjustment for any covariates. Model 2 was adjusted for age, sex, and socioeconomic status. Model 3 was adjusted for energy intake, age, sex, and socioeconomic status.

^bNot available.

Emotional eating behavior also influenced the selections of several food groups. After adjusting for age, sex, and socioeconomic status, higher EESs remained positively associated with a greater tendency to select beverages and snacks, while the selection of fruits and nuts was significantly fewer. After further adjustment for total energy intake, higher EESs were still associated with higher snack selection and fewer selections of fruits and nuts, with no significant associations observed for other food groups ([Table 2](#)).

Stratified analysis indicated that higher EESs remained associated with a greater energy intake tendency among the thin population, even after adjusting for age, sex, and socioeconomic status. In the overweight and obese population, a higher EES was associated with a lower tendency to select protein-rich foods, but only after adjusting for age, sex, and total energy intake tendency. In the normal weight population, no significant associations were found between EES and the intake tendencies for energy, protein, carbohydrates, or fat ([Figure 4](#)).

Figure 4. Associations between the emotional eating scale (EES) and intakes of energy and macronutrients stratified by BMI-for-age z score (BMI z score).



Additionally, dynamic scatter plots were used to provide a clear visual representation of how the association between EES and the energy intake tendency levels varied across different BMI z scores. However, there is no significant association observed between EES and the intake tendencies of energy, protein, carbohydrates, or fat across different BMI z score categories (Multimedia Appendix 4).

We also applied the GDA to investigate the association between dining environment tendencies and nutritional outcomes. In total, 525 (34.1%) participants selected “eat with peers”, while 1016 (65.9%) participants selected “solitary screen eating”, and

no significant difference in scenario selection was observed across age, sex, and socioeconomic status. The results presented in Table 3 indicated that participants who wished to eat alone while watching cartoons tended to select 6.9 g more carbohydrates, on average, than those who wished to eat with peers. Even after adjusting for age, sex, and socioeconomic status, those who wished to eat alone while watching cartoons still tended to select 7.7 g more carbohydrates than those who wished to eat with peers. Further adjusting for total energy intake, the difference remained significant, with those who wished to eat alone while watching cartoons tending to select 4.2 g more carbohydrates.

Table . Effects of dining environment tendencies on intakes of energy, macronutrients, and food groups.

Variables	β (95% CI)	Model 1 ^a	Model 2	Model 3
Energy (kcal)	4.0 (-1.8 to 9.8)		3.9 (-1.9 to 9.6)	— ^b
Protein (g)	2.3 (-0.7 to 5.4)		2.1 (-1.0 to 5.2)	0.3 (-1.0 to 1.5)
Fat (g)	1.4 (-1.8 to 4.7)		0.9 (-2.3 to 4.2)	-0.9 (-2.6 to 0.8)
Carbohydrate (g)	6.9 (0.9 to 13.0)		7.7 (1.6 to 13.8)	4.2 (1.2 to 7.1)
Cereals (g)	8.6 (-0.3 to 17.5)		10.2 (1.4 to 19.1)	8.1 (-0.2 to 16.3)
Tubers (g)	4.4 (-0.2 to 8.9)		4.8 (0.3 to 9.4)	4.1 (-0.4 to 8.5)
Vegetables (g)	1.3 (-9.7 to 12.2)		1.3 (-9.6 to 12.3)	-1.6 (-11.6 to 8.5)
Fruits (g)	2.2 (-2.5 to 6.9)		1.1 (-3.6 to 5.8)	0.7 (-4.0 to 5.4)
Livestock and poultry (g)	5.8 (-7.4 to 18.9)		3.5 (-9.7 to 16.7)	-3.2 (-11.8 to 5.5)
Aquatic products (g)	1.5 (-9.8 to 12.7)		1.9 (-9.4 to 13.2)	-0.7 (-11.4 to 9.9)
Eggs (g)	4.3 (0.9 to 7.7)		3.8 (0.4 to 7.3)	3.0 (-0.2 to 6.2)
Dairy products (g)	-2.8 (-12.9 to 7.3)		-3.1 (-13.2 to 7.0)	-5.1 (-14.7 to 4.6)
Legumes (g)	-1.3 (-6.5 to 3.9)		-1.0 (-6.2 to 4.2)	-1.8 (-6.8 to 3.3)
Nuts (g)	-1.0 (-2.5 to 0.4)		-1.1 (-2.5 to 0.4)	-1.2 (-2.7 to 0.2)
Beverages (mL)	15.4 (-0.1 to 30.8)		17.7 (2.3 to 33.2)	13.9 (-0.5 to 28.3)
Snacks (g)	2.4 (-2.7 to 7.5)		3.3 (-1.8 to 8.4)	2.2 (-2.6 to 7.0)

^aParticipants who selected “eating with peers” were used as reference. Model 1 made no adjustment for any covariates. Model 2 was adjusted for age, sex, and socioeconomic status. Model 3 was adjusted for energy intake, age, sex, and socioeconomic status.

^bNot available.

Regarding main food groups, participants who wished to eat alone while watching cartoons tended to select 4.3 g more eggs,

on average, compared to those who wished to eat with peers, a statistically significant difference. After controlling for age,

sex, and socioeconomic status, the difference remained significant, and the former group tended to consume 10.2 g more cereals, 4.8 g more tubers, 3.8 g more eggs, and 17.7 mL more beverages than the latter (Table 3). However, after further adjusting for energy intake, a willingness to eat alone while watching cartoons did not influence the selection of any food group.

Discussion

This study suggests that the GDA may be a promising approach for simulating dietary choices and assessing eating behavior tendencies, providing insights into developing novel dietary survey methods.

GDA Could Reflect Some Aspects of Dietary Intakes

In exploratory analysis, we observed moderate associations between the GDA and 24HR for energy, carbohydrate, cholesterol, and certain vitamins, indicating that adolescents' hypothetical food selections in the GDA may reflect some aspects of their actual intake. However, no significant correlations were found for other nutrients, such as protein, fat, and several micronutrients, suggesting that the GDA may capture only specific dimensions of dietary behavior and not serve as a comprehensive measure of actual intake. This discrepancy may be attributed to the limited food options in the GDA, which cannot fully capture the variety of real-world foods. Despite efforts to include representative foods from different regions of China, a gap remains between the actual food choices of children and adolescents and those offered in the GDA. Similarly, a study using virtual reality technology to simulate dietary choices also found differences between the foods selected in the real world and those chosen in the virtual environment [12].

Additionally, the exploratory study found that the GDA significantly overestimated protein intake, possibly reflecting children's preferences for protein-rich foods. This finding aligns with the results of the application study, which showed that livestock and poultry meat intake exceeded 200 g, significantly higher than the recommended value in the Dietary Reference Intakes for China [33]. The high proportion of high protein energy ratio distribution may also result from the overestimation. Over the past decades, rapid economic development has significantly increased meat consumption in China [35], and people have become more aware of the importance of protein for the growth and development of children and adolescents, which may have led to a preference for meat among them. Another explanation could be the biased visual presentation or portion sizes of protein-rich foods in the GDA. It may have made these items more appealing, potentially encouraging the selection over other food groups. Excessive meat consumption can increase the risk of obesity [36]; therefore, the meat intake preferences identified in children and adolescents through the GDA should be addressed promptly. Dietary patterns during childhood and adolescence are highly malleable [37], making it crucial to correct unhealthy food preferences at this stage to promote lifelong health. Several other nutrients also showed substantial discrepancies between the GDA and the 24HR, although these differences are significant. Notably, sodium was

markedly underestimated in the GDA, likely because the simulated environment includes fewer high sodium processed foods. For vitamin D, the GDA data showed a median (IQR) intake of 3.3 (2.4-4.3) μg , whereas 24HR's median (IQR) intake was 0 (0-0) since there were only 2 nonzero values. These differences further indicate that the GDA captures preference-driven hypothetical selections rather than actual nutrient intake, underlying the conceptual distinction between the 2 assessment methods.

Role of the GDA in Uncovering Food Preferences and Eating Behavior Tendencies

In the application analysis, we found the absence of significant differences in GDA-estimated energy selection across weight status groups. This may reflect both developmental and methodological factors. During adolescence, weight status does not necessarily correspond to energy requirements because of large variability in growth and maturation. Besides, the GDA captures hypothetical, preference-based food selection rather than physiological energy needs, allowing adolescents across weight groups to select some high-calorie foods in a virtual setting. Additionally, the limited range of food options in the GDA may further attenuate between-group differences. The primary aim of developing the GDA is to capture dietary behaviors and food preferences among children and adolescents. These aspects are often overlooked and challenging to obtain through traditional dietary surveys. In our application study, we found that the GDA can identify adverse eating behaviors. Individuals with high EESs tended to consume more carbohydrates and less protein, suggesting that they coped with their emotions by selecting high-carbohydrate foods. Regarding food groups, participants with higher EESs showed a greater tendency to choose more beverages and snacks and fewer fruits and nuts, despite median selections for some groups being zero. Previous studies have shown that emotional eating behavior in adolescents was positively associated with higher consumption of high-energy-dense foods, such as cake, ice cream, and beverages [38]. The mean frequency of high-energy-dense food intake was roughly 3 to 4 times higher than that of fruits and vegetables [39]. However, traditional dietary surveys collect data on school- or family-provided meals, potentially masking children's tendencies for emotional eating. The GDA may help reveal more authentic dietary behavior tendencies, such as emotional eating, by allowing children and adolescents to make food choices without the external constraints typically present in real-life settings—such as parental supervision or school meal policies. In this autonomy-simulated environment, their intrinsic preferences and decision-making patterns—such as a tendency to select high-calorie or high-fat foods—can be more clearly observed. This enables the following targeted psychological counseling and nutrition education, which not only enhances the mental health of children and adolescents but also promotes healthy eating patterns. More interestingly, a positive association between EESs and higher energy selection tendency was observed only in the thin subgroup. This finding suggests that the GDA may be particularly sensitive to identifying latent emotional eating tendencies before they translate into weight gain. Emotional eating behaviors may occur earlier than changes in BMI, especially among thin

adolescents. However, this result was preliminary and based on a relatively small sample size of thin groups, which should thereby be interpreted with caution.

In this study, we also identified that the dining environment can impact the nutritional outcomes of children and adolescents [40]. In this study, children and adolescents who selected the solitary screen eating scenario tended to choose more carbohydrates than those in the social eating scenario. Screen time during meals may be the critical influencing factor. A number of studies in Europe [41] and the United States [42] have demonstrated that screen exposure during meals, such as watching television, is significantly associated with obesity in children and adolescents. Therefore, the observed association could be attributable to the simulated screen time, which is known as an adverse eating behavior. Although the observed 7 g difference in daily carbohydrate intake may not be clinically significant in isolation, it may reflect meaningful behavioral tendencies—such as watching screens while eating food—that could cumulatively influence dietary patterns and health outcomes over time, particularly during adolescence. Our finding aligns with previous studies [43], as screen time, such as TV viewing, has been associated with increased consumption of sugar-rich foods, leading to higher carbohydrate intake [44].

Further Development

In the future, to fully realize the potential of the GDA in dietary assessment, ongoing improvements and upgrades to the GDA are needed. First, to better simulate children's eating processes and assess dietary intake, the GDA should incorporate a wider variety of foods and regularly update precise nutritional data for each item. Second, the mealtime environment can significantly influence the nutritional and behavioral outcomes of children and adolescents [45,46]. The GDA should therefore enrich meal setting options, such as meal duration and dining location, to maximize its effectiveness in addressing adverse eating behavior tendencies in advance. Finally, gamification has already been successfully applied in developing certain nutrition education games [27,47]. The GDA could also incorporate a food and nutrition education module, making the nutrition education process more engaging and accessible while offering users personalized, practical dietary improvement suggestions.

Strengths and Limitations

The study used a dual-phase approach, integrating both an exploratory study and an application study. This design not only explored the application of gamification in dietary assessment but also provided insights into the dietary behavior tendencies and food preferences of children and adolescents in a simulated context. However, we acknowledged some limitations. First, this study was conducted solely among populations in 2 Chinese cities, limiting the representativeness of the findings. Future research should involve a larger, multicenter population sample to enhance the generalizability of the results. Second, the exploratory study did not compare the GDA with the weighed record method, which was regarded as the gold standard for dietary assessment but was challenging to implement. The applicability of the findings may also be limited as EES was not specifically validated for our study population of Chinese

children and adolescents. Third, this exploratory study included a small sample size ($n=30$), which may limit the generalizability of the findings. However, the primary aim was to preliminarily assess the consistency between the GDA and 24HR rather than to provide definitive validation or precise estimates. The thin subgroup ($n=103$) in stratified analysis was also small. This limited sample size reduced the statistical power of the regression models and resulted in wide confidence intervals, suggesting that the observed associations in this subgroup should be interpreted cautiously. Fourth, as participants in the application study were middle school students in an older age range, further research is needed to determine the applicability of the GDA for younger children. Another limitation is that the 24HR did not include weekend days, which might have introduced bias, as dietary patterns can vary between weekdays and weekends. This was due to the on-site data collection schedule during school days. Besides, this study did not investigate user acceptability, which may vary across ages and cultural backgrounds in response to the GDA. In addition, the high rate of exclusion due to failed attention checks (373 out of 2241) suggested potential selection bias since participants who failed attention checks were significantly younger. Besides, in the GDA, the solitary eating scenario involves screen exposure. Therefore, the independent effects of eating alone and screen exposure could not be disentangled. Finally, there are some methodology limitations. As the GDA is designed to capture preference-based dietary behavior tendencies rather than actual intake, several food groups exhibited zero-inflated distributions with median values of zero. These patterns reflect selective hypothetical choices rather than true absence of consumption and may limit the suitability of standard linear regression. There are conceptual and methodological mismatches between the GDA and 24HR. While 24HR aims to quantify 3-day averaged actual nutrient and energy intake, which smooths out daily variations, the GDA is designed to capture 1-time hidden dietary behavior tendencies in a simulated environment. This difference reflects the distinct purposes of the tools and should be considered when interpreting their comparison. Moreover, the gamified nature of the virtual buffet may have encouraged curiosity-driven selections that do not fully reflect habitual dietary behavior, potentially introducing bias unrelated to dietary behavior.

Conclusions

This study compared the new gamified dietary assessment tool with the interview-led 24HR among children and adolescents for intakes of food groups, nutrients, and energy. While the GDA cannot fully reflect actual food and nutrient intake, it is primarily designed to capture dietary intake tendencies, particularly adverse eating behaviors that may not be easily detected by traditional dietary assessment methods. Despite some discrepancies with 24HR data, the GDA provides valuable insights into patterns of food selection and eating habits tendencies among children and adolescents, which could be hidden in traditional dietary survey methods. The GDA can thus serve as a complementary tool to traditional dietary assessment instruments in the field of dietary survey methods, thereby enhancing the comprehensiveness and accuracy of dietary survey outcomes.

Acknowledgments

The authors wish to acknowledge with gratitude the students who participated in this study and the teachers who assisted with the implementation of the project in Changzhou and Lanzhou.

Funding

This work was supported by the Sanming Project of Medicine in Shenzhen (number SZSM202111001) and the Research Foundation of Medical Science and Technology of Guangdong Province (grant C2023108).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

YZ was responsible for study design, participant recruitment, data collection, statistical analysis, and manuscript preparation. TL was responsible for survey development, data collection, and manuscript preparation. YT, HY, XS, YY, and XN were responsible for participant recruitment. AZ was responsible for survey development and manuscript preparation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Cuisines and beverages that are included in the gamified dietary assessment tool (GDA).

[[DOCX File, 18 KB - games_v14i1e72387_app1.docx](#)]

Multimedia Appendix 2

Distribution of average gamified dietary assessment tool (GDA)–estimated intakes of energy, macronutrients, and food groups with different sociodemographic status.

[[DOCX File, 31 KB - games_v14i1e72387_app2.docx](#)]

Multimedia Appendix 3

Gamified dietary assessment tool (GDA)–estimated macronutrient energy ratios and their distributions (%).

[[DOCX File, 21 KB - games_v14i1e72387_app3.docx](#)]

Multimedia Appendix 4

Dynamic scatter plots indicating changing associations between emotional eating scores and intakes of energy and macronutrients with varying BMI *z* scores.

[[MP4 File, 2551 KB - games_v14i1e72387_app4.mp4](#)]

References

1. Chilot D, Belay DG, Merid MW, et al. Triple burden of malnutrition among mother-child pairs in low-income and middle-income countries: a cross-sectional study. *BMJ Open* 2023 May 9;13(5):e070978. [doi: [10.1136/bmjopen-2022-070978](https://doi.org/10.1136/bmjopen-2022-070978)] [Medline: [37160393](https://pubmed.ncbi.nlm.nih.gov/37160393/)]
2. Obesity and overweight. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> [accessed 2025-02-20]
3. Caleyachetty R, Thomas GN, Kengne AP, et al. The double burden of malnutrition among adolescents: analysis of data from the Global School-Based Student Health and Health Behavior in School-Aged Children surveys in 57 low- and middle-income countries. *Am J Clin Nutr* 2018 Aug 1;108(2):414-424. [doi: [10.1093/ajcn/nqy105](https://doi.org/10.1093/ajcn/nqy105)] [Medline: [29947727](https://pubmed.ncbi.nlm.nih.gov/29947727/)]
4. Malnutrition. World Health Organization. 2024. URL: <https://www.who.int/news-room/fact-sheets/detail/malnutrition> [accessed 2025-02-20]
5. Franks PW, Hanson RL, Knowler WC, Sievers ML, Bennett PH, Looker HC. Childhood obesity, other cardiovascular risk factors, and premature death. *N Engl J Med* 2010 Feb 11;362(6):485-493. [doi: [10.1056/NEJMoa0904130](https://doi.org/10.1056/NEJMoa0904130)] [Medline: [20147714](https://pubmed.ncbi.nlm.nih.gov/20147714/)]
6. Tirosh A, Shai I, Afek A, et al. Adolescent BMI trajectory and risk of diabetes versus coronary disease. *N Engl J Med* 2011 Apr 7;364(14):1315-1325. [doi: [10.1056/NEJMoa1006992](https://doi.org/10.1056/NEJMoa1006992)] [Medline: [21470009](https://pubmed.ncbi.nlm.nih.gov/21470009/)]

7. Llewellyn A, Simmonds M, Owen CG, Woolacott N. Childhood obesity as a predictor of morbidity in adulthood: a systematic review and meta-analysis. *Obes Rev* 2016 Jan;17(1):56-67. [doi: [10.1111/obr.12316](https://doi.org/10.1111/obr.12316)] [Medline: [26440472](https://pubmed.ncbi.nlm.nih.gov/26440472/)]
8. Martins VJB, Toledo Florêncio TMM, Grillo LP, et al. Long-lasting effects of undernutrition. *Int J Environ Res Public Health* 2011 Jun;8(6):1817-1846. [doi: [10.3390/ijerph8061817](https://doi.org/10.3390/ijerph8061817)] [Medline: [21776204](https://pubmed.ncbi.nlm.nih.gov/21776204/)]
9. Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. *Epidemiol Health* 2014;36:e2014009. [doi: [10.4178/epih/e2014009](https://doi.org/10.4178/epih/e2014009)] [Medline: [25078382](https://pubmed.ncbi.nlm.nih.gov/25078382/)]
10. Bradley J, Simpson E, Poliakov I, et al. Comparison of INTAKE24 (an online 24-h dietary recall tool) with interviewer-led 24-h recall in 11–24 year-old. *Nutrients* 2016 Jun 9;8(6):358. [doi: [10.3390/nu8060358](https://doi.org/10.3390/nu8060358)] [Medline: [27294952](https://pubmed.ncbi.nlm.nih.gov/27294952/)]
11. Lambert N, Plumb J, Looise B, et al. Using smart card technology to monitor the eating habits of children in a school cafeteria: 1. Developing and validating the methodology. *J Hum Nutr Diet* 2005 Aug;18(4):243-254. [doi: [10.1111/j.1365-277X.2005.00617.x](https://doi.org/10.1111/j.1365-277X.2005.00617.x)] [Medline: [16011560](https://pubmed.ncbi.nlm.nih.gov/16011560/)]
12. Cheah CSL, Barman S, Vu KTT, et al. Validation of a virtual reality buffet environment to assess food selection processes among emerging adults. *Appetite* 2020 Oct 1;153:104741. [doi: [10.1016/j.appet.2020.104741](https://doi.org/10.1016/j.appet.2020.104741)] [Medline: [32445771](https://pubmed.ncbi.nlm.nih.gov/32445771/)]
13. Persky S, Goldring MR, Turner SA, Cohen RW, Kistler WD. Validity of assessing child feeding with virtual reality. *Appetite* 2018 Apr 1;123:201-207. [doi: [10.1016/j.appet.2017.12.007](https://doi.org/10.1016/j.appet.2017.12.007)] [Medline: [29277518](https://pubmed.ncbi.nlm.nih.gov/29277518/)]
14. Smith AF, Baxter SD, Hardin JW, Guinn CH, Royer JA. Relation of children's dietary reporting accuracy to cognitive ability. *Am J Epidemiol* 2011 Jan 1;173(1):103-109. [doi: [10.1093/aje/kwq334](https://doi.org/10.1093/aje/kwq334)] [Medline: [21059806](https://pubmed.ncbi.nlm.nih.gov/21059806/)]
15. Thompson FE, Subar AF. Dietary assessment methodology. In: Coulston AM, Boushey CJ, Ferruzzi MG, Delahanty LM, editors. *Nutrition in the Prevention and Treatment of Disease*, 4th edition: Academic Press; 2017:5-48. [doi: [10.1016/B978-0-12-802928-2.00001-1](https://doi.org/10.1016/B978-0-12-802928-2.00001-1)]
16. Ortega RM, Pérez-Rodrigo C, López-Sobaler AM. Dietary assessment methods: dietary records. *Nutr Hosp* 2015 Feb 26;31 Suppl 3:38-45. [doi: [10.3305/nh.2015.31.sup3.8749](https://doi.org/10.3305/nh.2015.31.sup3.8749)] [Medline: [25719769](https://pubmed.ncbi.nlm.nih.gov/25719769/)]
17. Vereecken CA, Covents M, Sichert-Hellert W, et al. Development and evaluation of a self-administered computerized 24-h dietary recall method for adolescents in Europe. *Int J Obes (Lond)* 2008 Nov;32 Suppl 5:S26-S34. [doi: [10.1038/ijo.2008.180](https://doi.org/10.1038/ijo.2008.180)] [Medline: [19011650](https://pubmed.ncbi.nlm.nih.gov/19011650/)]
18. Adamo KB, Brett KE. Parental perceptions and childhood dietary quality. *Matern Child Health J* 2014 May;18(4):978-995. [doi: [10.1007/s10995-013-1326-6](https://doi.org/10.1007/s10995-013-1326-6)] [Medline: [23817727](https://pubmed.ncbi.nlm.nih.gov/23817727/)]
19. Klesges RC, Stein RJ, Eck LH, Isbell TR, Klesges LM. Parental influence on food selection in young children and its relationships to childhood obesity. *Am J Clin Nutr* 1991 Apr;53(4):859-864. [doi: [10.1093/ajcn/53.4.859](https://doi.org/10.1093/ajcn/53.4.859)] [Medline: [2008864](https://pubmed.ncbi.nlm.nih.gov/2008864/)]
20. Fisher JO, Birch LL. Restricting access to palatable foods affects children's behavioral response, food selection, and intake. *Am J Clin Nutr* 1999 Jun;69(6):1264-1272. [doi: [10.1093/ajcn/69.6.1264](https://doi.org/10.1093/ajcn/69.6.1264)] [Medline: [10357749](https://pubmed.ncbi.nlm.nih.gov/10357749/)]
21. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". Presented at: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek 2011); Sep 28-30, 2011. [doi: [10.1145/2181037.2181040](https://doi.org/10.1145/2181037.2181040)]
22. Caponetto I, Earp J, Ott M. Gamification and education: a literature review. Presented at: Proceedings of the 8th European Conference on Games Based Learning (ECGBL 2014); Oct 9-10, 2014 URL: <https://iris.cnr.it/handle/20.500.14243/226324?mode=complete> [accessed 2026-02-13]
23. Muangsrinoon S, Boonbrahm P. Game elements from literature review of gamification in healthcare context. *J Technol Sci Educ* 2019;9(1):20. [doi: [10.3926/jotse.556](https://doi.org/10.3926/jotse.556)]
24. AlMarshedi A, Wills GB, Ranchhod A. The Wheel of Sukr: a framework for gamifying diabetes self-management in Saudi Arabia. 2015 Presented at: The 5th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2015); Sep 27-30, 2015; Berlin, Germany p. 475-480. [doi: [10.1016/j.procs.2015.08.370](https://doi.org/10.1016/j.procs.2015.08.370)]
25. Keung C, Lee A, Lu S, O'Keefe M. BunnyBolt: a mobile fitness app for youth. Presented at: Proceedings of the 12th International Conference on Interaction Design and Children (IDC 2013); Jun 24-27, 2013. [doi: [10.1145/2485760.2485871](https://doi.org/10.1145/2485760.2485871)]
26. Dennis TA, O'Toole L. Mental health on the go: effects of a gamified attention bias modification mobile application in trait anxious adults. *Clin Psychol Sci* 2014 Sep 1;2(5):576-590. [doi: [10.1177/2167702614522228](https://doi.org/10.1177/2167702614522228)] [Medline: [26029490](https://pubmed.ncbi.nlm.nih.gov/26029490/)]
27. de Vlieger NM, Sainsbury L, Smith SP, et al. Feasibility and acceptability of 'VitaVillage': a serious game for nutrition education. *Nutrients* 2021 Dec 31;14(1):189. [doi: [10.3390/nu14010189](https://doi.org/10.3390/nu14010189)] [Medline: [35011063](https://pubmed.ncbi.nlm.nih.gov/35011063/)]
28. Beltran A, O'Connor T, Hughes S, et al. Alpha test of a videogame to increase children's vegetable consumption. *Games Health J* 2012 Jun;1(3):219-222. [doi: [10.1089/g4h.2011.0027](https://doi.org/10.1089/g4h.2011.0027)] [Medline: [24761315](https://pubmed.ncbi.nlm.nih.gov/24761315/)]
29. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. *Bull World Health Organ* 2007 Sep;85(9):660-667. [doi: [10.2471/blt.07.043497](https://doi.org/10.2471/blt.07.043497)] [Medline: [18026621](https://pubmed.ncbi.nlm.nih.gov/18026621/)]
30. Tylka TL, Kroon Van Diest AM. The intuitive eating scale-2: item refinement and psychometric evaluation with college women and men. *J Couns Psychol* 2013 Jan;60(1):137-153. [doi: [10.1037/a0030893](https://doi.org/10.1037/a0030893)] [Medline: [23356469](https://pubmed.ncbi.nlm.nih.gov/23356469/)]
31. Ding Y, Yang Y, Li F, et al. Development and validation of a photographic atlas of food portions for accurate quantification of dietary intakes in China. *J Hum Nutr Diet* 2021 Jun;34(3):604-615. [doi: [10.1111/jhn.12844](https://doi.org/10.1111/jhn.12844)] [Medline: [33406287](https://pubmed.ncbi.nlm.nih.gov/33406287/)]

32. Yang Y. China Food Composition Tables [Book in Chinese], 6th edition: Peking University Medical Press; 2018, Vol. 1. URL: <http://jx.tsxfw.com/bookshow.asp?id=1516720> [accessed 2026-02-14]
33. Chinese Nutrition Society. Dietary Reference Intakes for China [Book in Chinese]: People's Medical Publishing House; 2023. URL: <https://baike.baidu.com/en/item/Dietary%20Reference%20Intakes%20for%20Chinese%20Residents%3A%202023%20Edition/51892> [accessed 2026-02-14]
34. Growth reference data for 5-19 years. World Health Organization. 2006. URL: <https://www.who.int/tools/growth-reference-data-for-5to19-years> [accessed 2025-02-20]
35. Wang ZH, Zhang B, Wang HJ, et al. Trend in meat consumption patterns among Chinese children in nine provinces (autonomous region) aged 6 to 17 years between 1991 and 2011 [Article in Chinese]. Zhonghua Yu Fang Yi Xue Za Zhi 2013 Sep;47(9):826-831. [doi: [10.3760/cma.j.issn.0253-9624.2013.09.012](https://doi.org/10.3760/cma.j.issn.0253-9624.2013.09.012)] [Medline: [24351564](https://pubmed.ncbi.nlm.nih.gov/24351564/)]
36. Wang Y, Beydoun MA. Meat consumption is associated with obesity and central obesity among US adults. Int J Obes (Lond) 2009 Jun;33(6):621-628. [doi: [10.1038/ijo.2009.45](https://doi.org/10.1038/ijo.2009.45)] [Medline: [19308071](https://pubmed.ncbi.nlm.nih.gov/19308071/)]
37. Ventura AK, Worobey J. Early influences on the development of food preferences. Curr Biol 2013 May 6;23(9):R401-R408. [doi: [10.1016/j.cub.2013.02.037](https://doi.org/10.1016/j.cub.2013.02.037)] [Medline: [23660363](https://pubmed.ncbi.nlm.nih.gov/23660363/)]
38. Lopez-Cepero A, Frisard CF, Lemon SC, Rosal MC. Association between emotional eating, energy-dense foods and overeating in Latinos. Eat Behav 2019 Apr;33:40-43. [doi: [10.1016/j.eatbeh.2019.03.001](https://doi.org/10.1016/j.eatbeh.2019.03.001)] [Medline: [30889426](https://pubmed.ncbi.nlm.nih.gov/30889426/)]
39. Nguyen-Michel ST, Unger JB, Spruijt-Metz D. Dietary correlates of emotional eating in adolescence. Appetite 2007 Sep;49(2):494-499. [doi: [10.1016/j.appet.2007.03.005](https://doi.org/10.1016/j.appet.2007.03.005)] [Medline: [17466408](https://pubmed.ncbi.nlm.nih.gov/17466408/)]
40. Fiese BH, Schwartz M. Reclaiming the family table: mealtimes and child health and wellbeing. Soc Policy Rep 2008;22(4):1-20. [doi: [10.1002/j.2379-3988.2008.tb00057.x](https://doi.org/10.1002/j.2379-3988.2008.tb00057.x)]
41. Roos E, Pajunen T, Ray C, et al. Does eating family meals and having the television on during dinner correlate with overweight? A sub-study of the PRO GREENS project, looking at children from nine European countries. Public Health Nutr 2014 Nov;17(11):2528-2536. [doi: [10.1017/S1368980013002954](https://doi.org/10.1017/S1368980013002954)] [Medline: [24642340](https://pubmed.ncbi.nlm.nih.gov/24642340/)]
42. Gingras V, Rifas-Shiman SL, Taveras EM, Oken E, Hivert MF. Dietary behaviors throughout childhood are associated with adiposity and estimated insulin resistance in early adolescence: a longitudinal study. Int J Behav Nutr Phys Act 2018 Dec 17;15(1):129. [doi: [10.1186/s12966-018-0759-0](https://doi.org/10.1186/s12966-018-0759-0)] [Medline: [30558613](https://pubmed.ncbi.nlm.nih.gov/30558613/)]
43. Cameron JD, Maras D, Sigal RJ, et al. The mediating role of energy intake on the relationship between screen time behaviour and body mass index in adolescents with obesity: the HEARTY study. Appetite 2016 Dec 1;107:437-444. [doi: [10.1016/j.appet.2016.08.101](https://doi.org/10.1016/j.appet.2016.08.101)] [Medline: [27545672](https://pubmed.ncbi.nlm.nih.gov/27545672/)]
44. Ramos E, Costa A, Araújo J, Severo M, Lopes C. Effect of television viewing on food and nutrient intake among adolescents. Nutrition 2013;29(11-12):1362-1367. [doi: [10.1016/j.nut.2013.05.007](https://doi.org/10.1016/j.nut.2013.05.007)] [Medline: [24103514](https://pubmed.ncbi.nlm.nih.gov/24103514/)]
45. Fulkerson JA, Friend S, Horning M, et al. Family home food environment and nutrition-related parent and child personal and behavioral outcomes of the Healthy Home Offerings via the Mealtime Environment (HOME) Plus Program: a randomized controlled trial. J Acad Nutr Diet 2018 Feb;118(2):240-251. [doi: [10.1016/j.jand.2017.04.006](https://doi.org/10.1016/j.jand.2017.04.006)] [Medline: [28578900](https://pubmed.ncbi.nlm.nih.gov/28578900/)]
46. Fulkerson JA, Neumark-Sztainer D, Story M, et al. The Healthy Home Offerings via the Mealtime Environment (HOME) Plus study: design and methods. Contemp Clin Trials 2014 May;38(1):59-68. [doi: [10.1016/j.cct.2014.01.006](https://doi.org/10.1016/j.cct.2014.01.006)] [Medline: [24480729](https://pubmed.ncbi.nlm.nih.gov/24480729/)]
47. Riasatian M, Karimian Z, Mohsenpour MA, Faghhih S. Effects of nutrition education via a mobile-based game on nutritional knowledge, dietary habits and anthropometric indices in preschool children: a study protocol. MethodsX 2023;10:102104. [doi: [10.1016/j.mex.2023.102104](https://doi.org/10.1016/j.mex.2023.102104)] [Medline: [36970028](https://pubmed.ncbi.nlm.nih.gov/36970028/)]

Abbreviations

- 24HR:** 24-hour dietary recall
AMDR: acceptable macronutrient distribution range
BMI z scores: BMI-for-age z scores
EES: emotional eating scale
GDA: gamified dietary assessment tool

Edited by A Coristine; submitted 26.Feb.2025; peer-reviewed by A Moyen, M Boaz; accepted 30.Jan.2026; published 13.Mar.2026.

Please cite as:

Zhang Y, Lin T, Tan Y, Shi X, Yang Y, Yang H, Na X, Zhang J, Zhao A

Development of a Novel Dietary Assessment Method Using Gamification Concepts: Exploratory and Application Study

JMIR Serious Games 2026;14:e72387

URL: <https://games.jmir.org/2026/1/e72387>

doi: [10.2196/72387](https://doi.org/10.2196/72387)

© Yulin Zhang, Tianwu Lin, Yuefeng Tan, Xiaojin Shi, Yucheng Yang, Haibing Yang, Xiaona Na, Junhan Zhang, Ai Zhao. Originally published in JMIR Serious Games (<https://games.jmir.org>), 13.Mar.2026. 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.

Longitudinal Association Between Internet Gaming Disorder and School Refusal Among Adolescents Using a Random Intercept Cross-Lagged Panel Model: Three-Wave Prospective Cohort Study

Peng Zheng* ; Zi Tao* ; Luoxiang Fang* ; Tao Yao; Lingling Zhou; Yutian Wang; Yuannan Lin

Second Affiliated Hospital & Yuying Children's Hospital of Wenzhou Medical University, No. 109 Xueyuan West Road, Wenzhou, Zhejiang, China

*these authors contributed equally

Corresponding Author:

Yuannan Lin

Second Affiliated Hospital & Yuying Children's Hospital of Wenzhou Medical University, No. 109 Xueyuan West Road, Wenzhou, Zhejiang, China

Abstract

Background: Internet gaming disorder (IGD) and school refusal are increasingly prevalent during adolescence, yet limited research has examined how they influence each other over time. Moreover, it is unclear whether the association differs by sex.

Objective: This study aimed to investigate the within-person longitudinal associations between school refusal and IGD and potential sex differences in these associations among Chinese adolescent gamers.

Methods: A 3-wave prospective cohort longitudinal study was conducted among students from 4 middle schools in Zhejiang Province, China. Adolescents who had played online games in the past 12 months were involved. Participants were recruited using convenience sampling. Data were collected at 6-month intervals: time 1 (T1, March 2024), time 2 (T2, September 2024), and time 3 (T3, March 2025). A total of 918 Chinese adolescents ($n=609$, 66.3% male; mean 13.55, SD 0.75) participated in the 3-wave survey. IGD and school refusal were assessed using the 9-item DSM-5 (*Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*) IGD checklist and the School Refusal Assessment Scale-Revised. A random intercept cross-lagged panel model and multigroup analyses were conducted.

Results: All statistical tests were 2-tailed with $\alpha=.05$. The random intercept cross-lagged panel model demonstrated good fit to the data ($\chi^2_5=16.02$; comparative fit index 0.970; root mean square error of approximation 0.035). Cross-lagged effects indicated a reciprocal association between school refusal and IGD. School refusal predicted later IGD (T1→T2: $\beta=.06$, 95% CI .02-.08; $P=.02$ and T2→T3: $\beta=.06$, 95% CI .02-.08; $P=.02$), and IGD predicted subsequent school refusal (T1→T2: $\beta=.12$, 95% CI .09-.15; $P<.001$ and T2→T3: $\beta=.13$, 95% CI .08-.16; $P<.001$). Multigroup analyses revealed significant sex differences. Wald tests showed sex-specific effects: among females, school refusal predicted later IGD (T1→T2: $\beta=.10$, 95% CI .05-.15; $P=.006$ and T2→T3: $\beta=.11$, 95% CI .04-.17; $P=.004$), whereas these paths were nonsignificant for males. IGD predicted later school refusal in both sexes, but the effect was stronger for males (T1→T2: $\beta=.15$, 95% CI .12-.18; $P<.001$ and T2→T3: $\beta=.15$, 95% CI .12-.18; $P<.001$) than females (T1→T2: $\beta=.07$, 95% CI .03-.10; $P=.02$ and T2→T3: $\beta=.07$, 95% CI .03-.10; $P=.02$).

Conclusions: This study provides novel insights into the reciprocal within-person associations between school refusal and IGD among adolescents. By accounting for stable between-person differences, the model clarifies how within-person fluctuations in one behavior relate to subsequent changes in the other. Sex-specific patterns were observed: school refusal predicted subsequent IGD among females, whereas IGD predicted later school refusal more strongly among males. These findings highlight the dynamic interplay between school disengagement and problematic gaming and underscore the importance of sex-sensitive intervention strategies.

(*JMIR Serious Games* 2026;14:e89619) doi:[10.2196/89619](https://doi.org/10.2196/89619)

KEYWORDS

adolescent gamers; internet gaming disorder; school refusal; reciprocal association; sex differences

Introduction

Internet gaming disorder (IGD) refers to persistent and maladaptive engagement in online gaming that leads to

significant functional impairment and is increasingly recognized as a pressing public health concern [1]. IGD is characterized by impaired control over gaming, prioritization of gaming over other responsibilities, and sustained engagement despite negative

consequences [2]. Prevalence estimates highlight its public health relevance during adolescence, with meta-analyses reporting pooled rates ranging from 6.7% to 8.8% [3,4], and studies across different regions document similar elevated levels [5]. In China, a recent meta-analysis reported an IGD prevalence of approximately 10% [6]. IGD has been associated with numerous psychological and social difficulties, including depressive symptoms, loneliness, aggressive behaviors, and strained family relationships [7-10]. These patterns highlight that IGD encompasses broader developmental risks beyond excessive gaming alone. Given its multifaceted impact, understanding the processes that may reinforce or exacerbate IGD is essential. One understudied yet potentially important factor is school refusal, which may interact dynamically with IGD and form reciprocal influences over time, particularly during adolescence.

School refusal refers to difficulties in attending or remaining in school, including tardiness, skipping classes, morning misbehaviors designed to avoid school, and attending school with significant distress or reluctance [11,12]. School refusal is often accompanied by marked negative emotions such as anxiety, fear, and emotional distress [13]. According to the Functional Behavioral Model [14], school refusal is maintained by 4 key reinforcement-based functions: avoidance of negative affectivity (ANA), escape from aversive situations (EAS), attention-seeking (AS), and PTR. Existing evidence suggests that school refusal is highly prevalent, with estimates ranging from 28% to 35% among children and adolescents in Western and global samples [15] and from 22.5% to 30% in various regions of China [16]. Importantly, adolescents who exhibit school refusal are more likely than their peers to develop subsequent behavioral and adjustment difficulties [17,18].

School refusal is closely related to IGD among adolescents. However, the directionality between these 2 behaviors remains unclear. Given these uncertainties, it is necessary to consider how school refusal and IGD might relate to each other in both directions. First, school refusal may increase IGD. According to escape-avoidance theory, individuals tend to withdraw from environments that elicit stress or negative affect [19]. Adolescents who experience academic pressure, social-evaluative concerns, or negative emotional states at school often show higher levels of school refusal [20]. Online gaming, which offers immersion, controllability, and immediate relief from distress, may thus become an attractive alternative to aversive school experiences, thereby school refusal and increased gaming engagement within a shared pattern of avoidance-oriented coping [21]. In other words, when school elicits persistent distress, gaming may function as an emotionally soothing and controllable environment, reinforcing avoidance tendencies and increasing adolescents' susceptibility to IGD.

Second, IGD may similarly increase school refusal. According to self-determination theory [22], individuals have 3 basic psychological needs: autonomy, competence, and belonging. Gaming environments frequently satisfy these needs through structured challenges, opportunities for volitional action, and social interaction [23,24]. When these experiences are perceived as more rewarding than those available in school, adolescents may experience lower motivation or diminished engagement

in academic settings [25]. Sustained involvement in gaming may further weaken school-related motivation by displacing academic goals, reducing tolerance for effortful learning, and amplifying the contrast between gratifying digital environments and demanding school contexts. Under such conditions, stronger gaming involvement may be accompanied by increased tendencies toward school disengagement or avoidance. Together, these theoretical perspectives suggest that school refusal and IGD may mutually reinforce each other over time.

Empirical findings may provide support for this potential bidirectional association between school refusal and IGD. Consistent with theoretical expectations, studies across different cultural contexts show that adolescents who struggle with school engagement are more likely to rely on gaming or digital media. For example, digital game addiction was positively associated with school refusal among adolescents in Turkey [21], and excessive internet engagement was strongly related to both truancy and medically certified absences in a large Finnish national sample of more than 86,000 students [26]. Research also indicates that problematic internet use is a salient external correlate of school refusal; a latent class study among Chinese adolescents found that youth with school refusal behaviors frequently use gaming or online activities to escape academic and social pressures [27]. Additional evidence from Japan further shows that problematic internet use is highly prevalent among adolescents clinically identified with school refusal [28]. Conversely, research on school-related protective factors shows that stronger school identification predicts lower IGD risk [29]. This finding indirectly suggests that weakened school connectedness, which is often observed among youth with school refusal, may heighten vulnerability to gaming-related problems.

Reviews and meta-analytic work also highlight consistent associations between problematic internet use and school-related outcomes. A recent review synthesizing 10 studies concluded that problematic internet use represents a risk factor for school refusal [13], although no meta-analysis to date has examined school refusal and IGD specifically. Longitudinal findings further illustrate that digital behaviors are relevant to school functioning: problematic social media use and problematic smartphone use have been associated with declines in academic performance [30,31], problematic social network site use has predicted subsequent school absenteeism [32], and problematic gaming has been prospectively linked to increased truancy [33]. Despite these findings, no longitudinal studies have directly examined the potential bidirectional association between school refusal and IGD. This gap highlights the need for research that clarifies how school refusal and IGD may co-develop over time.

Emerging evidence suggests that the association between school refusal and IGD across developmental pathways may differ by sex. Epidemiological studies consistently show that male adolescents exhibit a higher prevalence of IGD than females [34,35]. Furthermore, male participants with IGD tend to be more sensitive to gaming-related rewards, whereas female participants show stronger affective dysregulation and mood-related symptoms [36]. These sex-specific patterns in gaming motivation and emotional processes may contribute to differential pathways linking school-related distress and

problematic gaming behavior. At the same time, sex differences have also been documented in school attendance problems. For instance, female students generally report higher levels of school nonattendance compared with male students [37]. Evidence from a latent profile analysis further indicates that although most school refusal profiles show no significant sex differences, females are overrepresented in the mixed school refusal behavior profile, albeit with a small effect size [38]. This suggests that while sex disparities in school refusal are small, females may be more likely to exhibit complex or multifaceted forms of school refusal.

Although no study has explicitly tested whether sex moderates the association between school refusal and IGD, the existing evidence provides a plausible basis for sex-specific pathways. School-related distress may have a stronger impact on IGD among females due to heightened emotional vulnerability [39], whereas IGD may more strongly contribute to school-related difficulties among males, given their higher reward sensitivity and greater IGD prevalence [40]. However, it remains unclear whether these potential sex-specific patterns develop over time. This study, therefore, examines sex differences in the longitudinal association between school refusal and IGD, with the aim of informing sex-sensitive prevention strategies and advancing understanding of underlying developmental processes.

To advance the understanding of how school refusal and IGD develop over time, the present study used a 3-wave longitudinal design among Chinese adolescent gamers. A random intercept cross-lagged panel model (RI-CLPM) was applied to disentangle stable between-person differences from within-person fluctuations. Focusing on the within-person level, this study aimed to clarify the reciprocal relationship between school refusal and IGD. Based on previous empirical findings and theoretical perspectives, three hypotheses were proposed: (H1) Increases in school refusal would predict higher IGD at the next time point. (H2) Higher IGD would predict increases in school refusal over time. (H3) The reciprocal within-person association between school refusal and IGD would differ between male and female adolescent gamers. Findings are expected to provide deeper evidence on how short-term changes in school refusal and IGD develop and to inform sex-sensitive prevention strategies.

Methods

Study Design, Participants, and Setting

This study was designed as a 1-year prospective cohort study and conducted among students from 4 middle schools in Zhejiang Province, China. Data were collected at three 6-month intervals: time 1 (T1, March 2024), time 2 (T2, September 2024), and time 3 (T3, March 2025). The target population comprised adolescents who had played online games in the past 12 months. The participating schools were selected using convenience sampling. In collaboration with school administrators, all eligible students in the selected classes were invited to participate. Trained field workers administered the survey during regular classroom sessions. Prior to completing the questionnaire, the field workers briefed the students on the

objectives and procedures of the study, assured them of the anonymity and confidentiality of their responses, and emphasized the voluntary nature of participation. Parents or legal guardians were informed about the study in advance.

To enable linkage of repeated responses across the 3 assessment waves, students provided their school identification numbers, which were subsequently encoded for matching purposes. These identifiers were permanently removed from the analytical dataset after linkage to ensure participant anonymity.

At baseline (T1), 983 adolescents who met the inclusion criteria completed the survey. By the final wave (T3), 918 participants remained in the cohort ($n=609$, 66.3% male; mean 13.55, SD 0.75). Among the 918 participants who completed all 3 waves, baseline demographic variables had the following missing values: sex (2/918, 0.2%), father's educational level (8/918, 0.9%), mother's educational level (5/918, 0.5%), self-reported academic performance (9/918, 1.0%), and single-parent family status (3/918, 0.3%). There were no missing data for perceived family financial situation or gaming time per month. A test for missing completely at random indicated that the missingness was completely random.

Ethical Considerations

Prior to each wave of data collection, detailed information regarding the aims, procedures, potential risks, and confidentiality protections of the study was provided to both students and their parents or legal guardians. Written informed consent was obtained from parents or legal guardians, and written assent was obtained from the students. Participation was entirely voluntary, and it was clearly communicated that students had the right to decline or withdraw from the study at any time without any academic or administrative consequences. All questionnaires were self-administered and completed anonymously. An information sheet outlining participants' rights and data protection measures was distributed before survey administration. No incentive was provided to the students. No identifiable images of participants were collected or used in this study. The study protocol and consent procedures were reviewed and approved by the Institutional Ethics Committee of the Second Affiliated Hospital & Yuying Children's Hospital of Wenzhou Medical University (approval 2024-K-239 - 02).

Studied Variables and Measurements

Background Factors

Background factors were collected, including age, sex, parental educational level, self-reported academic performance, perceived family financial situations, single-parent family status, and gaming time per month.

IGD

IGD was assessed using the 9-item *DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition)* IGD checklist [1]. The checklist captures the presence of IGD symptoms over the past 12 months (yes/no response). The 9 criteria include preoccupation with gaming, withdrawal, tolerance, loss of control, giving priority to gaming over other activities, persistence despite negative consequences, deception regarding gaming time, escapism, and functional impairment

due to gaming. It has been validated among adolescents and demonstrated satisfactory psychometric properties [41]. In this study, the Cronbach α of IGD was 0.78 at T1, 0.84 at T2, and 0.81 at T3.

School Refusal

The School Refusal Assessment Scale-Revised (SRAS-R) was used to assess school refusal [11]. It is a 24-item scale comprising 4 dimensions: ANA, EAS, AS, and PTR. Sample items include the following: “How often do you have bad feelings about going to school because you are afraid of something related to school (eg, tests, school bus, teacher, fire alarm)?,” “How often do you stay away from school because it is hard to speak with the other kids at school?,” “How often do you feel you would rather be with your parents than go to school?,” and “When you are not in school during the week (Monday to Friday), how often do you leave the house and do something fun?.” The items are rated on a 7-point Likert scales (1=never to 7=always). It has been validated among adolescents and demonstrated satisfactory psychometric properties [42]. In this study, the Cronbach α of school refusal was 0.88 at T1, 0.90 at T2, and 0.94 at T3.

Data Analysis

Attrition analyses were conducted to compare participants who completed all 3 waves with those who discontinued participation at follow-up. Independent-sample *t* tests were used to examine sex differences in the 4 school refusal dimensions and IGD, and Pearson correlations were performed among the key study variables. These preliminary analyses were performed in IBM SPSS Statistics (version 28.0) using listwise deletion, as they were primarily descriptive and exploratory in nature and aimed to summarize observed data based on complete cases.

To evaluate the temporal stability of the measurement models, longitudinal measurement invariance tests were conducted for both school refusal and IGD. Criteria for acceptable fit followed established recommendations [43], including comparative fit index (CFI) and Tucker-Lewis Index (TLI) values ≥ 0.90 and root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) values ≤ 0.08 . Metric and scalar invariance were assessed by examining changes in model fit across nested models, with $\Delta\text{CFI} \leq 0.01$ and $\Delta\text{RMSEA} \leq 0.015$ indicating invariance. Results supported longitudinally consistent factor structures, confirming both metric and scalar invariance across waves.

The reciprocal associations between school refusal and IGD were examined using RI-CLPM. In these models, school refusal was specified as a latent construct indicated by the 4 functional dimensions of the SRAS-R: ANA, EAS, AS, and PTR. The 4 subscale scores were used as observed indicators of a higher-order latent school refusal factor at each wave, consistent with the functional behavioral model [14] of school refusal and

allowing the shared variance across dimensions to be captured at the construct level. IGD was operationalized as a continuous symptom count score by summing the 9 DSM-5 IGD checklist items, which is a commonly adopted approach in IGD research [44-46]. Prior to model estimation, observed variables were grand-mean centered so that intercepts reflected overall mean levels across the 3 assessment waves. Model specification began with an unconstrained RI-CLPM in which all autoregressive and cross-lagged paths were freely estimated across time. To obtain a more parsimonious and stable model, cross-lagged paths were then constrained to be equal across adjacent time intervals. In a subsequent step, both autoregressive and cross-lagged paths were constrained to be time-invariant, and model fit was compared across nested models to identify the optimal specification. After identifying the optimal model, covariates were incorporated as predictors of the between-person random intercepts, including age, sex, self-reported academic performance, perceived family financial status, and single-parent family status. Gaming time per month was included as a time-varying covariate at the within-person level, such that gaming time at each wave predicted the within-person deviations of school refusal and IGD at the corresponding time point, thereby controlling for concurrent fluctuations in gaming behavior.

The multigroup RI-CLPM approach was adopted to test potential sex differences in the longitudinal association between school refusal and IGD. An unconstrained model allowing all structural paths to vary across groups was compared with a model in which the cross-lagged effects were constrained to equality. A significant χ^2 difference indicated the presence of sex-specific effects.

All structural equation models, including the RI-CLPM and multigroup models, were estimated using Mplus version 8.3 (Muthén & Muthén). Full information maximum likelihood estimation was used to handle missing data in these models, given its advantages in longitudinal structural equation modeling for producing unbiased parameter estimates and maximizing data utilization under the missing at random assumption [47]. Statistical significance was set at 2-tailed $P < .05$ for all analyses. This study was reported in accordance with the American Psychological Association Reporting Standards for Studies Using Structural Equation Modeling [48].

Results

Attrition Analyses

As shown in Table 1, the 2 groups did not differ significantly with respect to sex, father’s and mother’s educational level, self-reported academic performance, perceived family financial status, single-parent status, gaming time per month, age, ANA, EAS, AS, PTR, and IGD.

Table . Attrition analyses comparing adolescent gamers who completed all 3 waves of the longitudinal study and those lost to follow-up^a.

Variables	Follow-up (n=918)	Lost to follow-up (n=65)	P value
Categorical variables, n (%)			
Sex			.38
Male	609 (66.3)	45 (69.2)	
Female	307 (33.4)	20 (30.8)	
Missing data	2 (0.2)	— ^b	
Father's educational level			.57
Middle school or below	630 (68.6)	45 (69.2)	
High school or equal level	208 (22.7)	17 (26.2)	
University or above	72 (7.8)	3 (4.6)	
Missing data	8 (0.9)	—	
Mother's educational level			.62
Middle school or below	656 (71.5)	47 (72.3)	
High school or equal level	171 (18.6)	14 (21.5)	
University or above	86 (9.4)	4 (6.2)	
Missing data	5 (0.5)	—	
Self-reported academic performance			.60
Bottom 20%	131 (14.3)	7 (10.8)	
21st-40th percentile	188 (20.5)	16 (24.6)	
41st-60th percentile	226 (24.6)	18 (27.7)	
61st-80th percentile	213 (23.2)	11 (16.9)	
Top 20%	151 (16.4)	13 (20)	
Missing data	9 (1)	—	
Perceived family financial situation			.36
Below average	119 (13)	9 (13.8)	
Average	622 (67.8)	39 (60)	
Above average	177 (19.3)	17 (26.2)	
Single-parent family status			.38
Yes	71 (7.7)	2 (3.1)	
No	682 (74.3)	51 (78.5)	
Choose not to report	162 (17.6)	12 (18.5)	
Missing data	3 (0.3)	—	
Gaming time per month			.59
<4 h	406 (44.2)	28 (43.1)	
4-8 h	257 (28)	19 (29.2)	
8-12 h	105 (11.4)	4 (6.2)	
12-16 h	93 (10.1)	8 (12.3)	
>16 h	57 (6.2)	6 (9.2)	
Continuous variables, mean (SD)			
Age (years)	13.55 (0.75)	13.65 (0.76)	.65
ANA ^c	12.43 (5.99)	12.55 (6.57)	.87

Variables	Follow-up (n=918)	Lost to follow-up (n=65)	<i>P</i> value
EAS ^d	10.21 (3.72)	9.77 (3.5)	.35
AS ^e	11.71 (5.87)	10.52 (5.55)	.12
PTR ^f	14.97 (6.36)	14.85 (6.5)	.88
IGD ^g	1.51 (1.72)	1.88 (2.16)	.10

^aANA, EAS, AS, and PTR are 4 dimensions of school refusal.

^bNot available.

^cANA: avoidance of negative affectivity.

^dEAS: escape from aversive situations.

^eAS: attention-seeking.

^fPTR: pursuit of tangible rewards.

^gIGD: internet gaming disorder.

Sex Differences in Main Variables

Male participants showed significantly lower mean scores (SD) than female participants across the 4 dimensions of school

refusal across the 3 waves, except EAS at T1. In contrast, male participants showed significantly higher mean scores (SDs) than female participants in IGD across the 3 waves (Table 2).

Table . Sex differences in school refusal dimensions and internet gaming disorder (IGD) across 3 waves among Chinese adolescent gamers^a.

Dimensions	Boys, mean (SD)	Girls, mean (SD)	<i>t</i> test (<i>df</i>)	<i>P</i> value	Cohen <i>d</i>
ANA ^b at T1	11.86 (5.71)	13.48 (6.27)	-3.92 (914)	<.001	-0.27
EAS ^c at T1	10.08 (3.64)	10.49 (3.87)	-1.57 (914)	.12	-0.11
AS ^d at T1	11.38 (5.80)	12.28 (5.83)	-2.21 (914)	.03	-0.15
PTR ^e at T1	14.57 (6.38)	15.70 (6.16)	-2.54 (914)	.01	-0.18
ANA at T2	11.60 (6.16)	14.21 (6.92)	-5.81 (914)	<.001	-0.40
EAS at T2	9.69 (3.73)	10.54 (4.11)	-3.15 (914)	.002	-0.22
AS at T2	10.52 (5.63)	12.00 (5.94)	-3.68 (914)	<.001	-0.26
PTR at T2	14.49 (6.68)	16.53 (7.07)	-4.28 (914)	<.001	-0.30
ANA at T3	12.16 (7.16)	15.64 (8.28)	-6.60 (914)	<.001	-0.45
EAS at T3	10.09 (5.05)	11.30 (5.10)	-3.42 (914)	<.001	-0.24
AS at T3	10.82 (6.50)	13.23 (6.94)	-5.16 (914)	<.001	-0.36
PTR at T3	14.98 (7.76)	17.21 (7.83)	-4.08 (914)	<.001	-0.29
IGD at T1	1.71 (1.80)	1.10 (1.46)	5.14 (914)	<.001	0.37
IGD at T2	1.62 (1.95)	1.19 (1.45)	3.44 (914)	<.001	0.25
IGD at T3	1.91 (2.31)	1.33 (1.71)	3.95 (914)	<.001	0.29

^aANA, EAS, AS, and PTR are 4 dimensions of school refusal.

^bANA: avoidance of negative affectivity.

^cEAS: escape from aversive social-evaluative situations.

^dAS: attention-seeking.

^e PTR: pursuit of tangible rewards.

Measurement Invariance Tests

Table 3 presents the results of longitudinal measurement invariance testing for school refusal and IGD. The configural invariance models again demonstrated acceptable fit, with CFI and TLI exceeding 0.90 and RMSEA and SRMR below 0.08, indicating that the basic factor structures of both constructs were

stable across the 3 waves. Metric invariance was supported, as changes in fit indices ($\Delta\text{CFI} \leq 0.01$; $\Delta\text{RMSEA} \leq 0.015$) fell within recommended thresholds, suggesting that the factor loadings remained consistent over time. Scalar invariance was also established, with minimal declines in model fit ($\Delta\text{CFI} \leq 0.01$; $\Delta\text{RMSEA} \leq 0.015$), indicating equivalence of item intercepts across measurement occasions.

Table . Longitudinal measurement invariance test for school refusal and internet gaming disorder (IGD).

Variables and model	CFI ^a	TLI ^b	RMSEA ^c	SRMR ^d	ΔCFI	ΔRMSEA	ΔSRMR
SR^e							
Configural invariance	0.957	0.947	0.043	0.034	— ^f	—	—
Metric invariance	0.952	0.941	0.043	0.038	0.005	0.000	0.004
Scalar invariance	0.948	0.937	0.045	0.041	0.004	0.002	0.003
IGD							
Configural invariance	0.942	0.931	0.038	0.039	—	—	—
Metric invariance	0.937	0.929	0.038	0.042	0.005	0.000	0.003
Scalar invariance	0.935	0.921	0.041	0.044	0.002	0.003	0.002

^aCFI: comparative fit index.

^bTLI: Tucker-Lewis Index.

^cRMSEA: root mean square error of approximation.

^dSRMR: standardized root mean square residual.

^eSR: school refusal.

^fNot available.

Measurement invariance across sex was examined for school refusal and IGD. The configural invariance models showed acceptable fit, with CFI values above 0.90 and RMSEA and SRMR values below 0.08, indicating similar factor structures for males and females. Constraining factor loadings across groups resulted in minimal changes in model fit ($\Delta\text{CFI} \leq 0.01$; $\Delta\text{RMSEA} \leq 0.015$), supporting metric invariance. Scalar invariance was further supported ($\Delta\text{CFI} \leq 0.01$; $\Delta\text{RMSEA} \leq 0.015$), suggesting that item intercepts were comparable across sex. These results indicate that observed sex differences reflect

substantive variation rather than measurement bias. Corresponding results are presented in Table S1 in [Multimedia Appendix 1](#).

Correlations

As shown in [Table 4](#), the 4 dimensions (ANA, EAS, AS, and PTR) of school refusal were positively intercorrelated across the 3 waves, with r ranging from 0.22 to 0.69 (all $P < .001$). Each dimension of school refusal was also significantly and positively associated with IGD at all waves, with r ranging from 0.21 to 0.40 (all $P < .001$).

Table . Pearson correlation coefficients among school refusal dimensions and internet gaming disorder (IGD) measured at 3 time points^a.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. ANA ^b at T1	1														
2. EAS ^c at T1	0.63 ^d	1													
3. AS ^e at T1	0.66 ^d	0.42 ^d	1												
4. PTR ^f at T1	0.52 ^d	0.42 ^d	0.43 ^d	1											
5. ANA at T2	0.56 ^d	0.36 ^d	0.44 ^d	0.36 ^d	1										
6. EAS at T2	0.38 ^d	0.42 ^d	0.26 ^d	0.24 ^d	0.66 ^d	1									
7. AS at T2	0.46 ^d	0.23 ^d	0.58 ^d	0.28 ^d	0.69 ^d	0.49 ^d	1								
8. PTR at T2	0.33 ^d	0.22 ^d	0.27 ^d	0.51 ^d	0.56 ^d	0.48 ^d	0.47 ^d	1							
9. ANA at T3	0.50 ^d	0.28 ^d	0.44 ^d	0.36 ^d	0.65 ^d	0.33 ^d	0.54 ^d	0.41 ^d	1						
10. EAS at T3	0.37 ^d	0.32 ^d	0.23 ^d	0.29 ^d	0.46 ^d	0.44 ^d	0.37 ^d	0.29 ^d	0.71 ^d	1					
11. AS at T3	0.44 ^d	0.23 ^d	0.52 ^d	0.28 ^d	0.51 ^d	0.30 ^d	0.66 ^d	0.34 ^d	0.79 ^d	0.64 ^d	1				
12. PTR at T3	0.34 ^d	0.23 ^d	0.29 ^d	0.48 ^d	0.41 ^d	0.27 ^d	0.36 ^d	0.55 ^d	0.67 ^d	0.58 ^d	0.60 ^d	1			
13. IGD at T1	0.44 ^d	0.32 ^d	0.30 ^d	0.31 ^d	0.24 ^d	0.29 ^d	0.28 ^d	0.33 ^d	0.24 ^d	0.25 ^d	0.19 ^d	0.21 ^d	1		
14. IGD at T2	0.34 ^d	0.24 ^d	0.30 ^d	0.28 ^d	0.39 ^d	0.31 ^d	0.31 ^d	0.23 ^d	0.31 ^d	0.28 ^d	0.25 ^d	0.24 ^d	0.53 ^d	1	
15. IGD at T3	0.31 ^d	0.25 ^d	0.29 ^d	0.21 ^d	0.36 ^d	0.31 ^d	0.30 ^d	0.24 ^d	0.40 ^d	0.36 ^d	0.33 ^d	0.30 ^d	0.48 ^d	0.55 ^d	1

^aANA, EAS, AS, and PTR are 4 dimensions of school refusal.

^bANA: avoidance of negative affectivity.

^cEAS: escape from aversive social-evaluative situations.

^d $p < .001$.

^eAS: attention-seeking.

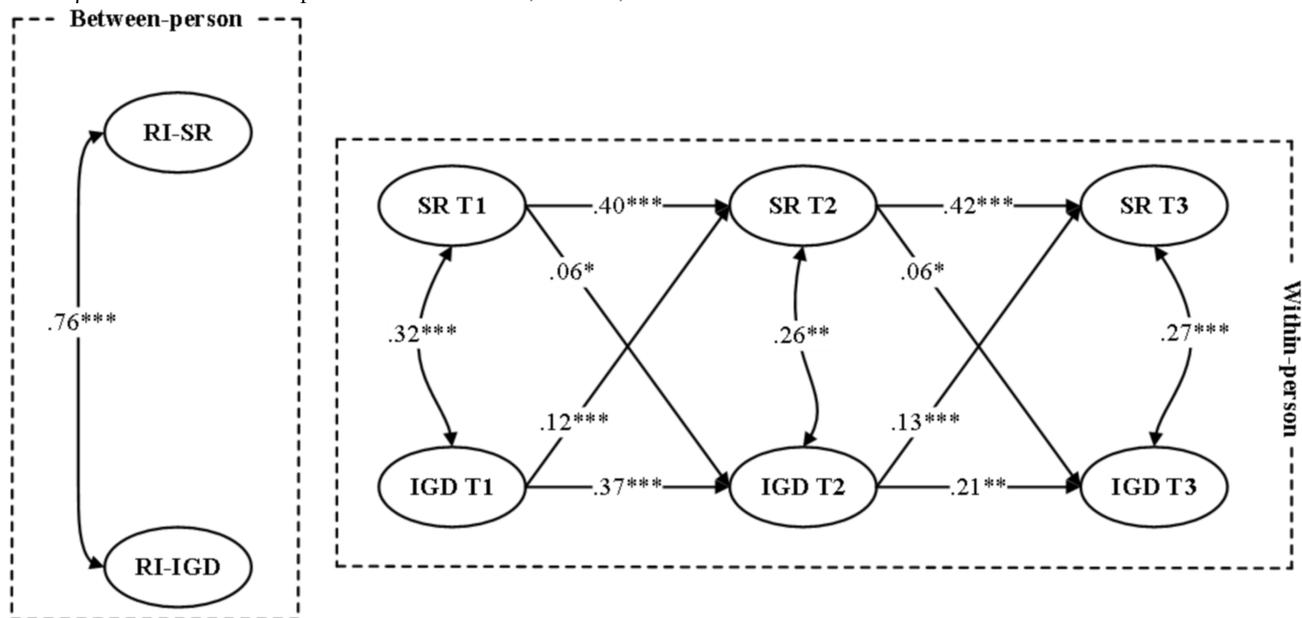
^fPTR: pursuit of tangible rewards.

RI-CLPM

The fit indices of the final selected model are presented in Table S2 in [Multimedia Appendix 1](#), demonstrating a good model fit

to the data ($\chi^2_5=16.02$; CFI=0.970; TLI=0.953; RMSEA=0.035; SRMR=0.029). The detailed results of model estimation are shown in [Figure 1](#).

Figure 1. Random intercept (RI) cross-lagged panel model illustrating the longitudinal reciprocal associations between school refusal (SR) and internet gaming disorder (IGD) across 3 waves in a sample of Chinese adolescent gamers. School refusal was modeled as a latent construct indicated by 4 functional dimensions. The model separates stable between-person differences from within-person fluctuations over time. Standardized path coefficients are shown. β indicates standardized path coefficients. * $P<.05$, ** $P<.01$, *** $P<.001$.



Within-person autoregressive paths were significant for both school refusal (T1 \rightarrow T2: $\beta=.40$, 95% CI .29-.51; $P<.001$ and T2 \rightarrow T3: $\beta=.42$, 95% CI .30-.54; $P<.001$) and IGD (T1 \rightarrow T2: $\beta=.37$, 95% CI .16-.58; $P<.001$ and T2 \rightarrow T3: $\beta=.26$, 95% CI .14-.38; $P=.003$), suggesting temporal stability.

Regarding cross-lagged effects, school refusal significantly predicted later IGD (T1 \rightarrow T2: $\beta=.06$, 95% CI .02-.08; $P=.02$ and T2 \rightarrow T3: $\beta=.06$, 95% CI .02-.08; $P=.02$). A reciprocal relationship was also observed, as IGD significantly and positively predicted school refusal (T1 \rightarrow T2: $\beta=.12$, 95% CI .09-.15; $P<.001$ and T2 \rightarrow T3: $\beta=.13$, 95% CI .08-.16; $P<.001$).

Within-person residual correlations among school refusal and IGD were positive and significant across all time points, with r ranging from .26 to .32. At the between-person level, the latent trait factors of school refusal and IGD were positively correlated ($r=.76$), indicating that adolescents who reported higher average

levels of school refusal across time also tended to report more severe IGD.

Multigroup RI-CLPM

In the first step, an unconstrained model was estimated in which all autoregressive and cross-lagged paths were freely estimated across males and females. The model demonstrated good fit to the data ($\chi^2_{10}=17.86$, CFI=0.994, TLI=0.982, RMSEA=0.045, SRMR=0.034). In the second step, a constrained model was tested in which all cross-lagged paths were constrained to be equal across groups. This model showed a less optimal fit ($\chi^2_{14}=48.84$, CFI=0.973, TLI=0.943, RMSEA=0.056, SRMR=0.061). A chi-square difference test indicated that the constrained model fit significantly worse than the unconstrained model ($\Delta\chi^2_4=30.98$, $P<.001$), suggesting that the within-person association between school refusal and IGD differed significantly by sex. Figures 2 and 3 present the standardized results.

Figure 2. Random intercept (RI) cross-lagged panel model depicting longitudinal associations between school refusal (SR) and internet gaming disorder (IGD) among male adolescent gamers across 3 measurement waves. The model estimates within-person cross-lagged effects while accounting for stable between-person differences. Standardized path coefficients are presented. ** $P < .01$, *** $P < .001$.

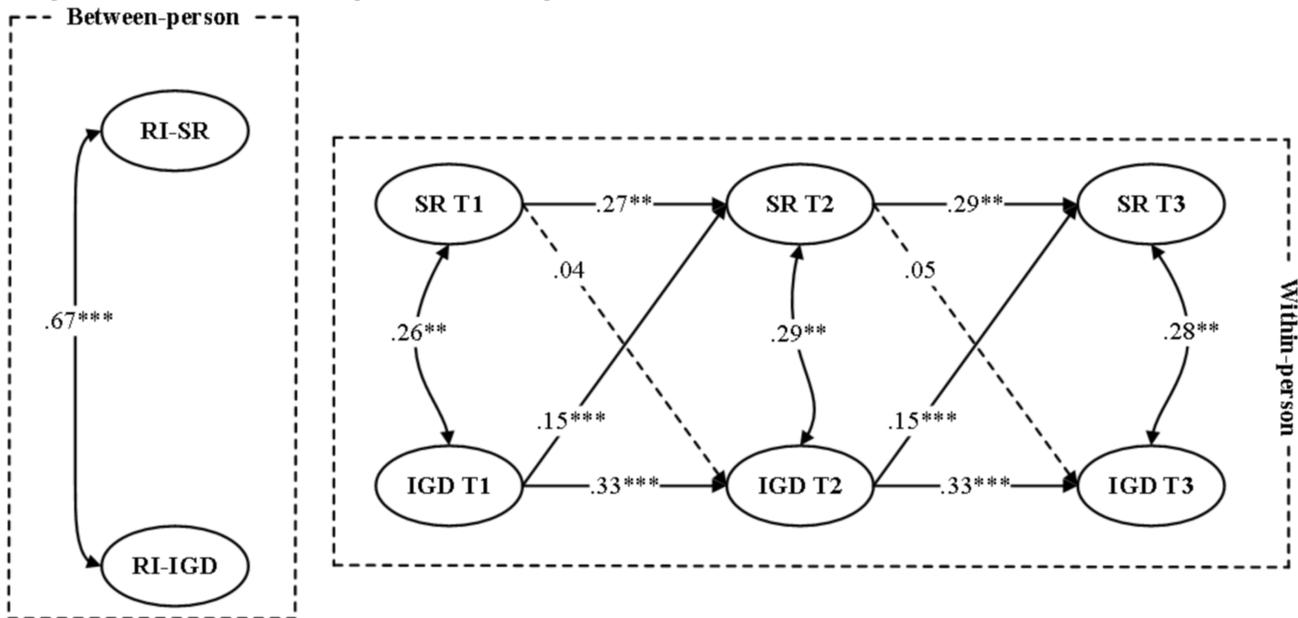
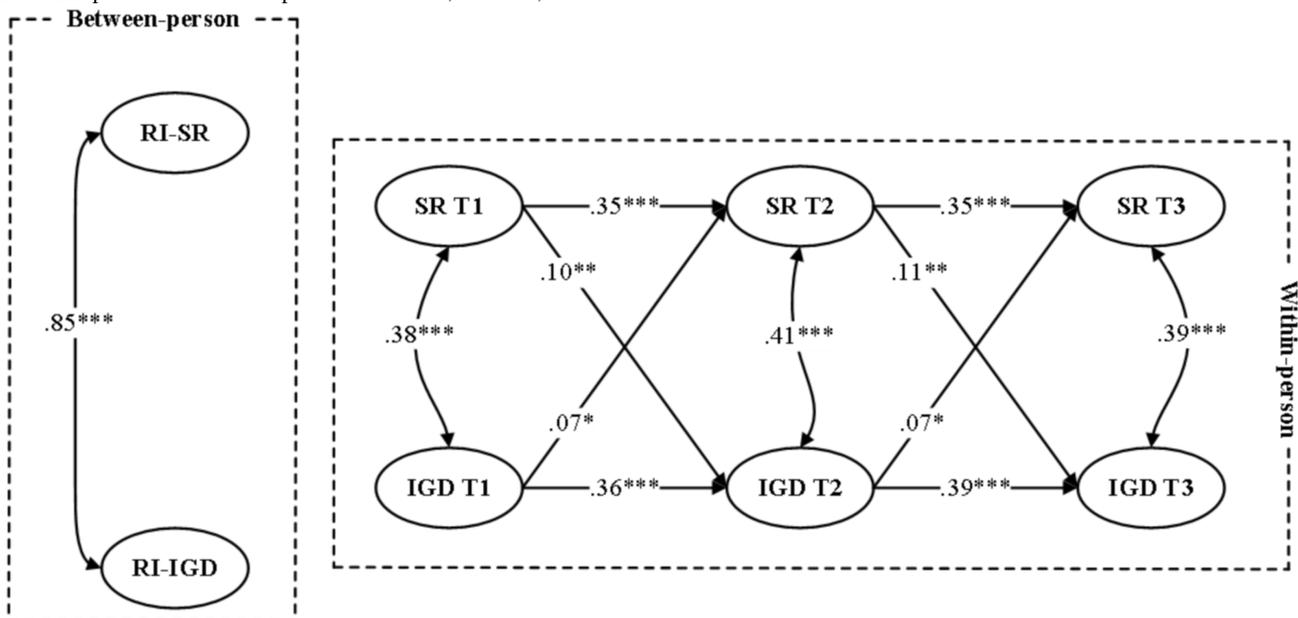


Figure 3. Random intercept (RI) cross-lagged panel model depicting longitudinal associations between school refusal (SR) and internet gaming disorder (IGD) among female adolescent gamers across 3 measurement waves. The model separates within-person temporal effects from between-person stability. Standardized path coefficients are presented. * $P < .05$, ** $P < .01$, *** $P < .001$.



Specifically, sex differences were observed in the pathway from school refusal predicting subsequent IGD. A Wald test of parameter constraints indicated a significant difference between males and females ($Wald \chi^2(1) = 10.74, P = .001$). Among females, school refusal significantly predicted later IGD at both time intervals ($T1 \rightarrow T2: \beta = .10, 95\% CI .05-.15; P = .006$ and $T2 \rightarrow T3: \beta = .11, 95\% CI .04-.17; P = .004$). In contrast, for males, the corresponding cross-lagged paths were not statistically significant. Sex differences were also found in the reverse pathway from IGD predicting school refusal, with a Wald test indicating a significant difference across sexes ($Wald$ test estimate = 8.96, $P = .003$). For males, the paths from IGD to later school refusal were significant ($T1 \rightarrow T2: \beta = .15, 95\% CI$

$.12-.18; P < .001$ and $T2 \rightarrow T3: \beta = .15, 95\% CI .12-.18; P < .001$). Among females, IGD significantly predicted school refusal ($T1 \rightarrow T2: \beta = .07, 95\% CI .03-.10; P = .02$ and $T2 \rightarrow T3: \beta = .07, 95\% CI .03-.10; P = .02$), but the effects were weaker.

Discussion

Principal Findings

The present study used RI-CLPM to investigate the longitudinal association between school refusal and IGD among Chinese adolescent gamers. By isolating within-person fluctuations from stable between-person differences, this study offers the novel longitudinal evidence that these 2 behaviors co-develop over

time in this population. Across 3 waves, school refusal was consistently associated with IGD at the subsequent wave, while IGD was likewise associated with subsequent increases in school refusal. Although the cross-lagged effects were modest in magnitude, their consistency across waves suggests a reliable within-person association. Multigroup RI-CLPM analyses indicated that females showed a significant path from school refusal to later IGD, whereas this path was not significant for males. In the opposite direction, IGD predicted later school refusal for both sexes, but the effects were stronger for males than for females. These findings suggest that school refusal and IGD are associated over time, with the direction and strength of these associations differing by sex.

Using RI-CLPM, the present study revealed significant within-person associations between school refusal and IGD across 3 waves. By separating stable between-person differences, the model captures how fluctuations relative to one's typical level are associated across time. The results showed that when adolescents reported higher-than-usual levels of school refusal, they also tended to report higher-than-usual levels of IGD at the subsequent wave. Likewise, when adolescents experienced temporary increases in IGD, they tended to show higher levels of school refusal at the next measurement point. Recent methodological work by Orth et al [49] suggests that cross-lagged effects of approximately 0.03, 0.07, and 0.12 can be interpreted as small, medium, and large, respectively, in longitudinal panel models. Within this framework, the association from school refusal to later IGD ($\beta=.06$) represents a small-to-medium effect, whereas the association from IGD to later school refusal ($\beta .12 - .13$) reaches the medium-to-large range. Considering that RI-CLPM typically yields conservative estimates, these coefficients suggest that the associations are meaningful at the within-person level rather than trivial.

The within-person association from school refusal to later IGD is consistent with the logic of the escape-avoidance framework and supports H1. Among adolescent gamers, school refusal often reflects attempts to disengage from distressing or overwhelming school-related contexts, such as academic pressure, teacher criticism, peer rejection, or bullying [50]. Avoiding school temporarily removes adolescents from environments that evoke anxiety, shame, or fear of failure and simultaneously creates large blocks of unstructured time with minimal external demands [13]. Online gaming becomes an appealing outlet in such circumstances because it offers immediate feedback, clear goals, rapid progress, and a sense of efficacy that contrasts with the uncertainty and evaluative nature of school settings [51]. When adolescents experience a within-person surge in school refusal, the psychological discomfort underlying avoidance, combined with increased discretionary time, may heighten the likelihood of turning to online games as a form of emotional relief and distraction [52]. Although the effect size for this pathway is small to medium, its repeated appearance across waves suggests a stable association. The finding emerging under the conservative RI-CLPM model further indicates that increases in school-related avoidance tend to be followed by increases in problematic gaming.

The reverse association, that is, from IGD to later school refusal, was stronger in magnitude and aligns with a self-determination perspective on adolescent gaming motivations, supporting H2. Online games often provide highly rewarding compensatory experiences, giving adolescents feelings of competence, relatedness, and autonomy that may be less accessible in the school environment [53]. As IGD intensifies, adolescents may increasingly structure their daily routines, emotional needs, and social relationships around virtual spaces [54]. Activities such as coordinating with guilds or teams, maintaining rankings, or progressing within a game's achievement system may become central to their identity and self-worth [55]. Under such circumstances, attending school can feel not merely unappealing but actively disruptive. School attendance interrupts gaming activities, reduces contact with gaming peers, and may expose adolescents to criticism about their gaming habits from adults [56]. Consequently, when adolescents exhibit within-person increases in IGD symptoms, school avoidance may function as a strategy to protect the psychological rewards and social belonging embedded in the gaming environment. The medium-to-large effect size suggests that increases in IGD are subsequently followed by higher school refusal.

These reciprocal associations suggest that increases in school refusal and IGD tend to occur together, creating a pattern that may become more stable over time if not addressed. Although each wave-to-wave association is modest in size, the fact that these associations appear consistently across multiple waves suggests that their effects may accumulate. During adolescence, even relatively small but persistent shifts in everyday behavior can, over the course of a school year, influence students' engagement with school, relationships with peers, and general well-being [57]. Rather than viewing school refusal and IGD as separate concerns, the present findings indicate that they often co-occur and may reflect a broader process through which adolescents manage stress and seek more controllable or rewarding environments. This close connection between the 2 behaviors highlights the value of approaches that consider both behaviors together rather than treating them as unrelated issues.

The *t* test results showed a familiar pattern often reported in previous studies. Across the 3 waves, males had lower mean scores on most dimensions of school refusal (except EAS at T1), whereas they consistently scored higher than females on IGD [34,37]. These mean-level differences suggest that females may be more likely to respond to school-related stress with emotional withdrawal, while males tend to spend more time in gaming. However, mean differences alone cannot tell us whether the 2 behaviors are associated in the same way for males and females over time. For this reason, a multigroup RI-CLPM was performed to examine whether the longitudinal association differed by sex.

The multigroup RI-CLPM results revealed clear sex-specific patterns in the longitudinal associations between school refusal and IGD, as confirmed by Wald tests comparing constrained and unconstrained models. Overall, the findings suggest a more reciprocal pattern among females, whereas for males, the dominant direction runs from IGD to subsequent school refusal, supporting H3. For females, school refusal significantly predicted later IGD at both intervals, whereas this longitudinal

path was not significant for males. This indicates that school-related avoidance may play a more central role in the development of problematic gaming among females. In line with the escape-avoidance perspective, females may be more inclined to use gaming as a way of managing school-related distress, such as anxiety about performance, worry about teacher evaluation, or tension in peer relationships [58]. When school experiences become emotionally overwhelming, online games may offer a relatively safe and controllable context, making this coping function more common for females [59]. Importantly, this pattern remained after accounting for stable differences between individuals. This suggests that for females, short-term increases in school-related avoidance are followed by increases in IGD rather than merely reflecting general vulnerability.

In contrast, short-term increases in school refusal among males were not associated with subsequent within-person changes in IGD, despite males reporting higher levels of IGD on average. This pattern suggests that for males, school refusal and gaming are not associated in the same “avoid school then turn to games” sequence that appears for females. Instead, the reverse direction was more evident: for males, IGD significantly predicted later school refusal at both intervals, whereas for females, the same paths were significant but weaker. This pattern implies that among males, periods of higher IGD are more strongly tied to later problems with school participation. One possible interpretation is that males may be more likely to organize their daily routines around gaming. They may stay up late to play, give priority to in-game commitments such as matches or team activities, and gradually reduce the time and energy available for schoolwork and attendance [60]. Over time, these shifts in time use and focus may make school feel less relevant or more burdensome, increasing the likelihood of school refusal when gaming becomes difficult to balance with academic demands [61]. For females, IGD was also associated with later school refusal, but the weaker effect suggests that problematic gaming is less central as a driver of school avoidance than it is for males.

The findings have several implications. First, the reciprocal within-person association between school refusal and IGD

suggests that these behaviors should not be treated as independent problems. When adolescents show shifts in either school-related avoidance or gaming involvement, these changes may influence the other domain, which indicates that interventions may be more effective when targeting both areas simultaneously. Second, the sex-specific results further highlight the need for more individualized approaches. For females, providing support to help them manage academic stress, improve classroom experiences, and feel more connected at school may reduce the likelihood of turning to gaming as a coping strategy [62]. For males, paying attention to patterns such as sleep disruption, late-night play, and reduced time for schoolwork may help prevent later school disengagement [63]. More broadly, the study shows that RI-CLPM can help identify how short-term changes in adolescents’ school and gaming experiences relate to each other over time.

Conclusion

The present study used RI-CLPM to examine how school refusal and IGD develop over time among Chinese adolescent gamers. By separating within-person changes from stable between-person differences, the findings provide evidence of a reinforcing within-person association. The multigroup RI-CLPM further showed that this bidirectional pattern was evident among females, who showed significant paths in both directions. In contrast, males showed a different pattern: IGD predicted later school refusal, but school refusal did not predict later IGD. Given the observational design and the modest effect sizes, these findings should be interpreted cautiously and as indicative of developmental tendencies rather than deterministic processes. These findings provide a clear longitudinal pattern of mutual reinforcement and highlight the importance of considering school disengagement and problematic gaming together rather than as separate issues. The results also show that these processes develop differently for males and females. This suggests important implications for tailored prevention and intervention efforts, as sex-specific differences should be taken into account.

Acknowledgments

We thank the participants who participated in the study.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

PZ: writing-original draft, visualization, methodology, investigation, formal analysis, data curation, conceptualization.

ZT: writing-original draft, visualization, methodology, investigation, data curation.

LF: writing-original draft, visualization, methodology, investigation, formal analysis, data curation, conceptualization.

TY: methodology, formal analysis, data curation.

LZ: writing-review & editing, supervision.

YW: writing-original draft, visualization, methodology, investigation.

YL: writing-original draft, visualization, methodology, investigation, formal analysis, data curation, conceptualization.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Tables displaying measurement invariance across sex and fit statistics and model comparisons for random intercept cross-lagged panel models (RI-CLPMs) of school refusal and internet gaming disorder (IGD).

[[DOCX File, 16 KB - games_v14i1e89619_app1.docx](#)]

References

1. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Association; 2013. [doi: [10.1176/appi.books.9780890425596](https://doi.org/10.1176/appi.books.9780890425596)]
2. Luo T, Wei D, Guo J, et al. Diagnostic contribution of the DSM-5 criteria for internet gaming disorder. *Front Psychiatry* 2021;12:777397. [doi: [10.3389/fpsy.2021.777397](https://doi.org/10.3389/fpsy.2021.777397)] [Medline: [35069285](https://pubmed.ncbi.nlm.nih.gov/35069285/)]
3. Gao YX, Wang JY, Dong GH. The prevalence and possible risk factors of internet gaming disorder among adolescents and young adults: systematic reviews and meta-analyses. *J Psychiatr Res* 2022 Oct;154:35-43. [doi: [10.1016/j.jpsychires.2022.06.049](https://doi.org/10.1016/j.jpsychires.2022.06.049)] [Medline: [35926424](https://pubmed.ncbi.nlm.nih.gov/35926424/)]
4. Zhou R, Morita N, Ogai Y, et al. Meta-analysis of internet gaming disorder prevalence: assessing the impacts of DSM-5 and ICD-11 diagnostic criteria. *Int J Environ Res Public Health* 2024 May 29;21(6):700. [doi: [10.3390/ijerph21060700](https://doi.org/10.3390/ijerph21060700)] [Medline: [38928946](https://pubmed.ncbi.nlm.nih.gov/38928946/)]
5. Fam JY. Prevalence of internet gaming disorder in adolescents: a meta-analysis across three decades. *Scand J Psychol* 2018 Oct;59(5):524-531. [doi: [10.1111/sjop.12459](https://doi.org/10.1111/sjop.12459)] [Medline: [30004118](https://pubmed.ncbi.nlm.nih.gov/30004118/)]
6. Gou S, Zhang W, Tang Y, Zhang J, He Q. Prevalence of internet gaming disorder among Chinese adolescents: a systematic review and meta-analysis. *Asian J Psychiatr* 2024 Dec;102:104257. [doi: [10.1016/j.ajp.2024.104257](https://doi.org/10.1016/j.ajp.2024.104257)] [Medline: [39366164](https://pubmed.ncbi.nlm.nih.gov/39366164/)]
7. Ghali S, Afifi S, Suryadevara V, et al. A systematic review of the association of internet gaming disorder and excessive social media use with psychiatric comorbidities in children and adolescents: is it a curse or a blessing? *Cureus* 2023 Aug;15(8):e43835. [doi: [10.7759/cureus.43835](https://doi.org/10.7759/cureus.43835)] [Medline: [37736454](https://pubmed.ncbi.nlm.nih.gov/37736454/)]
8. Lee SJ, Jeong EJ, Choi JI, Park MS. Social intelligence and pathological gaming: a longitudinal study of the associations among negative emotions, social intelligence, aggression, and pathological gaming in adolescents. *Front Psychiatry* 2024;15:1353969. [doi: [10.3389/fpsy.2024.1353969](https://doi.org/10.3389/fpsy.2024.1353969)] [Medline: [38903650](https://pubmed.ncbi.nlm.nih.gov/38903650/)]
9. Li S, Wu Z, Zhang Y, Xu M, Wang X, Ma X. Internet gaming disorder and aggression: a meta-analysis of teenagers and young adults. *Front Public Health* 2023;11:1111889. [doi: [10.3389/fpubh.2023.1111889](https://doi.org/10.3389/fpubh.2023.1111889)] [Medline: [37089492](https://pubmed.ncbi.nlm.nih.gov/37089492/)]
10. She R, Zhang Y, Yang X. Parental factors associated with internet gaming disorder among first-year high school students: longitudinal study. *JMIR Serious Games* 2022 Nov 8;10(4):e33806. [doi: [10.2196/33806](https://doi.org/10.2196/33806)] [Medline: [36346660](https://pubmed.ncbi.nlm.nih.gov/36346660/)]
11. Kearney CA. Identifying the function of school refusal behavior: a revision of the school refusal assessment scale. *J Psychopathol Behav Assess* 2002 Dec;24(4):235-245. [doi: [10.1023/A:1020774932043](https://doi.org/10.1023/A:1020774932043)]
12. Kearney CA. Confirmatory factor analysis of the school refusal assessment scale-revised: child and parent versions. *J Psychopathol Behav Assess* 2006 Sep;28(3):139-144. [doi: [10.1007/s10862-005-9005-6](https://doi.org/10.1007/s10862-005-9005-6)]
13. Di Vincenzo C, Pontillo M, Bellantoni D, et al. School refusal behavior in children and adolescents: a five-year narrative review of clinical significance and psychopathological profiles. *Ital J Pediatr* 2024 May 30;50(1):107. [doi: [10.1186/s13052-024-01667-0](https://doi.org/10.1186/s13052-024-01667-0)] [Medline: [38816858](https://pubmed.ncbi.nlm.nih.gov/38816858/)]
14. Kearney CA. *School Refusal Behavior in Youth: A Functional Approach to Assessment and Treatment*: American Psychological Association; 2001. [doi: [10.1037/10426-000](https://doi.org/10.1037/10426-000)]
15. Kearney CA, Albano AM. *When Children Refuse School: A Cognitive-Behavioral Therapy Approach, Parent Workbook*, 2nd edition: Oxford University Press; 2007. [doi: [10.1093/med:psych/9780195308297.001.0001](https://doi.org/10.1093/med:psych/9780195308297.001.0001)]
16. Xu X, Lu H, Du M, et al. A potential vicious cycle between school refusal and depression among Chinese adolescents: a cross-lagged panel model analysis. *Int J Ment Health Promot* 2025;27(10):1423-1437. [doi: [10.32604/ijmh.2025.068840](https://doi.org/10.32604/ijmh.2025.068840)]
17. González C, Martín M, Vicent M, Sanmartín R. School refusal behavior and aggression in Spanish adolescents. *Front Psychol* 2021;12:669438. [doi: [10.3389/fpsyg.2021.669438](https://doi.org/10.3389/fpsyg.2021.669438)] [Medline: [33995227](https://pubmed.ncbi.nlm.nih.gov/33995227/)]
18. González C, Díaz-Herrero Á, Vicent M, Sanmartín R, del Pilar Aparicio-Flores M, García-Fernández JM. Typologies of Spanish youth with school refusal behavior and their relationship with aggression. *J Aggress Maltreat Trauma* 2023 Dec 2;32(12):1779-1797. [doi: [10.1080/10926771.2023.2234840](https://doi.org/10.1080/10926771.2023.2234840)]
19. Dinsmoor JA. Escape, avoidance, punishment: where do we stand? *J Exp Anal Behav* 1977 Jul;28(1):83-95. [doi: [10.1901/jeab.1977.28-83](https://doi.org/10.1901/jeab.1977.28-83)] [Medline: [16812016](https://pubmed.ncbi.nlm.nih.gov/16812016/)]
20. Pérez - Marco M, González C, Fuster - Rico A, Vicent M, Fernández - Sogorb A, Aparicio - Flores MDP. Adolescent school refusal and academic self - efficacy: a latent profile and regression analysis. *Psychol Sch* 2024 Oct;61(10):4041-4056. [doi: [10.1002/pits.23263](https://doi.org/10.1002/pits.23263)]

21. Koşan Y, Ulaş S. Mediating effect of self-regulation on the relationship between digital game addiction and school refusal in adolescents. *J STEM Teach Inst* 2023 Oct 12;3(2):75-86 [[FREE Full text](#)]
22. Deci EL, Ryan RM. Self-determination theory: a macrotheory of human motivation, development, and health. *Can Psychol* 2008 Aug;49(3):182-185. [doi: [10.1037/a0012801](#)]
23. Luján-Barrera A, Cervera-Ortiz L, Chóliz M. In-game need satisfaction, frustration, and gaming addiction patterns across subgroups of adolescents through structural equation modeling: cross-sectional and instrument validation study of the youth gaming experience scales. *JMIR Serious Games* 2025 Nov 12;13(1):e63612. [doi: [10.2196/63612](#)] [Medline: [41223412](#)]
24. Ryan RM, Rigby CS, Przybylski A. The motivational pull of video games: a self-determination theory approach. *Motiv Emot* 2006 Dec 12;30(4):344-360. [doi: [10.1007/s11031-006-9051-8](#)]
25. Tiraboschi GA, Garon-Carrier G, Madigan S, Smith J, Surprenant R, Fitzpatrick C. Longitudinal associations between gaming and academic motivation during middle childhood. *Psychol Med* 2025 Aug 25;55:e235. [doi: [10.1017/S0033291725101153](#)] [Medline: [40851392](#)]
26. Kosola S, Kullberg M, Melander K, Engblom J, Ranta K, Alanko K. Associations of excessive internet use, sleep duration and physical activity with school absences: a cross-sectional, population-based study of adolescents in years 8 and 9. *Arch Dis Child* 2024 Jun 19;109(7):570-575. [doi: [10.1136/archdischild-2023-326331](#)] [Medline: [38627028](#)]
27. Ren Y, Jia S, Zheng Q, Liang S, He L. Latent subtypes of school refusal behavior and their social-ecological correlates among Chinese adolescents. *WSR* 2025 Oct 31;1(3):170-179. [doi: [10.54844/wsr.2025.1072](#)]
28. Fujita J, Aoyama K, Saigusa Y, et al. Problematic internet use and daily difficulties among adolescents with school refusal behaviors: an observational cross-sectional analytical study. *Medicine (Abingdon)* ;101(7):e28916. [doi: [10.1097/MD.00000000000028916](#)] [Medline: [35363214](#)]
29. Yu Y, Yen SHY, Wang DB, et al. School climate and school identification as determinants of internet gaming disorder among Chinese adolescent internet gamers: cross-sectional mediation study. *JMIR Serious Games* 2024 Nov 12;12:e50418. [doi: [10.2196/50418](#)] [Medline: [39622695](#)]
30. Amez S, Vujić S, De Marez L, Baert S. Smartphone use and academic performance: first evidence from longitudinal data. *New Media Soc* 2023 Mar;25(3):584-608. [doi: [10.1177/14614448211012374](#)]
31. Masood A, Luqman A, Feng Y, Ali A. Adverse consequences of excessive social networking site use on academic performance: explaining underlying mechanism from stress perspective. *Comput Human Behav* 2020 Dec;113:106476. [doi: [10.1016/j.chb.2020.106476](#)]
32. Chen J. The effect of social network sites usage on absenteeism and labor outcomes: longitudinal evidence. *Int J Manpow* 2023 Nov 9;44(6):1168-1195. [doi: [10.1108/IJM-06-2021-0338](#)]
33. Veerman CACD, van Vuuren CL, Verhoeff AP. Hangen gamen en sociale-mediagebruik samen met schoolverzuim? *Tijdschr Gezondheidswet* 2017 Jul;95(4):169-179. [doi: [10.1007/s12508-017-0049-2](#)]
34. Deng X, Hu YB, Liu CY, et al. Psychological distress and aggression among adolescents with internet gaming disorder symptoms. *Psychiatry Res* 2024 Jan;331:115624. [doi: [10.1016/j.psychres.2023.115624](#)] [Medline: [38039647](#)]
35. Macur M, Pontes HM. Internet gaming disorder in adolescence: investigating profiles and associated risk factors. *BMC Public Health* 2021 Aug 12;21(1):1547. [doi: [10.1186/s12889-021-11394-4](#)] [Medline: [34384394](#)]
36. Yu Y, Mo PKH, Zhang J, Li J, Lau JTF. Why is internet gaming disorder more prevalent among Chinese male than female adolescents? The role of cognitive mediators. *Addict Behav* 2021 Jan;112:106637. [doi: [10.1016/j.addbeh.2020.106637](#)] [Medline: [32919322](#)]
37. Havik T, Bru E, Ertesvåg SK. Assessing reasons for school non-attendance. *Scand J Educ Res* 2015 May 4;59(3):316-336. [doi: [10.1080/00313831.2014.904424](#)]
38. González C, Díaz-Herrero Á, Sanmartín R, Vicent M, Fernández-Sogorb A, García-Fernández JM. Testing the functional profiles of school refusal behavior and clarifying their relationship with school anxiety. *Front Public Health* 2020;8:598915. [doi: [10.3389/fpubh.2020.598915](#)] [Medline: [33344402](#)]
39. Dong GH, Potenza MN. Considering gender differences in the study and treatment of internet gaming disorder. *J Psychiatr Res* 2022 Sep;153:25-29. [doi: [10.1016/j.jpsychires.2022.06.057](#)] [Medline: [35793576](#)]
40. Zhang J, Hu Y, Wang Z, Wang M, Dong GH. Males are more sensitive to reward and less sensitive to loss than females among people with internet gaming disorder: fMRI evidence from a card-guessing task. *BMC Psychiatry* 2020 Jul 7;20(1):357. [doi: [10.1186/s12888-020-02771-1](#)] [Medline: [32635911](#)]
41. Yang X, Jiang X, Wu AMS, et al. Validation of the internet gaming disorder symptoms checklist based on the Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders in Chinese adolescents. *Child Psychiatry Hum Dev* 2023 Feb;54(1):26-33. [doi: [10.1007/s10578-021-01213-7](#)] [Medline: [34338962](#)]
42. Fu J. The reliability and validity of Chinese version of the School Refusal Assessment Scale-Revised [Master's thesis in Chinese]. : China Medical University; 2013 URL: <https://d.wanfangdata.com.cn/thesis/Y2297551> [accessed 2026-02-28]
43. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equ Modeling* 2002 Apr;9(2):233-255. [doi: [10.1207/S15328007SEM0902_5](#)]
44. Dang L, Zhai J, Kam CCS, Su X, Wu AMS. Gaming for self-searching? A three-wave longitudinal investigation on the correlations and predictions among self-concept clarity, player-avatar identification, and internet gaming disorder. *J Youth Adolescence* 2026 Jan;55(1):14-26. [doi: [10.1007/s10964-025-02258-w](#)]

45. Yu Y, Li S, Lyu X, Zhang J, Li JB, Lau JTF. A cross-lagged panel model of the bidirectional relationships between gaming-specific maladaptive cognitions and internet gaming disorder among adolescents in China. *Addict Behav* 2026 Jan;172:108518. [doi: [10.1016/j.addbeh.2025.108518](https://doi.org/10.1016/j.addbeh.2025.108518)] [Medline: [41075582](https://pubmed.ncbi.nlm.nih.gov/41075582/)]
46. Teng Z, Pontes HM, Nie Q, Xiang G, Griffiths MD, Guo C. Internet gaming disorder and psychosocial well-being: a longitudinal study of older-aged adolescents and emerging adults. *Addict Behav* 2020 Nov;110:106530. [doi: [10.1016/j.addbeh.2020.106530](https://doi.org/10.1016/j.addbeh.2020.106530)] [Medline: [32683173](https://pubmed.ncbi.nlm.nih.gov/32683173/)]
47. Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Equ Modeling* 2001 Jul 1;8(3):430-457. [doi: [10.1207/S15328007SEM0803_5](https://doi.org/10.1207/S15328007SEM0803_5)]
48. Appelbaum M, Cooper H, Kline RB, Mayo-Wilson E, Nezu AM, Rao SM. Journal article reporting standards for quantitative research in psychology: the APA publications and communications board task force report. *Am Psychol* 2018 Jan;73(1):3-25. [doi: [10.1037/amp0000191](https://doi.org/10.1037/amp0000191)] [Medline: [29345484](https://pubmed.ncbi.nlm.nih.gov/29345484/)]
49. Orth U, Meier LL, Bühler JL, et al. Effect size guidelines for cross-lagged effects. *Psychol Methods* 2024 Apr;29(2):421-433. [doi: [10.1037/met0000499](https://doi.org/10.1037/met0000499)] [Medline: [35737548](https://pubmed.ncbi.nlm.nih.gov/35737548/)]
50. Ochi M, Kawabe K, Ochi S, Miyama T, Horiuchi F, Ueno SI. School refusal and bullying in children with autism spectrum disorder. *Child Adolesc Psychiatry Ment Health* 2020;14(1):17. [doi: [10.1186/s13034-020-00325-7](https://doi.org/10.1186/s13034-020-00325-7)] [Medline: [32419839](https://pubmed.ncbi.nlm.nih.gov/32419839/)]
51. Boyle EA, Connolly TM, Hainey T, Boyle JM. Engagement in digital entertainment games: a systematic review. *Comput Human Behav* 2012 May;28(3):771-780. [doi: [10.1016/j.chb.2011.11.020](https://doi.org/10.1016/j.chb.2011.11.020)]
52. Faílde Garrido JM, Dapía Conde MD, Isorna Folgar M, Braña Rey F. Problematic use of video games in schooled adolescents: the role of passion. *Behav Sci (Basel)* 2024 Oct 24;14(11):992. [doi: [10.3390/bs14110992](https://doi.org/10.3390/bs14110992)] [Medline: [39594292](https://pubmed.ncbi.nlm.nih.gov/39594292/)]
53. Filippello P, Buzzai C, Costa S, Sorrenti L. School refusal and absenteeism: perception of teacher behaviors, psychological basic needs, and academic achievement. *Front Psychol* 2019;10(1471):1471. [doi: [10.3389/fpsyg.2019.01471](https://doi.org/10.3389/fpsyg.2019.01471)] [Medline: [31316431](https://pubmed.ncbi.nlm.nih.gov/31316431/)]
54. Wang X, Liu Y, Chu HKC, Wong SYS, Yang X. Relationships of internet gaming engagement, history, and maladaptive cognitions and adolescent internet gaming disorder: a cross-sectional study. *PLoS ONE* 2023 Sep 8;18(9):e0290955. [doi: [10.1371/journal.pone.0290955](https://doi.org/10.1371/journal.pone.0290955)]
55. Brailovskaia J, Bierhoff HW. Cross-cultural narcissism on Facebook: relationship between self-presentation, social interaction and the open and covert narcissism on a social networking site in Germany and Russia. *Comput Human Behav* 2016 Feb;55:251-257. [doi: [10.1016/j.chb.2015.09.018](https://doi.org/10.1016/j.chb.2015.09.018)]
56. Mahmud S, Jobayer MAA, Salma N, Mahmud A, Tamanna T. Online gaming and its effect on academic performance of Bangladeshi university students: a cross-sectional study. *Health Sci Rep* 2023 Dec;6(12):e1774. [doi: [10.1002/hsr2.1774](https://doi.org/10.1002/hsr2.1774)] [Medline: [38107152](https://pubmed.ncbi.nlm.nih.gov/38107152/)]
57. Salmela - Aro K, Tang X, Symonds J, Upadyaya K. Student engagement in adolescence: a scoping review of longitudinal studies 2010–2020. *J Res Adolesc* 2021 Jun;31(2):256-272. [doi: [10.1111/jora.12619](https://doi.org/10.1111/jora.12619)]
58. Kiszka P, Strojny A, Strojny P. High immersion/escapism motivation makes gaming disorder risk less dependent of playtime among highly engaged male gamers. *Front Psychiatry* 2024;15:1443091. [doi: [10.3389/fpsyg.2024.1443091](https://doi.org/10.3389/fpsyg.2024.1443091)] [Medline: [39386891](https://pubmed.ncbi.nlm.nih.gov/39386891/)]
59. Li L, Niu Z, Griffiths MD, Mei S. Relationship between gaming disorder, self-compensation motivation, game flow, time spent gaming, and fear of missing out among a sample of Chinese university students: a network analysis. *Front Psychiatry* 2021;12(5):761519. [doi: [10.3389/fpsyg.2021.761519](https://doi.org/10.3389/fpsyg.2021.761519)] [Medline: [34790137](https://pubmed.ncbi.nlm.nih.gov/34790137/)]
60. Ward MR. Cutting class to play video games. *Inf Econ Policy* 2018 Mar;42:11-19. [doi: [10.1016/j.infoecopol.2017.10.001](https://doi.org/10.1016/j.infoecopol.2017.10.001)]
61. Hamre R, Smith ORF, Samdal O, Haug E. Gaming behaviors and the association with sleep duration, social jetlag, and difficulties falling asleep among Norwegian adolescents. *Int J Environ Res Public Health* 2022 Feb 4;19(3):1765. [doi: [10.3390/ijerph19031765](https://doi.org/10.3390/ijerph19031765)] [Medline: [35162788](https://pubmed.ncbi.nlm.nih.gov/35162788/)]
62. Cheng Y, Jiang S, Chen J. Academic expectation stress and online gaming disorder among Chinese adolescents: the mediating role of psychological distress and the moderating role of stress mindset. *Child Youth Serv Rev* 2024 Mar;158:107492. [doi: [10.1016/j.childyouth.2024.107492](https://doi.org/10.1016/j.childyouth.2024.107492)]
63. Lériida-Ayala V, Aguilar-Parra JM, Collado-Soler R, Alférez-Pastor M, Fernández-Campoy JM, Luque-de la Rosa A. Internet and video games: causes of behavioral disorders in children and teenagers. *Children (Basel)* 2022 Dec 31;10(1):86. [doi: [10.3390/children10010086](https://doi.org/10.3390/children10010086)] [Medline: [36670637](https://pubmed.ncbi.nlm.nih.gov/36670637/)]

Abbreviations

ANA: avoidance of negative affectivity

AS: attention-seeking

CFI: comparative fit index

DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*

EAS: escape from aversive situations

IGD: internet gaming disorder

PTR: pursuit of tangible rewards

RI-CLPM: random intercept cross-lagged panel model
RMSEA: root mean square error of approximation
SRAS-R: School Refusal Assessment Scale-Revised
SRMR: standardized root mean square residual
TLI: Tucker-Lewis Index

Edited by S Brini; submitted 15.Dec.2025; peer-reviewed by C Hu, J Sun; accepted 07.Jan.2026; published 04.Mar.2026.

Please cite as:

Zheng P, Tao Z, Fang L, Yao T, Zhou L, Wang Y, Lin Y

Longitudinal Association Between Internet Gaming Disorder and School Refusal Among Adolescents Using a Random Intercept Cross-Lagged Panel Model: Three-Wave Prospective Cohort Study

JMIR Serious Games 2026;14:e89619

URL: <https://games.jmir.org/2026/1/e89619>

doi: [10.2196/89619](https://doi.org/10.2196/89619)

© Peng Zheng, Zi Tao, Luoxiang Fang, Tao Yao, Lingling Zhou, Yutian Wang, Yuannan Lin. Originally published in JMIR Serious Games (<https://games.jmir.org>), 4.Mar.2026. 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.

Quantifying Slowness in Parkinson Disease Using a Serious Game: Cross-Sectional Study

Luanne Cardoso Mendes¹, Prof Dr, PhD; Ariana Moura Cabral¹, BSc; Camille Marques Alves¹, PhD; Isabela Alves Marques¹, PhD; Alberto López Delis², Prof Dr, PhD; Yann Morère³, Prof Dr, PhD; Adriano de Oliveira Andrade¹, Prof Dr, PhD

¹Centre for Innovation and Technology Assessment in Health, Faculty of Electrical Engineering, Universidade Federal de Uberlândia, 2121 João Naves de Ávila Avenue, Santa Mônica, Uberlândia, Brazil

²Medical Biophysics Center, Universidad de Oriente, Santiago de Cuba, Cuba

³Laboratoire de Conception, d'Optimisation et de Modélisation des Systèmes, Université de Lorraine, Metz, France

Corresponding Author:

Adriano de Oliveira Andrade, Prof Dr, PhD

Centre for Innovation and Technology Assessment in Health, Faculty of Electrical Engineering, Universidade Federal de Uberlândia, 2121 João Naves de Ávila Avenue, Santa Mônica, Uberlândia, Brazil

Abstract

Background: Slowness in voluntary movements is a hallmark of Parkinson disease (PD); yet, objective measurement outside clinical settings is limited. Serious games represent a promising alternative to extract motor performance metrics during interactions. However, evidence on the effectiveness of these games in discriminating motor performance between individuals with and those without PD is still scarce.

Objective: This study aimed to objectively assess slowness using the RehaBEElitation serious game, based on in-game estimated measurements.

Methods: The study included 15 individuals with mild to moderate PD (Hoehn and Yahr I-III), assessed in both ON and OFF medication states, and 15 age- and gender-matched healthy controls (10 men and 5 women; mean 66.27, SD 9.13 years; range 45-82 years). All participants played each phase of the game on the easiest level. Slowness was evaluated by detecting the voluntary movement of the gyroscope signals using the singular spectrum analysis method. The response time (RT) and angular velocity (AV) of the participants while playing RehaBEElitation were estimated. Group-level comparisons were performed to investigate the presence of slowness patterns across conditions. The Kruskal-Wallis test and Wilcoxon signed rank test with Bonferroni correction were used to confirm the differences between groups, and the effect size was estimated using eta square (η^2). Spearman correlation analyses were conducted to examine associations between the RT and AV and the Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) motor items.

Results: Groups were age-homogeneous ($P > .05$). Participants with PD had significantly higher scores on MDS-UPDRS Part III in the OFF state compared to the ON state of medication (mean difference 2.86, 95% CI 2-12; $P < .05$). RT was generally shorter and AV higher in controls than in participants with PD. In the PD group, RT decreased and AV increased from OFF to ON states, reflecting an improvement in motor performance. Significant differences in RT were observed between groups in all phases of the game, with effect sizes ranging from small to moderate ($\eta^2 = 0.0239 - 0.0650$). AV differed markedly between groups in phase 4, with a large effect size ($\eta^2 = 0.404$). Correlation analyses revealed weak positive associations between RT and MDS-UPDRS items, while AV showed strong negative correlations with each motor item and the summary score for bradykinesia.

Conclusions: This study proposes a new method to assess slowness in PD by using inertial sensor data to extract objective motor measures (RT and AV) during a serious game, allowing continuous and quantitative evaluation beyond traditional clinical scales and tests. The findings demonstrate that RT and AV extracted from gameplay can detect slowness-related motor differences, supporting the RehaBEElitation serious game as an alternative and objective digital biomarker, with potential applications in both clinical and home-based monitoring of symptom progression.

(JMIR Serious Games 2026;14:e79463) doi:[10.2196/79463](https://doi.org/10.2196/79463)

KEYWORDS

Parkinson disease; bradykinesia; objective evaluation; serious game; slowness

Introduction

Serious game-based interventions have attracted considerable attention from the medical community. These technologies are defined as digital games developed with a primary purpose beyond entertainment, typically aimed at supporting education, training, or promoting behavioral change. In addition to providing rewarding experiences, these technologies can encourage interaction and enjoyable gameplay [1]. People with a variety of illnesses, including Parkinson disease (PD), can benefit from exercise training, symptom monitoring, and health improvement through serious games' engaging gaming dynamics [2]. The loss of dopaminergic neurons in the mesencephalon's substantia nigra is a hallmark of PD [3]. The lack of dopamine hinders the ability to execute and control movements, causing a number of identifiable motor symptoms, including bradykinesia, tremor, and rigidity [4].

The primary motor deficit seen in people with PD is bradykinesia, which is the incapacity to prepare, start, and carry out motions [5]. More specifically, bradykinesia can be understood as a failure of the basal ganglia to reinforce the cortical mechanisms that prepare and execute commands to move [6]. In this context, slowness in performing voluntary movements, prolonged reaction time (defined as the time between the stimulus presentation and the start of the movement), prolonged movement time (defined as the time between the start and finish of the movement), and, ultimately, prolonged response time (RT), which is the sum of reaction time and movement time, are the main characteristics of bradykinesia [7]. RT represents the capacity of an individual to identify pertinent information sources, interpret that information, and apply it to produce a suitable movement response [8]. Since PD impairs both reaction and movement time [7], RT could be an accurate indicator of slowness severity.

The most popular and well-accepted technique for the clinical evaluation of bradykinesia is the Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) [9]. The capacity of the patient to make quick, repeated, and alternating movements—like tapping fingers, opening and closing hands, and pronating and supinating the forearms—is the base for the clinical assessment of this symptom. To determine the severity of the condition, the clinician grades these movements from 0=normal to 4=severe [10]. However, this kind of evaluation is subjective and inaccurate due to a number of factors, including the evaluator's expertise level, patient compliance, and individual bias [11]. Furthermore, the clinical evaluation is unable to precisely characterize the severity of a symptom in a patient [12].

Inertial sensors can address the limitations of qualitative visual scales by providing objective motion measurements [12]. Recent studies confirm that inertial measurement units (IMUs) accurately quantify bradykinesia in PD in standardized tasks and remote monitoring. In upper limbs, studies such as Bremm et al [13] confirm that the extraction of metrics of speed, amplitude, and rhythm in tasks such as finger tapping, pronation-supination, and opening or closing the hand showed consistent correlations with clinical scales (such as the

MDS-UPDRS). In terms of mobility and overall function, Hong et al [14] used IMUs to capture walking speed, cadence, and stride length, as well as sit-to-stand performance, suggesting that wearable sensors provide a comprehensive and objective approach to assessing a wide range of motor symptoms in PD.

Recent studies in the field of continuous monitoring have highlighted the potential of wearable devices and smartphones to remotely detect patterns of slow movement and reduced range of motion throughout the day. The results demonstrated the feasibility of more accurate and personalized monitoring, reducing the need for frequent in-person visits and supporting therapeutic decisions [15].

Despite covering different tasks and contexts, and demonstrating that the use of devices with IMUs is a promising alternative that can identify and quantify slowness, no study has been conducted to evaluate this symptom with data collected from people with PD during interaction with serious games. In other words, IMU-based strategies are already valid and sensitive for assessing bradykinesia, but they have not yet been applied in gamified protocols, highlighting a methodological gap and a clear opportunity for innovation. Objective evaluation supplements the clinical evaluation [16].

This shortcoming motivated us to develop a serious game, so-called RehaBEElitation (Institute of Industrial Property—INPI registration: BR512021001975-0), which features 4 phases, to assess bradykinesia objectively and collaboratively [17]. The system uses a custom wearable glove interface with conductive threads and inertial sensors. According to the evaluation of the game's usability, RehaBEElitation is a serious game with an easy-to-understand narrative that meets adequately with the players' mental models, which aligns with the players' expectations of movement-control correspondence (eg, hand extension and flexion to move the bee up and down, adduction and abduction to move it left and right, and a pinch gesture to collect nectar—represented in the game as a droplet). This facilitated the interaction of individuals with the serious game and helped the system as a whole gain widespread acceptance [17,18].

This study extends a previous work [19] by evaluating bradykinesia across all 4 game phases (rather than just 2) using a time-based variable and introducing an improved human-machine interface (HMI) device. Additionally, because the signals captured by the HMI device consist of both the voluntary movement the person makes to play the game and the involuntary movement caused by the disease (ie, tremor), it is necessary to separate the involuntary movement from the detected signals since it could affect the slowness estimation. Another distinction in this work is the detailed explanation of the signal processing strategy used to exclude tremors from the collected data and, additionally, the inclusion of a correlation analysis between slowness-related variables and the clinical scale.

A key advantage of the RehaBEElitation serious game is the recording of real-time motion time series during interaction, enabling more in-depth and accurate analysis of the collected data. A recent literature review conducted by our group [20] revealed that previous studies evaluated PD symptoms or

monitored health conditions related to the disease using serious games combined with physical instruments, such as the Box and Blocks Test [21], the Purdue Pegboard Test [22], the Nine Hole Peg Test [23], the Time Up and Go test [24], the Maximum Step Length test [25], or the Trail Making Test [26]. None of these studies used data concerning the neuromuscular response of people with PD for the assessments. In the study, Sánchez-Herrera-Baeza et al [22], for example, players performed a test that assesses manual dexterity (the Box and Blocks Test) before and after interactions with the serious game developed, and the authors evaluated manual dexterity by comparing the results of the test performed at these different moments.

Similarly, Jäggi et al [27] conducted a pilot study with hospitalized patients with typical and atypical forms of PD using cognitive and motor exergames. Several cognitive tests (eg, Go/No-Go test, Reaction Time test, and Trail Making Test) and motor tests (eg, Preferred Gait Speed, Maximum Gait Speed, and Time Up and Go) were performed before and after playing the game to assess the motor and cognitive characteristics of the participants. Bégel et al [28] tested the effect of fine motor skills training, administered through serious games played on tablets, on the walking abilities of people with PD. Participants underwent motor performance tests before and after training with 2 games. The authors assessed the participants' gait and mobility using the Battery for the Assessment of Auditory Sensorimotor and Timing Abilities and the Beat Alignment Test, respectively. Finally, Ganzeboom et al [29] applied a serious game aimed at speech training in patients with PD-associated dysarthria. Each participant recorded 24 speech utterances before and after playing the game. The participants' speech was then evaluated for intelligibility by untrained listeners, comparing it with corresponding utterances made by a healthy speaker.

In contrast, the system developed in this study captures inertial sensor data that reflect neuromuscular activity to control the game avatar, enabling the objective quantification of motor symptoms, such as bradykinesia, through the extraction and analysis of variables that characterize the symptom. This distinctive capability constitutes a key differentiator of the present research. Accordingly, the purpose of this study was to objectively assess slowness in people with PD through their interaction with the RehaBEElitation serious game, using neuromuscular data to provide a precise and engaging way to quantify motor symptoms. Importantly, our intention is not to replace the clinical evaluation provided by the MDS-UPDRS, but rather to complement it by adding a fine-grained, performance-based measure obtained directly from the participants' motor behavior during the game.

Methods

This is an observational and cross-sectional study focused on the objective assessment of slowness, one of the main characteristics of PD. The experimental design involves collecting data from healthy individuals and individuals with PD in 2 different conditions: before taking medication that

controls the motor symptoms of the disease (OFF state) and after administering this medication (ON state).

Ethical Considerations

This research was approved by the Ethics Committee for Human Research at the Federal University of Uberlandia (UFU; protocol: 43229921.8.0000.5152). All participants provided written informed consent prior to enrollment in the study. The informed consent included authorization for the use of anonymized data for research and publication purposes. The experimental protocol was conducted in accordance with national and international ethical guidelines for human research, such as the Declaration of Helsinki. All collected data were deidentified prior to analysis. No information that could directly or indirectly identify participants was stored or included in the manuscript. Data were handled in secure and access-controlled institutional repositories, following best practices for confidentiality protection. In addition, participants did not receive any monetary compensation for participation.

Serious Game and Interface Device

RehaBEElitation is a bee-themed serious game designed to rehabilitate and monitor people with PD. The bees represent diligence, commitment, and hard work, qualities that people with PD really value during their recovery. The game was developed by a multidisciplinary team using Unity 3D. The player has to control the movements of a bee in a 3D environment. The game tasks were designed to require users to perform the identical motions as those found in the reference tool used to evaluate individuals with PD (MDS-UPDRS Part III): hand opening and closing, wrist extension and flexion, wrist adduction and abduction, finger tapping, and forearm supination and pronation. RehaBEElitation presents 4 phases (Multimedia Appendix 1), and each one symbolizes a real-world bee worker's task [17]. The objectives of each phase are as follows:

- Phase 1: pollinating the flowers—The objective is to collect pollen from a flower and deposit it in another flower. To move the bee up and down, the player must perform the movements of wrist extension and flexion; and to move the bee to the left and right, the movements of wrist adduction and abduction. The player must move the bee to a flower containing pollen (indicated by yellow arcs) and close the hand to catch it. Then, with the hand closed, the player must move the bee to a flower that does not have pollen (with green arcs) and open the hand to deposit it.
- Phase 2: feeding the larvae—The objective is to feed the larvae. The bee moves around the scene (up, down, left, and right) only if the player's hand is closed. The player must place the bee in front of a larva and open and close his or her hand to feed the larva.
- Phase 3: collecting the nectar—The objective is to collect nectar from flowers. Flowers that have nectar are indicated by drops of water. The player must guide the bee to a flower that has nectar and perform the finger-tapping movement to collect it.
- Phase 4: drying the nectar—The objective is to dry the nectar to produce honey. The player must go to honeycombs containing nectar (indicated by a luminous reflection) and

perform forearm supination and pronation movement (as a single movement) to make the bee flap its wings faster and dry the nectar.

An HMI was created to promote communication between the movements of the real environment and the game ([Multimedia Appendix 2](#)). It consists of a glove with inertial sensors, housed in a small case attached to the glove on the back of the hand, capable of estimating hand orientation; and conductive thread sewn on the glove, toward the palm of the hand and the fingers, capable of estimating finger tapping movements and hand opening and closing.

Experimental Protocol

Overview

The experimental group (EG) in this study consisted of 15 participants with PD, while the control group (CG) consisted of 15 healthy participants. Age and gender pairs were formed among the group volunteers (10 men and 5 women; mean 66.27, SD 9.13 years; range 45 - 82 years). The eligibility criteria for participants were as follows: a clinically confirmed diagnosis of PD, be aged between 40 and 100 years of age, regular use of PD medication (typically levodopa and/or dopamine agonists), have a mild to moderate PD stage (Hoehn and Yahr levels I, II, and III) [30], score higher than 19 on the Mini Mental State Examination [31], not have severe hearing or vision impairments (hearing acuity was assessed using the Self-Declaration Questionnaire on Hearing Loss, and visual acuity was assessed using the Snellen Optometric Scale), and not have another neurological disease history.

A convenience sampling strategy was adopted due to the specific eligibility criteria and availability of participants in the recruitment context. The sample size was defined based on previous studies with similar experimental designs that investigated motor performance and bradykinesia or slowness of the upper limbs in PD, and which demonstrated adequate power to detect differences between groups in movement-related variables [22,27,28]. The survey was disseminated at care centers for people with PD and through the communication channels of the UFU. Those interested in participating contacted the research team directly to schedule an appointment.

Data collection was conducted in a controlled environment at the Centre for Innovation and Technological Assessment in Health at UFU between February and March 2022. Upon arrival at the site, participants were welcomed by the research team, who presented the study objectives and provided necessary clarifications before signing the Informed Consent Form. Only after formal agreement to participate and verification of the inclusion criteria did the experimental procedures and data collection begin.

The participants were instructed to sit on a chair a meter from the screen that displayed the serious game. The limb most affected by the disease, as identified by a clinical assessment, was then fitted with the HMI device. To isolate the necessary movements and prevent compensating with another body part during movement execution, the limb was placed on a forearm support ([Multimedia Appendix 3](#)).

The participants were asked to play each phase of the serious game for 5 minutes or until they fulfilled their objectives. Data collection was performed in the ON and OFF states of medication for the EG. The first session was performed with individuals in the OFF state, after arriving at the collection site, in the morning. The ON state occurs when disease symptoms are under medication control, and the OFF state occurs when symptoms are not properly controlled by medication [16]. The EG volunteers were instructed to arrive in the morning having abstained from their PD medication for at least 8 hours, ensuring the re-emergence of motor symptoms and standardizing the OFF condition. After completing the first session in the OFF state, the EG participants took their usual medication (levodopa-based, such as Prolopa BD or other Prolopa presentations), and approximately 45 minutes later (once the therapeutic effects began to emerge), they performed the second session in the ON state, when symptoms were expected to be adequately controlled by medication.

The MDS-UPDRS Part III [9] was applied to the EG by an independent physiotherapist, who was not involved in the development or execution of the intervention protocols. Assessments were conducted in each session, before starting the game, that is, in both the ON and OFF states of the medication. The MDS-UPDRS was not applied to the CG. The CG participants interacted with the RehaBEElitation serious game in a single session using their dominant upper limb. All participants played the game at the easiest difficulty level.

Slowness Evaluation

First, bradykinesia was clinically assessed using the scores related to items 3.4a to 3.6b of the MDS-UPDRS Part III for the EG in both the ON and OFF states of medication. These items assess: 3.4a and 3.4b—tapping the fingers of the right and left hands, respectively; 3.5a and 3.5b—movements of the right and left hands, respectively; and 3.6a and 3.6b—pronation-supination movements of the right and left hands, respectively. The average scores of these items were considered for each medication state. In addition, slowness was objectively assessed for all participants of the study in 2 ways: by calculating the RT and by estimating the individuals' angular velocity (AV) of movement. Methodological details on how each variable was estimated are provided in the subsequent sections.

As described previously, RT involves both reaction and movement time, since RT is equal to reaction time added to movement time ([Multimedia Appendix 4](#)). Across all phases, when the participant correctly positioned the avatar on the target, a stimulus was presented to indicate the precise moment for movement execution. Accordingly, participants learned to align the avatar with the designated location and to initiate the movement only after stimulus onset.

AV (in degrees per second) represents the speed of movement execution [32] (ie, in relation to the RT variable, AV is linked to movement time). For a given movement, the higher the AV, the shorter the time taken to perform the voluntary movement. Thus, if it is verified that the velocity of movement execution was reduced after the use of medication, that is, in the ON state,

the time to perform the movement increased, which may suggest a failure in the medication effect.

In addition, RT allows considering aspects related to the player's reaction time, such as the learning acquired after the first section of interaction with the game. The participant may, for example, decrease his reaction time in the second game session (in the ON state) by learning that, immediately after receiving the visual and sound stimuli, he should perform the movement required to score points in the game in the shortest possible time.

Descriptive and inferential statistical analyses were performed to compare RT and AV between the EG and CG. The normality of data distribution was verified by the Shapiro-Wilk test prior to comparisons between groups, and appropriate parametric or nonparametric tests were adopted according to this result. In this study, the nonparametric Kruskal-Wallis test was used, followed by Wilcoxon signed rank tests with Bonferroni correction for pairwise group comparisons. The effect size was estimated using eta square (η^2) and interpreted according to Cohen criteria [33], with η^2 values of approximately 0.01 considered small, around 0.06 considered moderate, and values of ≥ 0.14 considered large.

Spearman rank-order correlation analyses were conducted to examine associations between the RT and AV variables and the MDS-UPDRS motor items previously described (items 3.4a/b, 3.5a/b, and 3.6a/b), as well as with the summed score of these items (BRADINDEX). For each participant, the median value of the estimated RT and AV data was calculated and used in the correlation analyses. The correlation coefficient (ρ) was classified into 5 categories: very weak (0.00 - 0.19), weak (0.20 - 0.39), moderate (0.40 - 0.59), strong (0.60 - 0.79), and very strong (0.80 - 1.00) [34]. Signal processing and statistical

procedures were performed using the R language for statistical computing (R Foundation for Statistical Computing).

Matching by age and gender during recruitment was used to minimize possible confounding factors, and no subgroup or interaction analyses were conducted due to the exploratory nature of the study and the matching strategy used. There were no missing data, as all attempts were successfully completed and validated during acquisition. Considering that sampling was by convenience, statistical analyses considered independent samples without additional weighting. Sensitivity analyses were not necessary, as all data were maintained and processed consistently throughout the study.

RT Estimation

RT was measured by defining a binary event in the game that signaled when the player approached the target and executed the movement. This event began when the player was supplied with visual and auditory stimuli, indicating that he or she could perform the movement and ended when the player performed the movement and scored in the game. Figure 1 shows the game screen (during phase 1) at the moment when the event responsible for calculating the RT occurs.

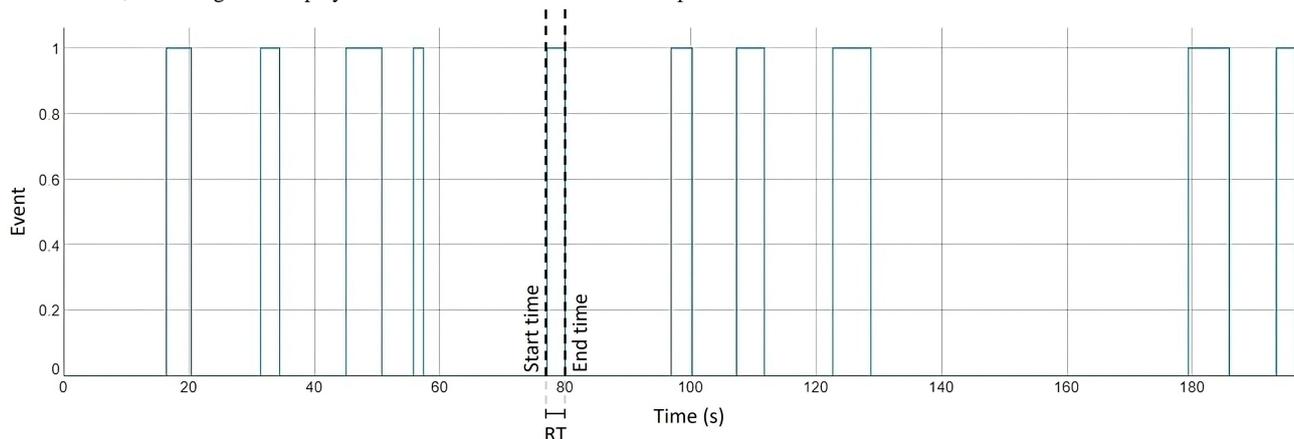
Figure 2 illustrates the occurrence of the described event of a volunteer from the EG in the OFF state during phase 4. This player arrived at the target and executed the movement 10 times (10 occurrences of the binary event). The start and end of each event were identified to determine the start and end times of the event. The duration of each event was then estimated by calculating the difference between the end and start times.

Thus, as RT is the sum of reaction time and movement time, both reaction time and movement time were taken into consideration to estimate the slowness of the individuals. This evaluation was performed for the 4 phases of the game.

Figure 1. Game screen showing the event to response time estimation. (A) Game screen while the player freely controlled the bee in the scenario. In this case, the player has not yet approached the target. (B) Game screen when the player reached the target and received the visual and sound stimuli, indicating that he/she could start executing the required movement. At this point, the bee remained static, the screen subtly changed color, a virtual hand appeared in the right corner of the screen, indicating to the player the movement to be performed, and a sound effect was emitted.



Figure 2. RT estimation from the defined event. Each blue rectangle represents an occurrence of the event, which indicates the duration of the RT. As stated, in this phase, the player hit the target and completed the expected movement 10 times, producing 10 RT values; the fourth occurrence showed the minimum RT, indicating that the player was faster in this instance. RT: response time.



AV Estimation

Overview

Gyroscope signals were used to measure the AV around a reference axis, which indicated how fast the movement was performed [32]. This characteristic was estimated only for phase 4 of the game because it is the only phase that requires the execution of a movement performed around one of the measurement axes of the inertial sensor (forearm supination and pronation) to score in the game. In [Multimedia Appendix 5](#), which shows the reference axes of the inertial sensor, it is observed that the supination and pronation movements occur around the Y axis.

As previously described, the motor symptoms of PD can be detected by inertial sensors and recorded as time-series data. These movement signals recorded by the sensors include both voluntary movements of the player to control the bee in the scenario and involuntary movements caused by the disease (such as tremor). The presence of involuntary movements may interfere with the estimation of slowness; hence, they should be disregarded from the collected signals. Thus, to extract only information related to players' voluntary movements, it was necessary to use a signal processing method capable of performing time series decomposition.

Time series decomposition involves decomposing a time series into several components, each representing one of the underlying oscillatory patterns in a signal. Singular spectrum analysis (SSA) is a decomposition method frequently used in several fields, such as economics, physics, and geography [35]; however, no study applying this method to signals collected from individuals with PD was found in the literature. In view of this gap and the technique's capacity to robustly decompose complex temporal signals, the SSA method was adopted in this study.

Singular Spectrum Analysis

SSA is a nonparametric method for the analysis of time series and digital images. This method decomposes a time series into a set of components that are grouped and interpreted, and the original time series is recovered by summing all its components [35]. SSA can be used to remove all periodic and trend components from a time series, leaving only a particular

component, which may have real and interpretable meaning [35]. In this study, decomposition was performed so that slowness could be evaluated objectively by using only the signals of voluntary movement, that is, excluding the influence of tremor.

A general description of the method is presented below [35-37]. SSA performs the 4 steps illustrated in [Multimedia Appendix 6](#). The input X is an ordered collection of N real numbers (eg, a time series or a digital image), and the output is a decomposition of X into a sum of identifiable components: $X = X_1 + \dots + X_m$.

- Step 1: embedding: The starting point of SSA is the embedding of the time series X into a vector space of dimension L . As an example, $X = (x_1, \dots, x_N)$ and $T = \text{SSA}$ maps \mathbb{R}^N to the space of Hankel matrices $L \times K$ with equal values on the antidiagonals. Thus, the embedding procedure constructs a sequence of vectors of the original time series using lagged copies of the scalar data. N is the length of the series, L is the length of the window (which is a parameter), and $K = N - L + 1$.

$$T(\text{SSA})(X) = (x_1 x_2 \dots x_L, x_2 x_3 \dots x_{L+1}, \dots, x_{N-L+1} x_{N-L+2} \dots x_N)$$

- Step 2: decomposition: The result of this step is the decomposition $X = \sum_i \sigma_i U_i V_i^T$, where $U_i \in \mathbb{R}^L$ and $V_i \in \mathbb{R}^K$ are vectors such that $\|U_i\| = 1$ and $\|V_i\| = 1$ for all i and σ_i are nonnegative numbers. This step performs the singular value decomposition (SVD) of the trajectory matrix and represents it as a sum of rank-one biorthogonal elementary matrices. Let $S = XX^T$, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d \geq 0$, be eigenvalues of the matrix S , $d = \text{rank } X = \max\{j: \lambda_j > 0\}$, U_1, \dots, U_d be the corresponding eigenvectors, and $V_j = X U_j / \lambda_j$, $j = 1, \dots, d$ be factor vectors. Denote $X_j = \lambda_j U_j V_j^T$. Then, the SVD of the trajectory matrix X can be written as $X = X_1 + \dots + X_d$. The triple (λ_i, U_i, V_i) consisting of the singular value $\sigma_j = \lambda_j$, the left singular vector U_j , and the right singular vector V_j of X is called the j th eigentriple.
- Step 3: grouping: The grouping step corresponds to the division of the elementary matrices X_j into several groups and summing the matrices within each group. The grouping procedure partitions the set of indices $\{1, \dots, d\}$ into m disjoint subsets $\{I_1, \dots, I_m\}$, with each group I_k containing

a set of principal components $\{i_1, \dots, i_p\}$, representing specific components of the signal. Let $I_k = \{i_1, \dots, i_p\}$. Then, the resulting matrix X_{I_k} corresponding to the group I_k is defined as $X_{I_k} = X_{i_1} + \dots + X_{i_p}$. These matrices are computed for $I_k = I_1, \dots, I_m$, and the SVD expansion leads to the decomposition $X = X_{I_1} + \dots + X_{I_m}$. Each X_{I_k} represents a set of eigentriples, which describes a specific component series in the original time series.

- Step 4: reconstruction: The last step is to reconstruct the components of the original series. This is achieved by diagonal averaging of each X_{I_k} to provide the k th component of the series X , where the n th sample is obtained by averaging over the cross-diagonal $i + j = \text{const} = n + 1$ of X_{I_k} . This is because each X_{I_k} can be seen as the Hankel matrix for the corresponding embedded component series. Then, the resulting decomposition of the initial object X is $X = X_1 + \dots + X_m$.

The SSA procedure was applied to identify the tremor-related oscillatory components, which were subsequently removed from the original time series. The resulting residual signal, obtained by subtracting the tremor components from the original data (since the original signal can be reconstructed as the sum of all SSA-decomposed components), represents the tremor-free voluntary movement used for slowness evaluation.

The estimate of SSA was computed in the R language by using the method `ssa` (from the *Rssa* package), which constructs the `ssa` object that holds the decomposition and various auxiliary information. The typical call of the `ssa` function in R is:

```
s <- ssa(x, L = (N+1) %/% 2, neig=NULL)
```

The main arguments are (N is the length of the series):

- x is an object to be decomposed (eg, time series).
- L is a window length, which by default is fixed to half the length of the series. However, it is also recommended to use a proportional value to the periodicity of the time series. Very high values of L may mix the trend component with the periodicity components. Therefore, $L=250$ was used (the sampling frequency of the data).
- $neig$ is the number of eigentriples desired. If $neig=NULL$, a default value, which depends on L and N , will be used. It was considered $neig=NULL$.

The function returns an `ssa` object. The precise layout of the object is hidden, but identifying the indices of the elementary components is necessary for the next step (reconstruction). This identification is performed to find the selection parameters of the principal components to group the elementary time series. The indices of the elementary components can be identified by

using the `$` operator and by analyzing the plots of the eigenvalues and the correlation matrix of the eigenvalues, namely, `s$sigma` returns the vector of single values, `(s$sigma)^2` returns the vector of eigenvalues, `plot(s)` plots the eigenvalues, and `plot(wcor(s))` plots the w -correlation matrix.

By definition, components that have similar eigenvalues are used for the grouping of elementary time series that represent the periodic or oscillatory components of the original time series. In other words, eigenvalues with very similar magnitudes are associated with a periodic or sinusoidal component of the signal, in this case, the PD tremor, which represents an involuntary movement.

The reconstruction step is performed with the `reconstruct` function. The typical call in R is:

```
r <- reconstruct(s, groups=list(Tremor=c(3:5, 13:16))
```

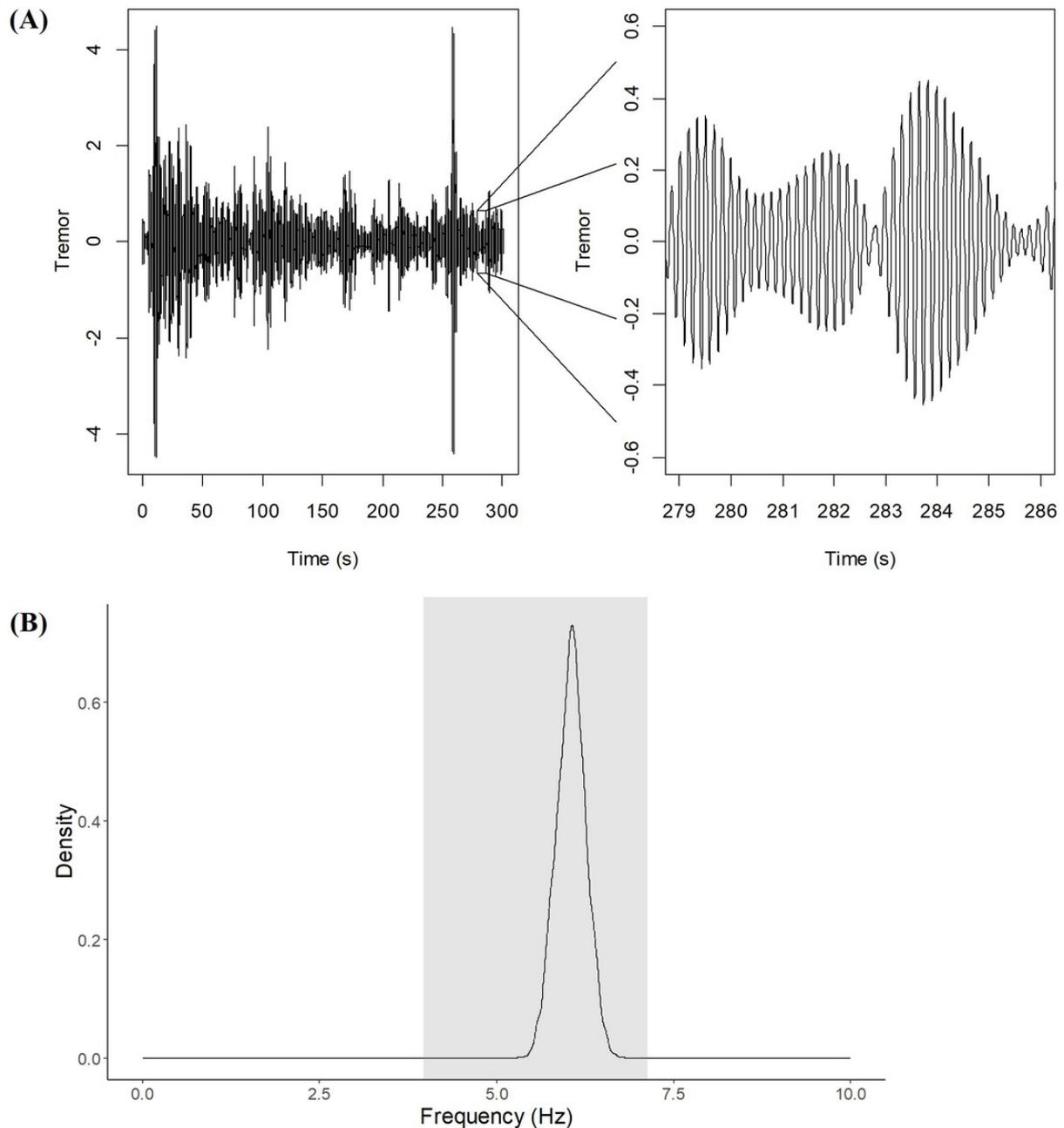
The main arguments are as follows:

- s is an `ssa` object that holds the decomposition.
- `groups` is a list of numeric vectors consisting of the indices of the elementary components used for the reconstruction; the list entry can be named, and in this example, it was called `Tremor`. The elementary components with indices 3, 4, and 5 and 13, 14, 15, and 16 (which have similar magnitudes) represent the oscillatory component of the time series, which characterizes the player's involuntary movement (`tremor`).

Thus, SSA was applied to the original signal collected during the experiment to extract the tremor. The residual signal after extraction of the involuntary movement represents the voluntary movement used to quantify the slowness of the participants. In other words, the decomposition allowed us to exclude the frequency range associated with tremor while retaining the frequencies representing voluntary motion, which were then used to quantify participants' slowness. [Multimedia Appendix 7](#) illustrates the decomposition of a gyroscope signal from a participant of the EG. The sum of the voluntary movement and tremor signals results in the original signal.

[Figure 3](#) illustrates a typical tremor signal resulting from the use of SSA. An enlarged view of the tremor signal extracted by the decomposition is shown in [Figure 3A](#) (on the right), and the power spectral density of the tremor is depicted in [Figure 3B](#). [Figure 3B](#) shows that the density of the tremor extracted by the decomposition ranges between 4 and 7 Hz, which is the frequency range of a tremor originating from PD. This indicates that the method is robust in cleanly separating the oscillatory components of the original signal that truly correspond to PD tremor.

Figure 3. Characteristics of the signal extracted by the singular spectrum analysis decomposition method. **(A)** On the left, the complete tremor signal, and on the right, a zoomed view of the signal. **(B)** Power spectral density of the tremor, confirming that it corresponds to a Parkinsonian tremor, as its dominant frequency lies within the characteristic 4 - 7 Hz range.

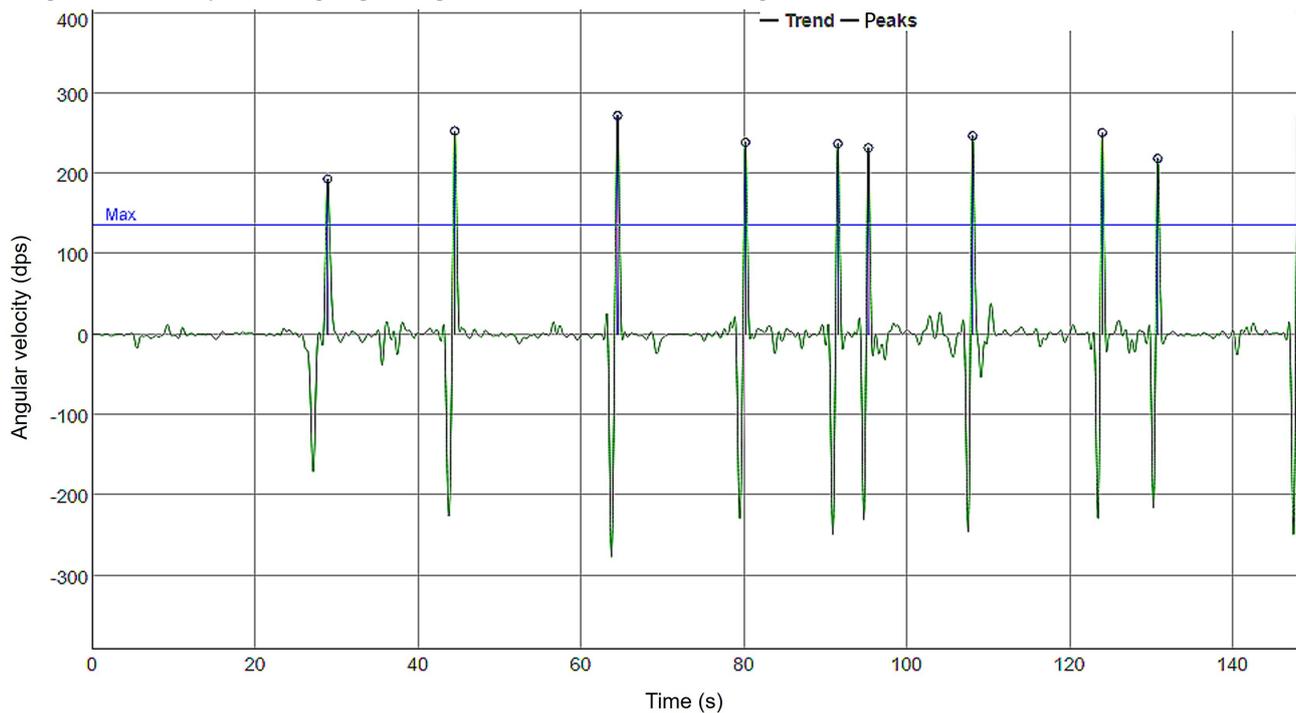


By minimizing the contribution of the higher-frequency components associated with tremor in the spectral profile of the movement, it becomes possible to isolate the signal that truly represents the participant's voluntary motion. Removing these high-frequency components is essential to ensure that slowness, which is manifested within the voluntary movement itself, can be estimated in a pure and uncompromised manner, without interference from other motor symptoms such as tremor. This allows the control signal used in the game to reflect only the

participant's voluntary movement characteristics, particularly their slowness.

After applying SSA for signal decomposition, the peaks of the signals representing the voluntary movement were identified and discriminated by using a threshold of 0.5 to capture only the regions of interest and evaluate the participants' slowness of movement, as depicted in Figure 4. The higher the peak value, the faster the movement was executed.

Figure 4. Peak detection in the gyroscope signal, representing the angular velocity of hand movement. Each peak corresponds to a directional change or a complete movement cycle, and higher peak amplitudes indicate faster and more vigorous motion execution.



On the basis of the processes outlined earlier, the influence of involuntary movement arising from PD was mitigated for the assessment of slowness.

To minimize potential sources of bias, certain methodological precautions were implemented. First, standardized protocols were applied during all data collection procedures, including instructions, task execution, and environmental conditions, ensuring consistency among participants. Assessment sessions were conducted individually to avoid external interference and influence from other participants' performance. Disease severity was controlled by including only individuals classified as having mild to moderate PD (Hoehn and Yahr I-III), thereby reducing heterogeneity in motor impairment. After taking their regular antiparkinsonian medication, the researchers waited for a period of time until each participant subjectively reported improvement in motor symptoms, indicating perception of adequate motor control caused by the medication, minimizing drug-related performance fluctuations. In addition, data processing followed predefined criteria, and the researcher responsible for extracting and organizing the results was unaware of the participants' identities to reduce observer bias.

Results

All individuals who expressed interest in participating in the study and were evaluated for eligibility criteria met the established requirements. In total, 15 participants with PD and 15 healthy participants were included and analyzed in the study.

Samples from healthy participants were collected only after the participation of an individual with PD of the same age and gender, ensuring pairing between the groups. There were no dropouts or exclusions during the collection process, so all participants fully completed the experimental protocol.

The primary demographic and clinical characteristics of the sample are presented in [Table 1](#). To assess potential differences in age between the PD group and the healthy CG, an ANOVA was conducted, showing no statistically significant difference between the groups ($P > .05$). The Kruskal-Wallis rank-sum test was applied to compare the total MDS-UPDRS Part III scores in the ON and OFF medication states, demonstrating a significant difference between these conditions ($P < .05$). Additionally, an ANOVA performed on the Hoehn and Yahr scores across medication states indicated no significant variation ($P > .05$).

[Table 2](#) displays the slowness assessment results for both the EG and CG. The average RT and AV values for each participant, considering the 4 phases of the game, are presented. The RT values were generally lower for the CG than for the EG, while the AV values were higher for the CG than for the EG in both the ON and OFF states. It is noteworthy that the CG participant with the highest RT also showed the lowest AV (participant 3). However, it is important to emphasize that RT and AV do not have a linear relationship; changes in one measure do not necessarily correspond proportionally to changes in the other, as they capture different aspects of movement performance.

Table. Main characteristics of study participants, including age, Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) Part III total score, and Hoehn and Yahr (H&Y) stage.

Control group		Experimental group					
ID	Age (years)	ON state				OFF state	
		ID	Age (years)	MDS-UPDRS Part III	H&Y	MDS-UPDRS Part III	H&Y
1	82	1	57	25	3	61	2
2	64	2	80	25	2	49	2
3	80	3	64	26	2	58	2
4	66	4	77	31	2	44	2
5	51	5	59	24	3	47	2
6	56	6	64	39	2	48	2
7	77	7	70	20	2	41	2
8	62	8	73	42	3	45	3
9	60	9	60	12	2	29	1
10	80	10	66	16	1	29	1
11	69	11	69	37	3	70	2
12	59	12	60	31	2	54	2
13	62	13	50	35	3	59	3
14	55	14	68	38	2	52	2
15	71	15	77	11	1	38	1

Table . Results of slowness evaluation^a.

Control group			Experimental group						
ID ^b	RT ^c (sec- onds)	AV ^d (de- grees per second)	ON state				OFF state		
			ID	MDS-UP- DRS ^e	RT (sec- onds)	AV (degrees per second)	MDS-UP- DRS	RT (sec- onds)	AV (degrees per second)
1	1.13	308.38	1	3	1.66	305.35	5	3.30	139.49
2	1.30	347.34	2	6	3.73	139.20	9	2.53	177.71
3	3.27	287.36	3	8	2.08	248.31	8	1.63	182.14
4	2.00	400.60	4	6	1.84	351.83	9	1.43	394.17
5	1.59	460.94	5	2	1.98	398.71	3	2.60	306.59
6	1.57	520.40	6	6	1.94	245.04	9	2.42	182.70
7	2.50	391.27	7	4	1.14	234.44	8	2.20	238.86
8	1.95	401.34	8	3	2.15	234.16	6	1.99	301.79
9	1.49	621.23	9	3	1.42	473.47	5	1.93	352.31
10	1.33	506.88	10	2	1.22	272.14	4	1.45	316.26
11	1.70	466.61	11	6	1.26	492.09	9	1.84	327.66
12	1.49	555.35	12	6	2.94	167.36	12	3.70	135.72
13	1.60	565.49	13	6	1.90	152.37	9	1.41	170.01
14	1.49	518.94	14	5	1.90	260.84	9	2.04	252.35
15	1.48	383.09	15	3	1.98	374.44	7	3.51	278.41

^aHigher RT and lower AV indicate greater slowness, and higher MDS-UPDRS scores reflect worse motor impairment.

^bID denotes participant identifier.

^cRT: response time.

^dAV: angular velocity.

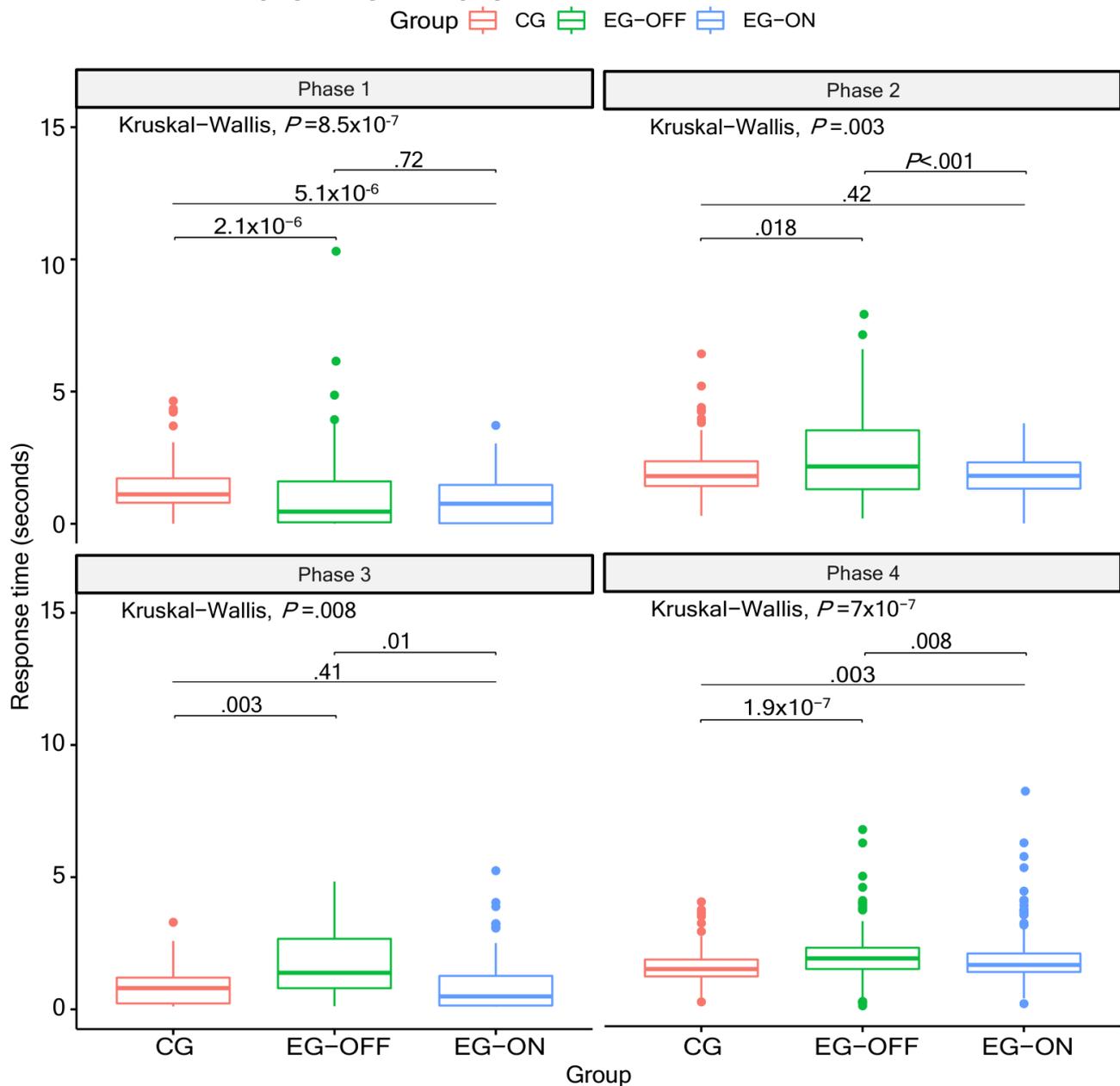
^eMDS-UPDRS: Movement Disorder Society-Unified Parkinson's Disease Rating Scale.

Regarding the EG, participants 1, 5, 6, 9, 11, 12, and 15 in the ON state of medication presented lower RT scores and higher AV scores compared to the OFF state. In the OFF state, the individual with the highest degree of disease impairment, according to the clinical evaluation, also presented the largest RT and the smallest AV (participant 12). Participants 3 and 14 increased their AV in the ON state compared to the OFF state but also increased their RT; and participants 2, 7, and 10 decreased their RT in the ON state compared to the OFF state

but also decreased their AV. Participants 4, 8, and 13 increased their RT and decreased their AV in the ON state compared to the OFF state.

The distribution of RT values for all participants is depicted in [Figure 5](#). Except for phase 1, in all other phases, the CG had the lowest RT, followed by the EG in the ON and OFF states. In phase 1, the EG in the ON state produced the best RT values, followed by the CG and then the EG in the OFF state.

Figure 5. Distribution of estimated response time across participants for each of the 4 game phases, comparing the CG and the EG in the ON and OFF medication conditions. CG: control group; EG: experimental group.

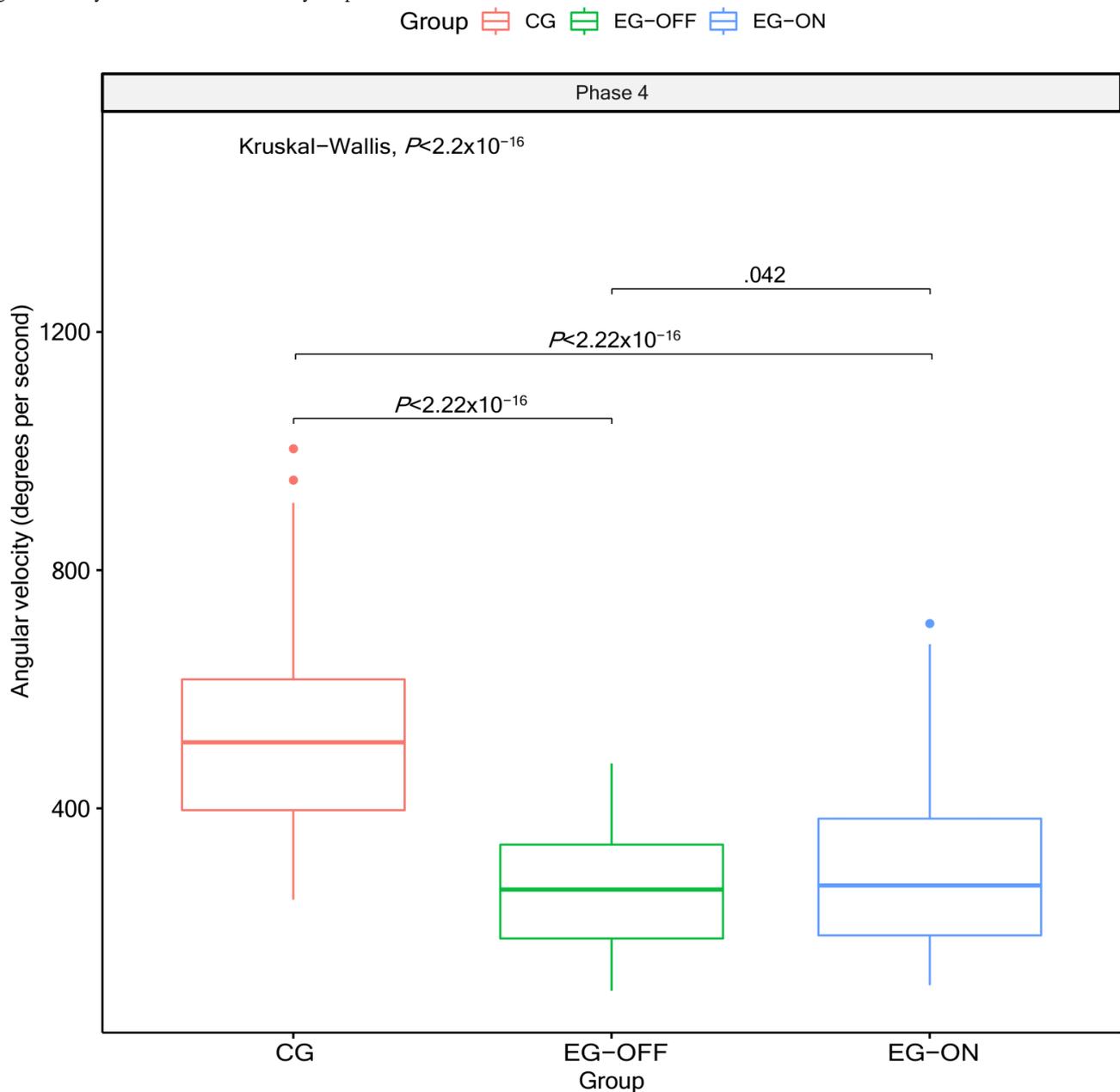


The Kruskal-Wallis test indicated statistically significant differences in RTs among the groups across all phases. Specifically, phase 1 showed $P=8.5 \times 10^{-7}$ with a small-to-moderate effect size ($\eta^2=0.0591$), and statistically significant differences were found among all groups except between the EG-OFF and EG-ON; phase 2 showed $P=.0028$ with a small effect size ($\eta^2=0.0239$), and no statistically significant differences were observed between the CG and the EG-ON; phase 3 showed $P=.0081$ with a small effect size

($\eta^2=0.0419$), and the comparison between the CG and the EG-ON did not reach statistical significance; and phase 4 showed $P=7 \times 10^{-7}$ with a moderate effect size ($\eta^2=0.0650$), and statistically significant differences were observed among all groups.

The distribution of AV data for all participants is shown in Figure 6. Again, the CG had the largest AV, followed by the EG in the ON state, and then by the EG in the OFF state.

Figure 6. Distribution of angular velocity among participants, comparing the control group and experimental group in both medication states. As noted, angular velocity was estimated exclusively for phase 4.



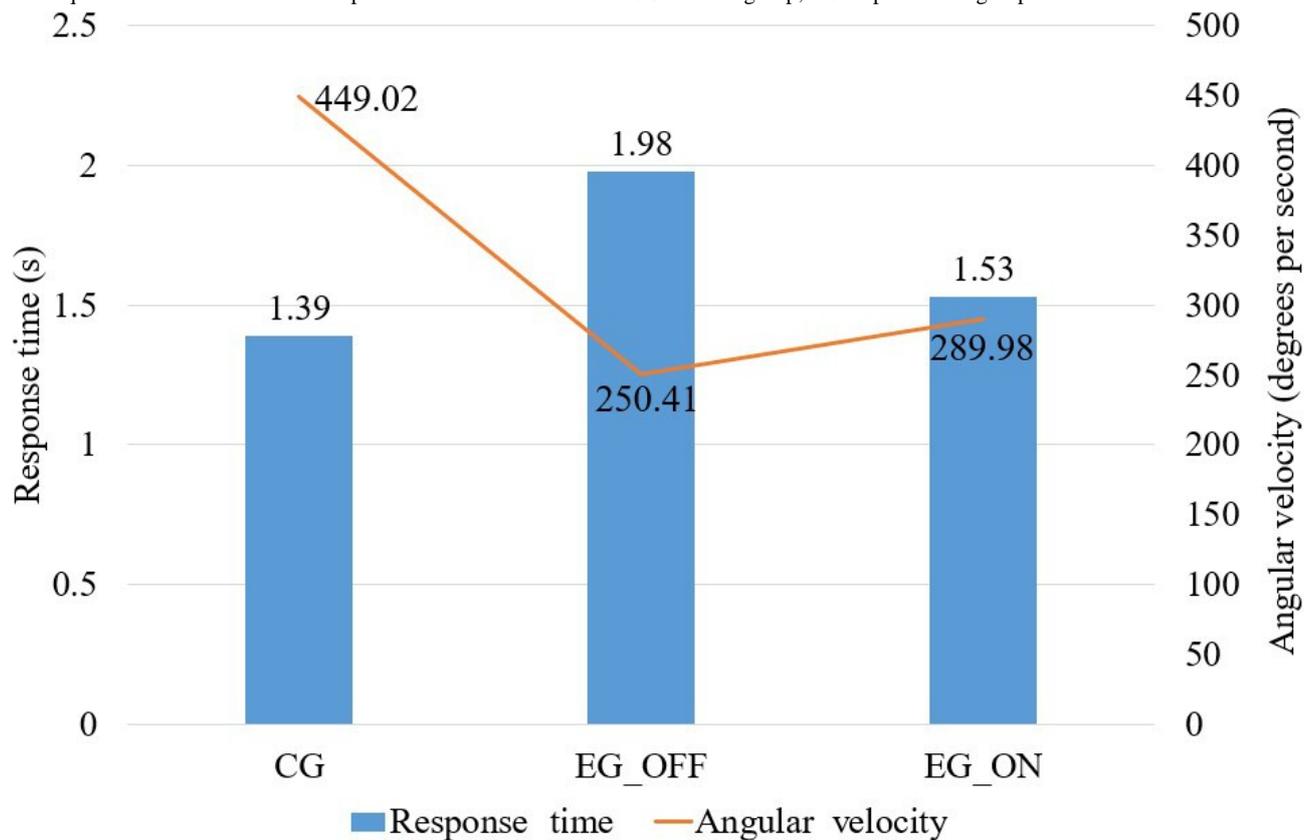
A statistically significant difference in AV among the groups was observed using the Kruskal-Wallis test ($P < 2.2 \times 10^{-16}$), with a large effect size ($\eta^2 = 0.404$).

It is worth noting that the analyses presented in Figures 5 and 6 were not conducted based on individual trials but rather using aggregate measures per group. Specifically, both figures were constructed based on the median values of each variable within each group, ensuring that differences in the number of trials per participant did not influence the results. Thus, Figure 5 represents the median RT per group in each phase, while Figure 6 presents the median AV per group in phase 4. This procedure

ensures a more robust and consistent interpretation of the data, minimizing biases related to intraindividual variability.

Figure 7 shows the mean values of RT and AV for all groups. This grouped visualization of the estimated variables (RT and AV) reflects an overall trend in motor performance between the groups. On average, healthy individuals performed faster movements in less time during interaction with the game when compared to participants with PD. In addition, for the EG, the RTs decreased from the OFF state to the ON state, and AVs increased from the OFF state to the ON state. This visualization highlights the overall difference between the groups in terms of RT and AV measurements.

Figure 7. Mean response time and angular velocity values for each group. The CG showed the fastest responses and highest movement speeds, while EG exhibited improved performance in the ON state compared to the OFF state, with lower response times and higher angular velocities. These values provide a quantitative overview of motor performance in all conditions. CG: control group; EG: experimental group.



The correlation analysis revealed distinct association patterns between RT and AV and the motor items of the MDS-UPDRS. Table 3 summarizes the corresponding correlation coefficients along with their classifications. RT showed very weak to weak

positive correlations with all individual items and with BRADINDEX. In contrast, AV demonstrated strong negative correlations across all MDS-UPDRS items and with the composite bradykinesia index.

Table . Spearman correlations (ρ) between response time (RT), angular velocity (AV), and the Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) items.

MDS-UPDRS item	ρ (RT×MDS-UPDRS)	ρ (AV×MDS-UPDRS)
3.4a	0.058	-0.656
3.4b	0.217	-0.617
3.5a	0.114	-0.687
3.5b	0.216	-0.625
3.6a	0.118	-0.725
3.6b	0.215	-0.665
BRADINDEX ^a	0.147	-0.698

^aBRADINDEX is the summary of the 3.4a to 3.6b item scores.

Discussion

On the basis of movement data collected by inertial sensors during participants' interaction with the serious game, it was possible to measure indicators of slowness in PD. These results demonstrate that metrics extracted from a gamified environment can capture motor changes characteristic of the disease and differentiate individuals with PD from participants in the CG.

Comparison to Prior Work

Unlike existing literature that relies on subjective scales or external motor tests, RehaBEElitation records inertial signals during gameplay for dual-purpose rehabilitation and monitoring. While the MDS-UPDRS remains the gold standard, its high interexaminer variability and inability to capture daily symptom fluctuations highlight the need for more reliable, objective tools [38,39]. Our system addresses this by capturing direct

neuromuscular data through a sensor-integrated glove, ensuring a more precise assessment of impairment.

Despite not addressing the use of serious games, recent studies have demonstrated the use of wearable sensors to assess and monitor the motor signs and symptoms of PD [40,41]. Such technologies can help researchers and health care professionals obtain a more reliable and objective view of disease progression and the severity of motor symptoms. Furthermore, according to Bhidayasiri and Tarsy [16], the successful implementation of wearables to evaluate PD symptoms, in conjunction with the use of established scales such as the MDS-UPDRS, promises to improve the quality of medical care for people with PD. However, despite numerous efforts to create new technologies for bradykinesia assessment, it has not yet been possible to establish one that performs this evaluation objectively, as considerable experience and effort are still required to operate the available technologies [10].

In this study, slowness was assessed using the individuals' RT, which was calculated by the duration of an event indicating the moment when the player reaches the target and performs the movement, as well as the gyroscope signal collected by the wearable HMI device, which represents the AV of movement. Statistical analysis indicated that the groups of healthy individuals and patients with PD were comparable in terms of age. Overall, as reflected by the MDS-UPDRS III and Hoehn and Yahr scores (Table 1), patients with PD in the OFF medication state exhibited greater impairment than in the ON state.

To obtain a reliable AV estimation, the SSA method was applied to the gyroscope signals to isolate voluntary activity. Several techniques can be used for time series decomposition, such as Fourier analysis, empirical mode decomposition, wavelet transform, kinematic decomposition by submovements, and singular spectrum decomposition (SSD). However, as the recorded signals are nonstationary time series, Fourier analysis is not the ideal option [42]. This approach assumes stationarity, risking the loss of transient information [43]. Moreover, Fourier analysis and empirical mode decomposition can provide data decomposition results with limited physical sense, that is, components with unclear physical meaning. The components retrieved by these techniques can also be affected by mode mixing, again interfering with the physical meaning of individual components, especially in cases where abrupt variations in movement occur [37]. Techniques such as wavelet transform, although suitable for nonstationary data, depend heavily on the prior choice of a mother wavelet, which can introduce bias and hinder the physiological interpretation of the extracted components [40]. The kinematic decomposition by submovements approach is limited by its dependence on ideal models that do not always represent real movements and by its sensitivity to noise during segmentation, which can lead to the misidentification of submovements [44]. Finally, the SSD technique was initially applied to decompose the original signal, but no successful results were obtained. The selection of fundamental parameters in SSD, including window length and elementary component indices, is completely automated, which generated components with no physical meaning.

The SSA method was able to satisfactorily disregard the signal that represents the involuntary movement of the players. The results suggested that this method is particularly useful for this application, as the extracted tremor signal represents a disease-induced tremor (Figure 3). Due to the decomposition, it is possible to separately and independently study the characteristics of the pure movement and of the tremor of individuals with PD by analyzing only the extracted signal of interest.

Principal Findings

The results indicated in Table 2 showed that participants 1, 5, 6, 9, 11, 12, and 15 in the ON state of medication presented lower RT scores and higher AV scores compared to the OFF state. This result was already expected, since EG participants in the ON state of medication have more reduced motor symptoms than in the OFF state [45]. However, in the ON state, some situations also occurred: participants 2, 7, and 10 reduced RT but also reduced AV, participants 3 and 14 increased AV but also RT, and participants 4, 8, and 13 increased RT and reduced AV. It is known that RT is the sum of reaction time and movement time; AV is more related to movement time.

After taking the medication, participants 2, 7, and 10 had lower RT but performed the movements more slowly; that is, they increased the movement time, which consequently resulted in a decrease in reaction time. Movement time may have increased because the medication is probably not bringing benefits regarding the improvement of motor symptoms in these individuals, which suggests the need to readjust the dosage of the medication or change it for another medication. The reaction time may have been reduced by the learning acquired in the first session of interaction with the game, that is, the players learned that after reaching the target and receiving the visual and sound stimuli, they should perform the required movement. Thus, although the player performed the movements more slowly, he was able to initiate the execution of the movements earlier.

After participants 3 and 14 took the medication, the movements were performed faster, that is, movement time was decreased, but RT increased, which consequently resulted in increased reaction time. Movement time may have been reduced due to the medication's effectiveness in reducing motor symptoms. The reaction time may have increased due to the lack of learning that could have been acquired in the first session; that is, the participants did not understand that the visual stimulus indicated the exact moment to start the execution of the movement. Thus, although the motor symptoms were reduced, the players did not learn enough, making them start the execution of the movements later.

Finally, after participants 4, 8, and 13 took the medication, their RT increased, and they performed the movements more slowly, that is, they increased the movement time. It can be suggested that the medication does not reduce the motor symptoms of these individuals; therefore, a readjustment of the medication dosage should be performed. Moreover, these individuals probably did not acquire learning in relation to the game.

Medication is the key treatment for the motor symptoms of PD, but the pharmacological response inevitably diminishes over time [46]. Following these results, it is possible to identify the individuals who possibly need medication adjustments, as well as those who require additional training with the serious game to better understand its features and rules. These findings may provide a more personalized patient follow-up by health care professionals and assist them in assessing individuals and monitoring the disease.

In phase 1, the median RT of the EG-OFF was lower than that of both the EG-ON and the CG. The EG-ON also showed lower RTs compared to the healthy controls. However, no statistically significant differences were found between the EG-OFF and EG-ON groups, whereas both EGs differed significantly from the CG. One possible explanation for this pattern is that participants in the EG had already been exposed to the RehaBEElitation serious game, which may have provided a degree of familiarity with the task demands. This prior exposure may have facilitated faster motor planning and reduced hesitancy during the initial phases of interaction, even among participants in the OFF-medication condition. In contrast, healthy controls, despite not presenting motor symptoms, were inexperienced with the task and may have required additional time to internalize the stimuli and response mapping, resulting in comparatively longer reaction times. It is worth noting, however, that the CG exhibited a narrower or flatter boxplot distribution in phase 1, indicating lower variability and greater consistency in performance compared to the EGs. A similar trend was observed in phase 3, in which the EG-ON demonstrated lower RTs than the healthy controls, although without reaching statistical significance. Together, these findings suggest that task familiarity may have contributed to the observed differences in RT performance.

Furthermore, the information indicated by [Figures 5-7](#) showed that, in general, the RT and AV scores for the CG participants were better than those for the EG in the ON state, which in turn were better than those for the EG in the OFF state, as already expected. In addition, EG participants in the OFF state of medication and CG participants who presented the worst results for RT also presented the worst results for AV, which reveals that the measures used for the assessment of slowness were quite consistent.

According to [Figures 5](#) and [6](#), phase 4 is the most suitable for assessing slowness using the RehaBEElitation serious game. [Figure 5](#) shows that, in this phase, the differences between the groups were better detected, since the boxplots overlapped less, and only in phase 4 was it possible to estimate the intentional AV of the participants ([Figure 6](#)). Movements performed around the other axes, such as flexion-extension around the X axis and adduction-abduction around the Z axis, were not used to score points but to steer the bee within the scenario. Because these actions served navigation, players could deliberately modulate them (eg, moving more slowly or with reduced amplitude) to fine-tune the bee's positioning. Consequently, AV derived from these axes would not reflect performance intent and was not analyzed. In contrast, when a movement is executed to score points, it is performed with intentional speed and amplitude,

which are suitable for AV estimation. For this, AV analyses were restricted to the gyroscope signals recorded in phase 4.

A further point to discuss is that phase 4 required the execution of the forearm supination and pronation movement, a more complex movement than the others and one that is not often performed during daily activities. This may have encouraged players to focus on the execution of the movement, which may have influenced a better estimation of the variables used in the study. These findings agree with those reported by Summa et al [47].

Effect size estimates provide important information about the practical relevance of the differences observed between groups. Although all phases of the RT analysis showed statistically significant differences, the associated effect sizes were small for phases 2 and 3, small-to-moderate for phase 1, and moderate for phase 4, suggesting that group membership explains only a limited part of the variability in RTs. This indicates that, although the groups are different, other factors such as individual variability, familiarity with the task, or motor strategy may play a substantial role in RT performance. In contrast, the AV results revealed a large effect size ($\eta^2=0.404$), indicating that the differences between groups in AV are not only statistically significant but also of great practical magnitude. This pattern also confirms that AV is a more sensitive marker of group-specific motor behavior than RT and, therefore, may serve as a stronger discriminator of motor performance differences in this context.

The grouped visualization shown in [Figure 7](#) highlights an overall pattern of motor performance between groups but does not capture individual nuances or intrasubject variations. These differences may be related to the heterogeneity of symptoms among people with PD, variability in response to medication, level of familiarity with the game, motivation, rigidity, and different difficulties present in tasks at each phase of the game. Thus, although the overall trend indicates greater slowness in the Parkinson group, future studies should consider more detailed analyses at the individual level, as well as increasing the number of trials per participant, to reduce motor compensations or momentary fluctuations and obtain even more robust measures.

Correlation analysis revealed distinct patterns between RT and AV variables and the motor items of the MDS-UPDRS. While RT showed very weak to weak correlations with items that reflect slowness on the clinical scale, AV showed strong and consistent correlations, indicating, once again, greater sensitivity of this metric to the typical motor impairment of PD. In this study, slowness was estimated from the sum of reaction time and movement time (RT variable), although reaction time is not part of the clinical assessment of this symptom. This conceptual difference may have contributed to the weaker correlations observed between RT and MDS-UPDRS scores. Additionally, RT seems to reflect broader aspects of interaction with the task, such as attention, familiarity, and individual strategies, which may explain the lower magnitude of the associations.

Strengths and Limitations

An important contribution of this study is the definition of a specific event within the game that functions as a behavioral marker, allowing the estimation of a variable indicative of slowness (RT) during task execution. Although not a biological marker, this event provides a structured and reproducible reference point from which RT can be calculated. As already discussed, this variable, despite not showing a strong correlation with the MDS-UPDRS, captures a central component of motor slowness, as it reflects both the preparation and initiation of movement, processes known to be impaired in PD. Without this event, it would not be possible to quantify slowness with the RT variable in the simple way that was carried out. There is a total of 23 predefined markers in the game, which were detailed in a study describing the architecture and organization of our system [17]. Thus, the game's internal markers play a key role in providing objective anchors within the interaction, allowing the derivation of meaningful measures of motor performance.

Study limitations include the lack of prior training for the CG and the single-day assessment for the EG in both ON and OFF medication states. While this scheduling avoided collection bias, the cumulative duration of tasks may have impacted performance during the ON-state trials for some participants. Additionally, the PD group used their most affected limb, while the CG used their dominant limb. Although this introduces variability, it preserves ecological validity, as bradykinesia is characteristically asymmetric. Since our goal was to characterize game interaction under realistic clinical conditions rather than compare absolute performance, we consider the impact of this difference on our findings to be limited.

Future Directions

In this study, the generalization of the results is influenced by the sample characteristics. Although the use of a serious game that presents a well-established motor task favors broader applicability, the results should be interpreted, considering that participants presented mild to moderate PD (Hoehn and Yahr I-III) and that all interactions with the system were conducted in a controlled environment. Therefore, extrapolation to individuals with advanced disease, other clinical contexts, or unsupervised home use should be done with caution. Future studies with larger and more heterogeneous samples are needed to confirm the external validity of these results.

Acknowledgments

The developed serious game won the prize at the 11th edition of the “Challenge Handicap & Technologie” event in Metz (France) on May 12, 2023. The first-place prize was awarded in the “Autonomy” category, related to promoting the autonomy of people in situations of dependency [48]. The project also received the Brazilian Society of Biomedical Engineering (SBEB) Boston Scientific Award for Innovation in Biomedical Engineering 2023. The award, created in a partnership between the SBEB and Boston Scientific, aims to value and encourage the academic production of projects and work in biomedical engineering aimed at tackling the difficulties of the Brazilian health system and public health services in Latin America [49]. No generative artificial intelligence was used in any part of this manuscript.

Funding

The work was carried out with the support of the National Council for Scientific and Technological Development (CNPq—444437/2024-0, 442150/2023-7, and 405365/2023-3), the Coordination for the Improvement of Higher Education

In comparison to other methodologies, using serious games is an appealing option for measuring Parkinson motor symptoms. This game can be applied frequently and easily in clinics and hospitals, which demonstrates excellent accessibility and is cost-effective. In addition, the methods applied can isolate voluntary activity from involuntary activity. Severe tremor interferes with the classification of bradykinesia.

This study advances the current literature by proposing a novel and objective method to quantify slowness in PD using inertial sensor data acquired during interaction with a serious game. While previous studies have explored serious games for PD assessment and rehabilitation, most relied on subjective clinical scales, questionnaires, or conventional physical tests to evaluate motor symptoms. In contrast, our approach directly analyzes sensor-derived metrics obtained during gameplay, enabling the extraction of quantitative measures of RT and AV without interrupting the user experience. To the best of our knowledge, this is the first study to evaluate slowness in PD by processing inertial sensor data collected during interaction with a serious game. This contribution strengthens the role of serious games as potential digital biomarkers and highlights their value as engaging tools for continuous and sensitive motor assessment. From a real-world perspective, this approach may facilitate remote monitoring in home environments, complement clinical evaluations, and support personalized follow-up of motor symptom progression, thereby expanding access to objective and technology-driven care for individuals with PD.

Conclusions

The findings of this study suggest that the RehaBEElitation serious game is a promising tool for objectively assessing motor slowness in individuals with PD. Interaction-derived variables, particularly AV, showed strong correlations with MDS-UPDRS motor scores, capturing motor impairment sensitively. The game also detected differences between ON and OFF medication states, reinforcing the ability of the serious game to detect changes associated with treatment. Overall, these results indicate that the RehaBEElitation serious game represents a viable and potentially valuable approach for the objective monitoring of slowness and for supporting strategies for tracking the progression and therapeutic management of PD.

Personnel, and the Foundation for Research Support of the State of Minas Gerais. AOA, AMC, and CMA are fellows of CNPq, Brazil (302942/2022-0, 174190/2023-9, and 153088/2024-9, respectively). The authors are grateful for these organizations that supported this study and the participants who participated in this research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical restrictions preventing public sharing of data but are available from the corresponding author on reasonable request. However, we provide an anonymized dataset deposited on Zenodo, which contains relevant information about the participants, including age, gender, clinical characteristics, and the in-game estimated variables used in this study, to assess slowness (response time and angular velocity) [50]. Additional materials are also freely accessible: the source code of the RehaBEElitation serious game (available on Zenodo [51]), the executable file of the RehaBEElitation serious game (available on Zenodo [52]), the introductory videos and videos showing the functioning of each phase of the RehaBEElitation serious game (available on Zenodo [53]), and anonymized videos demonstrating the gameplay interaction of a participant with Parkinson disease (available on Zenodo [54]).

Authors' Contributions

LCM and AOA were responsible for the conceptualization of the study; LCM, CMA, and IAM contributed to the development of the game and human-machine interface device and collected the data; LCM and AMC estimated the participants' response time and angular velocity, performed the statistical tests, developed the figures, and wrote the first version of the manuscript; CMA, IAM, AMC, ALD, YM, and AOA revised the manuscript; and AOA and YM supervised the team and managed the project. All authors read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Phases of the game. (A) Phase 1: pollinating the flowers. The image shows that the player has collected pollen and had his or her hand closed, as indicated by the yellow grains below the bee. (B) Phase 2: feeding the larvae. When the player opens his or her hand to deliver food to the larvae and feed them, the light arcs above the larvae disappear. (C) Phase 3: collecting the nectar. As the player approaches the water drops, it is necessary to make a pinching gesture to collect it. (D) Phase 4: drying the nectar. When the honeycombs containing nectar are dried (after pronation and supination movement execution), the light reflex disappears. [\[PNG File, 1248 KB - games_v14i1e79463_app1.png\]](#)

Multimedia Appendix 2

Wearable human-machine interface based on a glove in 2 different perspectives. A compact casing positioned on the back of the hand houses the electronic circuitry and related components responsible for transmitting motion data to the game. [\[PNG File, 345 KB - games_v14i1e79463_app2.png\]](#)

Multimedia Appendix 3

Schematic of the experimental scenario (lateral view). The participant was positioned comfortably in a chair to facilitate the best possible interaction with the game. The researcher intervened when necessary, fostering a more human-centered experimental session. [\[PNG File, 744 KB - games_v14i1e79463_app3.png\]](#)

Multimedia Appendix 4

Response time definition. It represents the total response duration, encompassing both the reaction time (from stimulus onset to movement initiation) and the movement time (from movement initiation to task completion). Lower values indicate faster performance. [\[PNG File, 391 KB - games_v14i1e79463_app4.png\]](#)

Multimedia Appendix 5

Sensor axis orientation. Movements on the other axes were used only to steer the bee in the scenario (not to score points) and could be intentionally slowed down or minimized; therefore, angular velocity analyses were restricted to phase 4. [\[PNG File, 326 KB - games_v14i1e79463_app5.png\]](#)

Multimedia Appendix 6

Overview of the singular spectrum analysis decomposition pipeline, covering 4 steps: embedding, decomposition, grouping, and reconstruction.

[PNG File, 54 KB - [games_v14i1e79463_app6.png](#)]

Multimedia Appendix 7

Decomposition of a gyroscope signal from an experimental group participant. The top signal represents the original time series, containing both components of voluntary movement (from gameplay) and involuntary movement (tremor). The middle signal isolates the voluntary movement, that is, the intentional control of the bee. The bottom signal isolates the disease-related tremor, that is, only the oscillatory or periodic components of the original time series.

[PNG File, 131 KB - [games_v14i1e79463_app7.png](#)]

References

- 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;17:3. [doi: [10.1186/s11556-019-0233-2](#)] [Medline: [31938075](#)]
- Wilkinson A, Tong T, Zare A, Kanik M, Chignell M. Monitoring health status in long term care through the use of ambient technologies and serious games. *IEEE J Biomed Health Inform* 2018 Nov;22(6):1807-1813. [doi: [10.1109/JBHI.2018.2864686](#)] [Medline: [30106702](#)]
- Vázquez-Vélez GE, Zoghbi HY. Parkinson's disease genetics and pathophysiology. *Annu Rev Neurosci* 2021 Jul 8;44:87-108. [doi: [10.1146/annurev-neuro-100720-034518](#)] [Medline: [34236893](#)]
- Palacios-Sánchez L, Nupan MT, Botero-Meneses JS. James Parkinson y su ensayo sobre la “parálisis agitante”, doscientos años después [James Parkinson and his essay on “shaking palsy,” two hundred years later]. *Arq Neuropsiquiatr* 2017;75:671-672. [doi: [10.1590/0004-282x20170108](#)]
- Bologna M, Paparella G, Fasano A, Hallett M, Berardelli A. Evolving concepts on bradykinesia. *Brain (Bacau)* 2020 Mar 1;143(3):727-750. [doi: [10.1093/brain/awz344](#)] [Medline: [31834375](#)]
- Berardelli A, Rothwell JC, Thompson PD, Hallett M. Pathophysiology of bradykinesia in Parkinson's disease. *Brain (Bacau)* 2001 Nov;124(Pt 11):2131-2146. [doi: [10.1093/brain/124.11.2131](#)] [Medline: [11673316](#)]
- Evarts EV, Teräväinen H, Calne DB. Reaction time in Parkinson's disease. *Brain (Bacau)* 1981 Mar;104(Pt 1):167-186. [doi: [10.1093/brain/104.1.167](#)] [Medline: [7470841](#)]
- Schmidt RA, Lee TD. Motor learning and performance: from principles to application. *Clin Geriatr Med* 2020;36:1-12. [doi: [10.1016/j.cger.2019.08.002](#)]
- Goetz CG, Tilley BC, Shaftman SR, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008 Nov 15;23(15):2129-2170. [doi: [10.1002/mds.22340](#)] [Medline: [19025984](#)]
- Lee WL, Sinclair NC, Jones M, et al. Objective evaluation of bradykinesia in Parkinson's disease using an inexpensive marker-less motion tracking system. *Physiol Meas* 2019 Feb 1;40(1):014004. [doi: [10.1088/1361-6579/aafef2](#)] [Medline: [30650391](#)]
- Luiz LMD, Marques IA, Folador JP, Andrade AO. Intra and inter-rater remote assessment of bradykinesia in Parkinson's disease. *Neurologia (Engl Ed)* 2024 May;39(4):345-352. [doi: [10.1016/j.nrleng.2021.08.006](#)] [Medline: [38616062](#)]
- Smid A, Elting JWJ, van Dijk JMC, et al. Intraoperative quantification of MDS-UPDRS tremor measurements using 3D accelerometry: a pilot study. *J Clin Med* 2022 Apr 19;11(9):2275. [doi: [10.3390/jcm11092275](#)] [Medline: [35566401](#)]
- Bremm RP, Pavelka L, Garcia MM, et al. Sensor-based quantification of MDS-UPDRS III subitems in Parkinson's disease using machine learning. *Sensors (Basel)* 2024 Mar 29;24(7):2195. [doi: [10.3390/s24072195](#)] [Medline: [38610406](#)]
- Hong G, Mao F, Zhang M, et al. Modeling and validation of wearable sensor-based gait parameters in Parkinson's disease patients with cognitive impairment. *Front Aging Neurosci* 2025;17:1590224. [doi: [10.3389/fnagi.2025.1590224](#)] [Medline: [40787670](#)]
- Sica M, Tedesco S, Crowe C, et al. Continuous home monitoring of Parkinson's disease using inertial sensors: a systematic review. *PLoS ONE* 2021;16(2):e0246528. [doi: [10.1371/journal.pone.0246528](#)] [Medline: [33539481](#)]
- Bhidayasiri R, Tarsy D. Parkinson's disease: “On-off” phenomenon. *Curr Clin Neurol* 2012;36:14-15. [doi: [10.1007/978-1-60327-426-5_7](#)]
- Mendes LC, Sá AAR, Marques IA, et al. RehaBEElitation: the architecture and organization of a serious game to evaluate motor signs in Parkinson's disease. *PeerJ Comput Sci* 2023;9:e1267. [doi: [10.7717/peerj-cs.1267](#)] [Medline: [37346638](#)]
- Mendes LC, Marques IA, Alves CM, et al. Évaluation de l'utilisabilité d'un jeu sérieux pour la rééducation des personnes atteintes de la maladie de parkinson. 2022 Presented at: Handicap 2022—Humaines et Artificielles, les Intelligences au Service du Handicap p. 40-45 URL: https://ifraith.fr/handicap2022/wp-content/uploads/sites/13/2023/11/proceedings_handicap2022_usb.pdf [accessed 2026-02-23]
- Mendes LC, Moura Cabral A, Alves CM, Morere Y, Marques IA, de Oliveira Andrade A. Objective evaluation of bradykinesia using the RehaBEElitation serious game. Presented at: 2022 IEEE Workshop on Complexity in Engineering (COMPENG); Jul 18-20, 2022; Florence, Italy p. 65-66. [doi: [10.1109/COMPENG50184.2022.9905437](#)]

20. de Oliveira LC, Mendes LC, de Lopes RA, et al. A systematic review of serious games used for rehabilitation of individuals with Parkinson's disease. *Res Biomed Eng* 2021 Dec;37(4):849-865. [doi: [10.1007/s42600-021-00162-3](https://doi.org/10.1007/s42600-021-00162-3)]
21. Oña Simbaña ED, Cuesta-Gomez A, Garcia JA, et al. Evaluating a VR-based box and blocks test for automatic assessment of manual dexterity: a preliminary study in Parkinson's disease. Presented at: 2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH); Aug 5-7, 2019. [doi: [10.1109/SeGAH.2019.8882472](https://doi.org/10.1109/SeGAH.2019.8882472)]
22. Sánchez-Herrera-Baeza P, Cano-de-la-Cuerda R, Oña-Simbaña ED, et al. The impact of a novel immersive virtual reality technology associated with serious games in Parkinson's disease patients on upper limb rehabilitation: a mixed methods intervention study. *Sensors (Basel)* 2020 Apr 11;20(8):2168. [doi: [10.3390/s20082168](https://doi.org/10.3390/s20082168)] [Medline: [32290517](https://pubmed.ncbi.nlm.nih.gov/32290517/)]
23. Fernández-González P, Carratalá-Tejada M, Monge-Pereira E, et al. Leap motion controlled video game-based therapy for upper limb rehabilitation in patients with Parkinson's disease: a feasibility study. *J NeuroEngineering Rehabil* 2019 Dec;16(1). [doi: [10.1186/s12984-019-0593-x](https://doi.org/10.1186/s12984-019-0593-x)]
24. Cai H, Lin T, Chen L, et al. Evaluating the effect of immersive virtual reality technology on gait rehabilitation in stroke patients: a study protocol for a randomized controlled trial. *Trials* 2021 Jan 25;22(1):91. [doi: [10.1186/s13063-021-05031-z](https://doi.org/10.1186/s13063-021-05031-z)] [Medline: [33494805](https://pubmed.ncbi.nlm.nih.gov/33494805/)]
25. Yuan RY, Chen SC, Peng CW, Lin YN, Chang YT, Lai CH. Effects of interactive video-game-based exercise on balance in older adults with mild-to-moderate Parkinson's disease. *J Neuroeng Rehabil* 2020 Jul 13;17(1):91. [doi: [10.1186/s12984-020-00725-y](https://doi.org/10.1186/s12984-020-00725-y)] [Medline: [32660512](https://pubmed.ncbi.nlm.nih.gov/32660512/)]
26. Robert P, Manera V, Derreumaux A, et al. Efficacy of a web app for cognitive training (MEMO) regarding cognitive and behavioral performance in people with neurocognitive disorders: randomized controlled trial. *J Med Internet Res* 2020 Mar 11;22(3):e17167. [doi: [10.2196/17167](https://doi.org/10.2196/17167)] [Medline: [32159519](https://pubmed.ncbi.nlm.nih.gov/32159519/)]
27. Jäggi S, Wachter A, Adcock M, et al. Feasibility and effects of cognitive-motor exergames on fall risk factors in typical and atypical Parkinson's inpatients: a randomized controlled pilot study. *Eur J Med Res* 2023 Jan 16;28(1):30. [doi: [10.1186/s40001-022-00963-x](https://doi.org/10.1186/s40001-022-00963-x)] [Medline: [36647177](https://pubmed.ncbi.nlm.nih.gov/36647177/)]
28. Bégel V, Puyjarinet F, Gény C, et al. Fine motor serious game training improves gait in Parkinson's disease: a pilot randomized controlled trial. *bioRxiv*. Preprint posted online on Sep 17, 2025. [doi: [10.1101/2025.09.16.676299](https://doi.org/10.1101/2025.09.16.676299)]
29. Ganzeboom M, Bakker M, Beijer L, Strik H, Rietveld T. A serious game for speech training in dysarthric speakers with Parkinson's disease: exploring therapeutic efficacy and patient satisfaction. *Int J Lang Commun Disord* 2022 Jul;57(4):808-821. [doi: [10.1111/1460-6984.12722](https://doi.org/10.1111/1460-6984.12722)] [Medline: [35338751](https://pubmed.ncbi.nlm.nih.gov/35338751/)]
30. Preedy VR, Watson RR. *Handbook of Disease Burdens and Quality of Life Measures*: Springer; 2009. [doi: [10.1007/978-0-387-78665-0](https://doi.org/10.1007/978-0-387-78665-0)]
31. 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](https://doi.org/10.1016/0022-3956(75)90026-6)] [Medline: [1202204](https://pubmed.ncbi.nlm.nih.gov/1202204/)]
32. Pasciuto I, Ligorio G, Bergamini E, Vannozzi G, Sabatini AM, Capozzo A. How angular velocity features and different gyroscope noise types interact and determine orientation estimation accuracy. *Sensors (Basel)* 2015 Sep 18;15(9):23983-24001. [doi: [10.3390/s150923983](https://doi.org/10.3390/s150923983)] [Medline: [26393606](https://pubmed.ncbi.nlm.nih.gov/26393606/)]
33. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edition: Routledge; 2013. [doi: [10.4324/9780203771587](https://doi.org/10.4324/9780203771587)]
34. de Oliveira Andrade A, Paixão APS, Cabral AM, et al. Task-specific tremor quantification in a clinical setting for Parkinson's disease. *J Med Biol Eng* 2020 Dec;40(6):821-850. [doi: [10.1007/s40846-020-00576-x](https://doi.org/10.1007/s40846-020-00576-x)]
35. Golyandina N, Korobeynikov A, Zhigljavsky A. *Singular Spectrum Analysis with R*: Springer; 2018. [doi: [10.1007/978-3-662-57380-8](https://doi.org/10.1007/978-3-662-57380-8)]
36. Hassani H. Singular spectrum analysis: methodology and comparison. *J Data Sci* 2007;5(2):239-257. [doi: [10.6339/JDS.2007.05\(2\).396](https://doi.org/10.6339/JDS.2007.05(2).396)]
37. Bonizzi P, Karel JMH, Meste O, Peeters RLM. Singular spectrum decomposition: a new method for time series decomposition. *Adv Adapt Data Anal* 2014 Oct;06(4):1450011. [doi: [10.1142/S1793536914500113](https://doi.org/10.1142/S1793536914500113)]
38. Regnault A, Borojerdi B, Meunier J, Bani M, Morel T, Cano S. Does the MDS-UPDRS provide the precision to assess progression in early Parkinson's disease? Learnings from the Parkinson's progression marker initiative cohort. *J Neurol* 2019 Aug;266(8):1927-1936. [doi: [10.1007/s00415-019-09348-3](https://doi.org/10.1007/s00415-019-09348-3)] [Medline: [31073716](https://pubmed.ncbi.nlm.nih.gov/31073716/)]
39. Evers LJW, Krijthe JH, Meinders MJ, Bloem BR, Heskes TM. Measuring Parkinson's disease over time: the real-world within-subject reliability of the MDS-UPDRS. *Mov Disord* 2019 Oct;34(10):1480-1487. [doi: [10.1002/mds.27790](https://doi.org/10.1002/mds.27790)] [Medline: [31291488](https://pubmed.ncbi.nlm.nih.gov/31291488/)]
40. di Biase L, Summa S, Tosi J, et al. Quantitative analysis of bradykinesia and rigidity in Parkinson's disease. *Front Neurol* 2018;9:121. [doi: [10.3389/fneur.2018.00121](https://doi.org/10.3389/fneur.2018.00121)] [Medline: [29568281](https://pubmed.ncbi.nlm.nih.gov/29568281/)]
41. Juanjuan H, zhiming Y, Jianguo W, Bochen L, Xianjun Y. An automatic detection method for bradykinesia in Parkinson's disease based on inertial sensor. Presented at: 2020 IEEE 3rd International Conference on Electronics Technology (ICET); May 8-12, 2020; Chengdu, China p. 166-169. [doi: [10.1109/ICET49382.2020.9119604](https://doi.org/10.1109/ICET49382.2020.9119604)]
42. Mandic DP, Rehman NU, Wu Z, Huang NE. Empirical mode decomposition-based time-frequency analysis of multivariate signals: the power of adaptive data analysis. *IEEE Signal Process Mag* 2013;30(6):74-86. [doi: [10.1109/MSP.2013.2267931](https://doi.org/10.1109/MSP.2013.2267931)]

43. Pietrosanti L, Patera M, Suppa A, et al. Relevance of harmonic content findings of hand motor (dys)functionalities in Parkinson's disease revealed by means of a sensory glove. *Biocybern Biomed Eng* 2025 Jul;45(3):507-514. [doi: [10.1016/j.bbe.2025.07.004](https://doi.org/10.1016/j.bbe.2025.07.004)]
44. Oubre B, Daneault JF, Whritenour K, et al. Decomposition of reaching movements enables detection and measurement of ataxia. *Cerebellum* 2021 Dec;20(6):811-822. [doi: [10.1007/s12311-021-01247-6](https://doi.org/10.1007/s12311-021-01247-6)] [Medline: [33651372](https://pubmed.ncbi.nlm.nih.gov/33651372/)]
45. Dhall R, Kreitzman DL. Advances in levodopa therapy for Parkinson disease: review of RYTARY (carbidopa and levodopa) clinical efficacy and safety. *Neurology (ECronicon)* 2016 Apr 5;86(14 Suppl 1):S13-S24. [doi: [10.1212/WNL.0000000000002510](https://doi.org/10.1212/WNL.0000000000002510)] [Medline: [27044646](https://pubmed.ncbi.nlm.nih.gov/27044646/)]
46. Olanow CW, Stern MB, Sethi K. The scientific and clinical basis for the treatment of Parkinson disease (2009). *Neurology (ECronicon)* 2009 May 26;72(21 Suppl 4):S1-136. [doi: [10.1212/WNL.0b013e3181a1d44c](https://doi.org/10.1212/WNL.0b013e3181a1d44c)] [Medline: [19470958](https://pubmed.ncbi.nlm.nih.gov/19470958/)]
47. Summa S, Tosi J, Taffoni F, et al. Assessing bradykinesia in Parkinson's disease using gyroscope signals. *IEEE Int Conf Rehabil Robot* 2017 Jul;2017:1556-1561. [doi: [10.1109/ICORR.2017.8009469](https://doi.org/10.1109/ICORR.2017.8009469)] [Medline: [28814041](https://pubmed.ncbi.nlm.nih.gov/28814041/)]
48. Jogo desenvolvido na UFU contribui para o tratamento de pacientes com parkinson. *Comunica UFU*. 2023. URL: <https://comunica.ufu.br/noticias/2023/06/jogo-desenvolvido-na-ufu-contribui-para-o-tratamento-de-pacientes-com-parkinson> [accessed 2026-02-23]
49. Prêmio SBEB boston scientific de inovação em engenharia biomédica 2023: confirma os vencedores. Sociedade Brasileira de Engenharia Biomédica (SBEB). 2023. URL: <https://sbeb.org.br/premio-sbeb-boston-scientific-de-inovacao-em-engenharia-biomedica-2023-confirma-os-vencedores/> [accessed 2026-02-23]
50. Mendes LC, Alves CM, Andrade A. Dataset for slowness assessment during interactions with the RehaBEElitation serious game. *Zenodo*. 2025. URL: <https://zenodo.org/records/17860725> [accessed 2026-02-24]
51. Mendes LC, Marques IA, Sá AAR, et al. RehaBEElitation—source code. *Zenodo*. 2022. URL: <https://zenodo.org/records/6985977> [accessed 2026-02-24]
52. Mendes LC, Marques IA, Sá AAR, et al. RehaBEElitation serious game. *Zenodo*. 2022. URL: <https://zenodo.org/records/6988572> [accessed 2026-02-24]
53. Mendes LC, Sá AAR, Marques IA, et al. Introductory videos and videos showing the functioning of each phase of the RehaBEElitation serious game. *Zenodo*. 2022. URL: <https://zenodo.org/records/6977456> [accessed 2026-02-24]
54. Mendes LC, Alves CM, Andrade A. Demonstration of gameplay interaction by a participant with Parkinson's disease with the RehaBEElitation serious game. *Zenodo*. 2025. URL: <https://zenodo.org/records/17859707> [accessed 2026-02-24]

Abbreviations

AV: angular velocity

CG: control group

EG: experimental group

HMI: human-machine interface

IMU: inertial measurement unit

MDS-UPDRS: Movement Disorder Society-Unified Parkinson's Disease Rating Scale

PD: Parkinson disease

RT: response time

SSA: singular spectrum analysis

SSD: singular spectrum decomposition

SVD: singular value decomposition

UFU: Federal University of Uberlândia

Edited by S Brini; submitted 23.Jun.2025; peer-reviewed by MD Mendonca, S Hill; accepted 05.Jan.2026; published 25.Feb.2026.

Please cite as:

Mendes LC, Cabral AM, Alves CM, Marques IA, Delis AL, Morère Y, Andrade ADO

Quantifying Slowness in Parkinson Disease Using a Serious Game: Cross-Sectional Study

JMIR Serious Games 2026;14:e79463

URL: <https://games.jmir.org/2026/1/e79463>

doi: [10.2196/79463](https://doi.org/10.2196/79463)

© Luanne Cardoso Mendes, Ariana Moura Cabral, Camille Marques Alves, Isabela Alves Marques, Alberto López Delis, Yann Morère, Adriano de Oliveira Andrade. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 25.Feb.2026. 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.

Co-Designing Mobile Serious Games to Support Patients With Psoriatic Arthritis and Chronic Pain: Mixed Methods Study

Bárbara Ramalho¹, MSc; Samuel Gomes¹, MSc; Marta Silva Vicente¹, MSc; Filipa Magalhães¹, MSc; Rodolfo Gonçalves Costa¹, MSc; Sandra Gama², PhD; Vasileios Charisis³, PhD; Leontios Hadjileontiadis^{3,4}, PhD; Sofia B Dias⁵, PhD; iPROLEPSIS Consortium³

¹HUMAN Lab, INESC-ID, Instituto Superior Técnico, Faculdade de Motricidade Humana, Universidade de Lisboa, Lisbon, Oeiras, Portugal

²HUMAN Lab, INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

³Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Thessaloniki, Central Macedonia, Greece

⁴Department of Biomedical Engineering and Biotechnology, Khalifa University, Abu Dhabi, United Arab Emirates

⁵Center of Interdisciplinary Study of Human Performance (CIPER), Faculdade de Motricidade Humana, Universidade de Lisboa, Estrada da Costa, 1495-751 Cruz Quebrada, Lisbon, Portugal

Corresponding Author:

Sofia B Dias, PhD

Center of Interdisciplinary Study of Human Performance (CIPER), Faculdade de Motricidade Humana, Universidade de Lisboa, Estrada da Costa, 1495-751 Cruz Quebrada, Lisbon, Portugal

Abstract

Background: Serious games offer promising avenues for clinical care by enhancing patient engagement and delivering therapeutic benefits. In psoriatic arthritis (PsA), chronic pain contributes to emotional distress, functional limitations, and reduced well-being. While symptom-tracking apps exist, few digital interventions directly address chronic pain through engaging, therapeutic experiences tailored to patients' cognitive and physical needs.

Objective: This study aimed to co-design mobile serious games—NoPain Games—to support patients with PsA in managing chronic pain. Conducted within the iPROLEPSIS Horizon Europe project, the study involved a multidisciplinary cocreation session with rheumatologists, researchers, and technical experts, followed by a usability feedback session with patients with PsA. The goal was to identify therapeutic priorities, refine game mechanics, and assess usability to inform the development of personalized, accessible digital interventions.

Methods: A sequential mixed methods design was used. First, a 90-minute remote cocreation session was held with 14 experts (6 rheumatologists, 4 technical experts, and 4 researchers) from 3 European countries. Participants reviewed game storyboards and discussed therapeutic and design priorities. Thematic analysis of transcribed discussions identified key insights. Next, a 60-minute remote usability feedback session was conducted with 5 patients with PsA (aged 25 - 64 y), who interacted with 2 high-fidelity game prototypes (Space Oddity and Four Seasons). Usability was assessed using the System Usability Scale (SUS), and qualitative feedback was collected through moderated discussion. Item-level analysis using item characteristic curves provided deeper insight into usability perceptions and item sensitivity. All ethics requirements were met for both study phases.

Results: Thematic analysis of the transcribed dialogs of the cocreation session revealed three core themes: (1) therapeutic benefits (pain distraction, memory enhancement, cognitive stimulation, stress reduction, and creative engagement); (2) game difficulty (balancing duration and complexity to sustain engagement without fatigue); and (3) accessibility and interaction (addressing physical limitations, optimizing touchscreen usability, and ensuring inclusive design). These insights informed the development of 2 NoPain Game prototypes, which received a SUS score of 79 (SD 10.4; 95% CI 69.89 - 88.11), indicating good usability. Item characteristic curve analysis showed strong discrimination for learnability, while ease of use and confidence exhibited ceiling effects. Items like support needs and inconsistency showed minimal variability, and the learning curve demonstrated delayed but meaningful responsiveness at higher usability levels. Qualitative feedback reinforced the relevance of difficulty adjustment, technical refinements, and game mechanics, offering actionable insights for future iterations and broader implementation.

Conclusions: This study uniquely contributes to the field by co-designing mobile serious games specifically for patients with PsA, integrating expert and patient input to address chronic pain through accessible, cognitively engaging digital interventions—an approach not previously explored in this population. Future work will refine game mechanics, integrate adaptive difficulty, and conduct clinical trials to evaluate therapeutic efficacy. NoPain Games may serve as complementary tools within digital care ecosystems, offering support tailored to the therapeutic needs and physical limitations of patients with PsA underserved by conventional therapies.

KEYWORDS

serious games; psoriatic arthritis; chronic pain; NoPain Games; mHealth; digital health care; iPROLEPSIS; mobile health

Introduction

In health care, serious games—designed for purposes beyond entertainment—have generated considerable interest and sparked ongoing debate [1]. Their potential has often been overlooked due to the still prevailing perception of games as purely recreational objects rather than valuable tools for skill development and health rehabilitation [2]. Research studies, however, have demonstrated the benefits of serious games, particularly their ability to reduce stress [3], enhance mobility [4,5], and alleviate pain [6,7].

Many of these benefits align with the symptomatology of rheumatic diseases, making it possible to use serious games as health care tools for conditions such as psoriatic arthritis (PsA). PsA, a degenerative rheumatic disease, is marked by joint stiffness and persistent pain, significantly affecting patients' quality of life [8]. As there is no cure, treatment primarily focuses on managing symptoms and holistically improving patients' well-being, often through interventions like physical therapy [9]. However, these traditional approaches can feel repetitive or burdensome for patients, leading to low adherence and reduced effectiveness. While various digital tools, particularly mobile apps for disease management, have been proposed [10], they primarily serve as symptom-tracking solutions and fail to address the ongoing joint pain experienced by patients.

Serious games offer a compelling alternative to complement traditional treatments by creating experiences that are both engaging and therapeutic [11]. For example, DaktylAct, a touch-based serious game, has been proposed as an innovative tool for assessing fine motor skills in PsA using novel digital biomarkers [12]. Furthermore, immersive experiences in chronic pain management have been extensively studied, with strong evidence supporting their pain-relieving capabilities, particularly in virtual reality (VR) environments [6,13]. Building on this, previous research has highlighted the effectiveness of exergames in altering pain perception by engaging multiple sensory modalities [14]. For instance, Gold et al [13] highlighted how VR can influence neurobiological mechanisms, showing that auditory, visual, and tactile stimuli reduce activity in pain-processing areas. In addition, Hoffman et al [15] further validated the effectiveness of immersive digital interventions by comparing VR and opioids in pain relief. Their findings suggest that while both approaches are beneficial independently, their combination leads to even greater analgesic effects. These findings, derived from subjective pain scales and functional magnetic resonance imaging, highlight the potential of mobile-based alternatives with similar engagement mechanisms to perform effectively without the need for specialized hardware.

Moreover, recent studies have further expanded the evidence base for serious games in pain management. Beltran-Alacreu et al [16] developed a task-oriented serious game for older adults

with chronic neck pain, demonstrating its suitability and therapeutic potential. Saragih et al [7] conducted a systematic review and meta-analysis confirming the efficacy of serious games in managing chronic pain among older adults, reinforcing their clinical relevance. Additionally, Peña et al [17] explored the use of digital art and attachment priming in a web-based serious game, showing promising results in reducing both pain and social disconnection in individuals with chronic pain and loneliness.

Immersion remains a critical factor in pain management games. Gromala et al [6] emphasized that realistic visual environments, sound effects, and interactive storytelling contribute significantly to the overall therapeutic experience. Their concept of imaginary immersion introduces in-game threats and challenges to sustain player focus and engagement. We believe that translating these principles into mobile platforms requires innovative game mechanics that maintain immersion despite limited sensory channels. For instance, the MyRelief smartphone app presents a serious game to relieve chronic low-back pain through physical and psychological activities [18]. Although their analysis suggests that serious games can help improve patient status, the authors recognize that further research is needed to validate all game components.

Incorporating motivational elements in mobile pain management games is another critical design consideration. In this line, Ijaz et al [19] demonstrated the power of competitive motivation in their cycling game, which used on-screen scores of real and artificial competitors to enhance user engagement. Similarly, Tuah et al [20] reviewed common gamification elements in rehabilitation games, identifying key components such as points, leaderboards, badges, progression systems, and avatars that contribute to sustained participation.

To the best of our knowledge, no serious game has been specifically designed to target the chronic pain experienced by individuals with PsA. This study aims to explore the co-design and development of mobile serious games, referred to as NoPain Games, specifically tailored to support patients with PsA in managing chronic pain. Developed using an agile co-design methodology [21], these serious games represent a new approach to chronic pain management, as part of the Horizon Europe iPROLEPSIS project [22,23], which aims to create a personalized suite of games tailored to the needs of individuals with PsA. The primary objective of this study is to assess the therapeutic potential, usability, and design requirements of these games through a multidisciplinary cocreation process involving clinicians, researchers, and technical experts, along with the patients' usability feedback. By integrating cognitive stimulation, stress reduction, and inclusive interaction mechanics, the study seeks to establish foundational design principles for mobile-based digital interventions that enhance patient engagement and well-being. The underlying hypothesis is that co-designed serious games incorporating adaptive difficulty and accessible interfaces can serve as effective pain

distractors and cognitive enhancers for individuals living with PsA.

Methods

Overview

In this study, we followed a sequential mixed methods design, integrating qualitative and quantitative data to inform the co-design and evaluation of mobile serious games for patients with PsA. This sequential design was selected to allow expert-derived qualitative insights to inform prototype development prior to patient usability testing. Integration occurred through iterative refinement of game mechanics based on themes from the cocreation session, followed by triangulation of usability scores and patient feedback to validate design decisions. In the qualitative phase, 14 experts (6 rheumatologists, 4 technical experts, 4 researchers) were recruited via purposive sampling from the iPROLEPSIS Consortium and participated in a 90-minute remote cocreation session. Eligibility required domain expertise in rheumatology, digital health, or game development. Participants reviewed game storyboards and contributed to design refinement through moderated discussion. In the quantitative phase, 5 patients with PsA (age 25 - 64 years) from 4 countries were recruited through targeted outreach. Eligibility required a confirmed PsA diagnosis and the ability to interact with mobile devices. Patients participated in a 60-minute remote usability session, interacting with 2 high-fidelity game prototypes. Usability was assessed using the System Usability Scale (SUS), and qualitative feedback was collected through a structured discussion. Study outcomes included: (1) identification of therapeutic priorities and design requirements (qualitative), (2) usability scores and item-level sensitivity analysis (quantitative), and (3) patient feedback on game experience and accessibility (qualitative). Thematic analysis followed Braun and Clarke's framework, and SUS data were analyzed using descriptive statistics and item characteristic curves (ICCs). The study received ethical approval, and informed consent was secured from all participants.

The following subsections provide a detailed account of the study context, co-design procedures, participant characteristics, data collection instruments, ethics, and analytic strategies used to generate and interpret the findings.

Study Context

This study is part of the iPROLEPSIS Horizon Europe research project, which seeks to investigate the progression from general health to PsA through multi-source data analysis, ultimately

creating an innovative, personalized digital care ecosystem. Within the framework of iPROLEPSIS, the project focuses on designing, developing, and validating cutting-edge digital biomarkers to assess and address PsA using an Artificial Intelligence-Personalized Game Suite (AI-PGS). Developed collaboratively with key stakeholders, the AI-PGS takes a comprehensive, multitargeted approach to managing PsA symptoms. It includes intervention activities aimed at enhancing breathing, mobility, stiffness, balance, coordination, fitness, diet, and mood. By embracing a holistic perspective, the AI-PGS addresses stress, anxiety, fatigue, and pain through various categories of serious games, such as NoPain Games, Exercise Games [24], Sensorimotor Art Games [25], Breathing Games [26], Dietary Games, and Emotional Games. This study emphasizes NoPain Games, showcasing the critical role of engaging patients in therapeutic activities that incorporate cognitive processes—such as memory, coordination, and visual perception tasks—to help alleviate and manage chronic pain and discomfort associated with PsA symptoms.

NoPain Games Agile Co-Design Process

The design and development process for the proposed games followed an agile approach [21], emphasizing early feedback to iteratively refine game concepts and prototypes. The ideation and storyboards phase will guide the foundation for subsequent stages, including prototyping, software development, and expert evaluation, while ensuring continuous cocreation efforts aimed at clinical validation (see Figure 1). Following the product backlog, sprint planning meetings, and backlog refinement phases, the design process began with the Crazy8s ideation approach [27]. More specifically, stakeholders—including patients, clinicians, researchers, and technicians—collaboratively sketched ideas for various game categories within the AI-PGS framework, including NoPain Games (see Figure 2). This phase captured diverse requirements and expectations, culminating in the initial development of a game design document outlining the NoPain Games' key aspects, such as their concepts, mechanics, design, controls, and clinical relevance. This study highlights the subsequent agile cocreation sessions that brought together health care professionals, researchers, and technical experts as active cocreators and co-designers (see "Study participants and data collection" section) and patients with PsA as usability feedback providers (see "Feedback From Patients With PsA" section). In particular, using the storyboard technique [28], a collaborative session focused on refining and assessing the proposed designs for the NoPain Games (Figure 3), while another later session focused on the usability and general feedback of functional prototypes (Figure 4).

Figure 1. Schematic representation of the development framework for the proposed NoPain Games, illustrating the cocreation and agile methodologies used to achieve a minimum viable product. The process includes iterative 2 - 3 month sprints involving product backlog refinement (led by a technical coordinator), sprint planning, ideation, game development, expert evaluation, and clinical validation. All phases, except product backlog refinement, actively engaged patients, clinicians, researchers, and developers to ensure a multidisciplinary and user-centered approach.

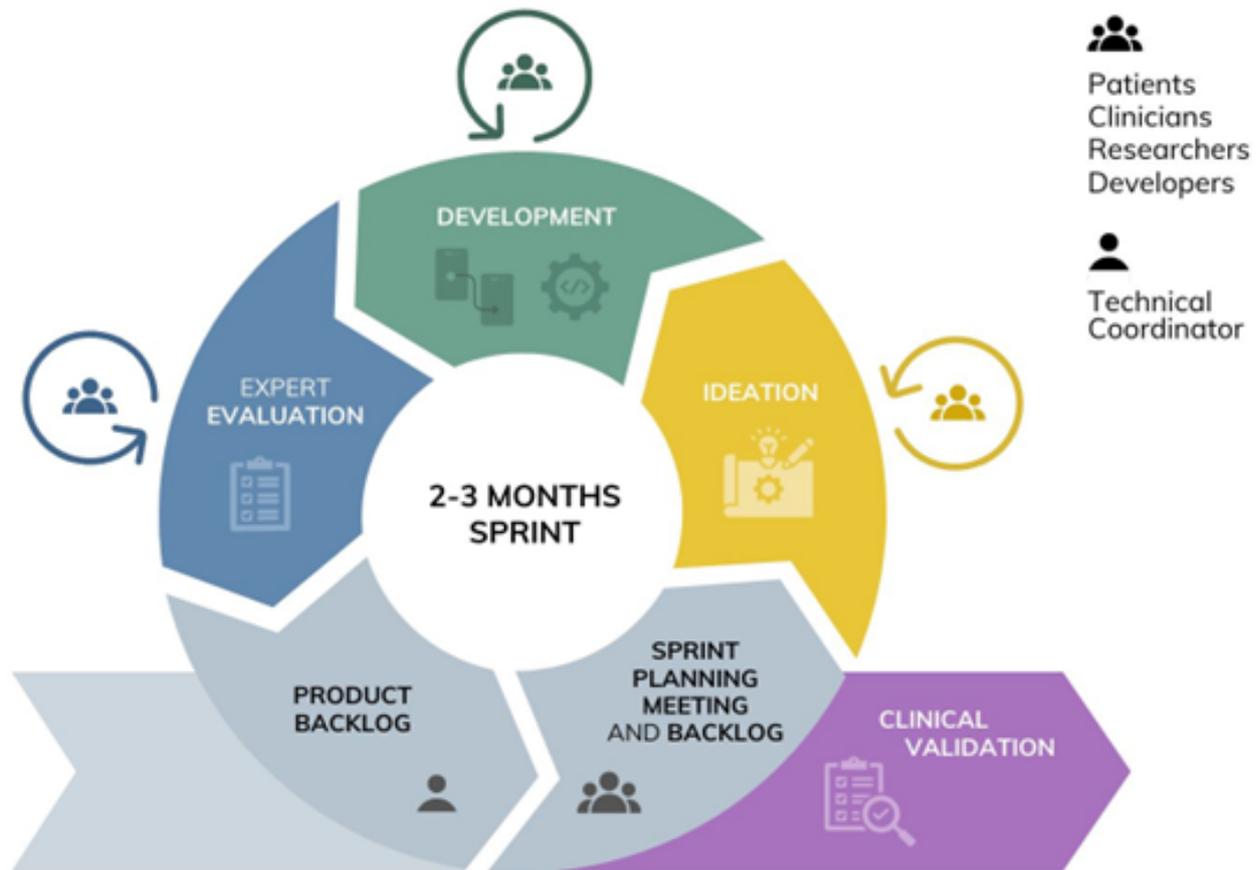


Figure 2. Co-design session involving patients with psoriatic arthritis during the ideation process of the games using Crazy8s method. The initial sketch of the NoPain Games category is marked with a dashed blue square.



Initially Proposed NoPain Games Storyboards

Three different NoPain Games were initially designed: “Lightning,” “Flashing Stars,” and “The Garden” with related storyboards (Figure 3). In the “Lightning” game, players/patients are required to replicate a given pattern by lighting up rooms in a 2D house, reinforcing memory and coordination (Figure 3A). The “Flashing Stars” game engages players in counting

falling stars within a night-sky scenario, fostering relaxation through gentle visual and cognitive stimulation (Figure 3B). Finally, “The Garden” game involves identifying a specific flower type within a vast field of blooms, promoting a sense of calm and concentration (Figure 3C). Overall, by incorporating the repetitive and rhythmic interactions of identifying flowers, lighting up rooms, or counting falling stars, the proposed NoPain Games encourage users to engage in predictable, structured

movements that promote relaxation and a sense of control. These are simple and rewarding mechanics aimed to provide immediate feedback, reinforcing positive engagement, and promoting a sense of accomplishment, while enabling a calm environment. Designed for both iOS and Android operating systems, the proposed NoPain Games high-fidelity prototypes present similar interaction controls, with players/patients using a touchscreen for interaction, with calming audio to match each scenario.

Study Participants

Cocreation Session

To obtain feedback on the related storyboards, a cocreation session was held. The latter occurred in October 2023 and was conducted digitally via the Microsoft Teams platform with shared storyboard visuals. The session included 14 experts from 3 European countries, namely the United Kingdom, Greece, and Portugal. Participants were selected using purposive sampling [29], intentionally inviting individuals with expertise in relevant fields. They were identified through institutional affiliations within the iPROLEPSIS Consortium and invited via email by the coordinating team. Participation was based on voluntary engagement and without the provision of compensation. Exclusion criteria for experts included a lack of direct experience with PsA or digital health interventions. The group included 6 rheumatologists, 4 researchers, and 4 technical experts from the iPROLEPSIS Consortium, with multidisciplinary perspectives, enriching the discussions. The cocreation session was moderated by the last author, a qualitative researcher with expertise in participatory design. To minimize bias, the moderator did not participate in game development. Two qualitative coders (first and last author) independently reviewed transcripts and discussed potential biases during theme development (see Results).

Feedback Session

To maintain a patient-centered and clinically relevant design process, individuals living with PsA were actively engaged in

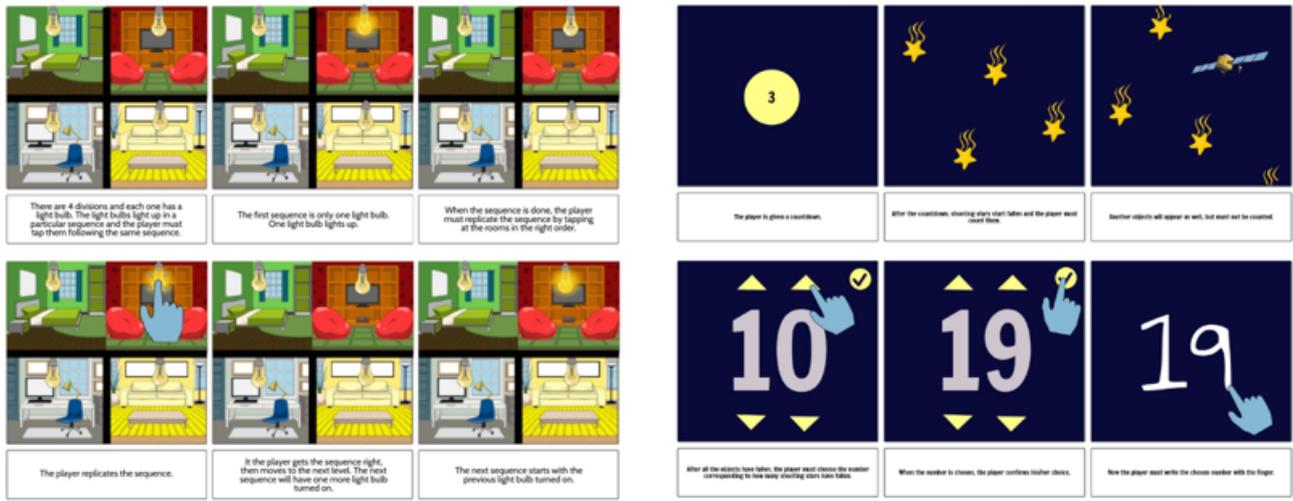
a dedicated feedback session conducted in May 2025. During the session, participants evaluated a pilot version of Space Oddity and a design prototype of Four Seasons (Figure 4). A 60-minute remote feedback session was conducted via Zoom with 5 patients with PsA from Portugal, the United Kingdom, the Netherlands, and Greece, using screen sharing for gameplay and an embedded SUS questionnaire via Google Forms. Similar to the cocreation session, participants were recruited through targeted outreach within the iPROLEPSIS network, through clinician referrals and patient advocacy networks affiliated with the project, using email and phone outreach. Participation was based on voluntary engagement and without the provision of compensation. For patients with PsA, exclusion criteria included cognitive impairments, severe visual/motor limitations preventing touchscreen interaction, or the inability to participate in digital sessions. The feedback session was also moderated by the last author, who did not participate in the quantitative analysis to minimize bias in the findings.

Data Collection

Cocreation Session

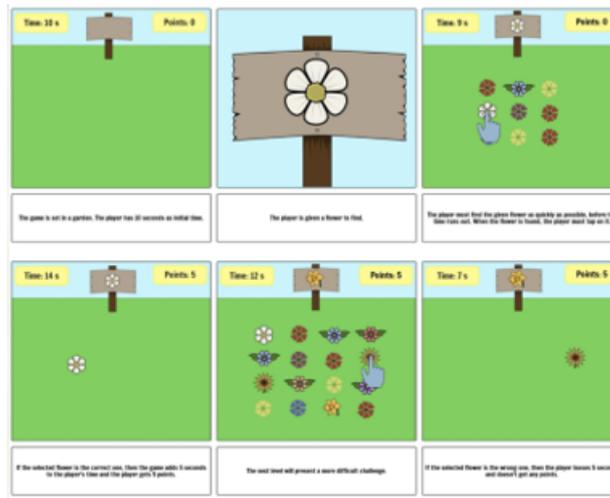
Audio recordings of the cocreation session were transcribed verbatim and then pseudoanonymized to preserve participant confidentiality. The digital setting facilitated active engagement from all attendees. The session was led by a female researcher specializing in qualitative research (last author), who guided discussions by posing questions and fostering group interactions to gather participants' opinions and experiences. To enhance participants' understanding of the session topics, relevant storyboards for each NoPain game were presented (Figure 3). The session followed a semistructured script with key questions designed to explore participants' views on the design of the proposed NoPain Games (Multimedia Appendix 1). A 90-minute timeframe was allowed for in-depth discussions and thorough exploration of the main topics.

Figure 3. Screenshot of the storyboards for the proposed NoPain games: (A) The “Lightning” game illustrates a memory challenge where players light up rooms to replicate a specific pattern; (B) The “Flashing Stars” game presents a relaxation activity focused on counting falling stars; and (C) “The Garden” game features a calming experience in which players identify specific flowers in a vibrant field.



(A)

(B)



(C)

Figure 4. Screenshot of the prototypes used in the usability evaluation, featuring (A) the “Space Oddity” rockets minigame and (B) “The Four Seasons” design prototype.



(A)



(B)

Feedback Session

At the feedback session, after the playing experience and discussion, participants completed a System Usability Score (SUS) questionnaire [30,31], and verbally provided their feedback on refinements for improving game experience and accessibility. SUS comprises 10 statements, each evaluated on a 5-point Likert scale ranging from “Strongly Disagree” to “Strongly Agree.” It serves as a standardized tool for assessing user experience and pinpointing usability challenges. Widely adopted across academic and industry settings, the SUS has

been validated as a reliable benchmark for interface evaluation [32]. Final scores range from 0 to 100, with values above 68 and above 80.3 generally indicating acceptable and excellent usability, respectively [32].

Table 1 tabulates the Implementation Matrix of the study. In the latter, qualitative and quantitative data sources are systematically organized by type, timing of data collection, participant group, and associated research aims. The matrix also outlines anticipated outcomes for each data stream, facilitating transparency in methodological integration and supporting the interpretive logic of the mixed methods design.

Table . Implementation matrix: overview of qualitative and quantitative data sources, collection timelines, participants, and study aim.

Data source	Type of data	Timing of collection	Participants	Study aim or research question	Anticipated outcome
Expert cocreation session	Qualitative	October 2023	14 experts (6 rheumatologists, 4 researchers, 4 technical experts)	What therapeutic priorities and design requirements should guide the development of NoPain Games?	Identification of core themes: therapeutic benefits, game difficulty, accessibility
Transcribed session dialogs	Qualitative	Postsession (Oct 2023)	Same as above	How do experts perceive the cognitive and physical needs of patients with PsA ^a in game design?	Thematic analysis yielding subthemes across cognitive and emotional engagement, gameplay adaptability, and inclusive interaction
Patient usability feedback session	Quantitative and qualitative	May 2025	5 patients with PsA (Portugal, the United Kingdom, Netherlands, Greece)	How usable and engaging are the NoPain Game prototypes from a patient perspective?	SUS ^b score, feedback on game mechanics, difficulty, and interaction design
SUS	Quantitative	May 2025	Same as above	What is the perceived usability of the NoPain Games?	Average SUS score >68
Patient feedback discussion	Qualitative	May 2025	Same as above	What refinements do patients suggest for improving game experience and accessibility?	Insights into difficulty adjustment, technical refinements, and inclusive interaction

^aPsA: psoriatic arthritis.

^bSUS: System Usability Scale.

Data Analysis

Qualitative Analysis

The qualitative data analysis process used an inductive thematic approach, with 2 independent researchers (first and last authors) analyzing all qualitative data. After thoroughly familiarizing themselves with the data, the researchers developed an initial set of codes, which were subsequently organized into themes and subthemes. These were carefully reviewed to ensure alignment with the data. The final themes and subthemes were crafted using clear and precise language, adhering to the methodology established by Braun and Clarke [33]. To enhance credibility and reliability, the themes underwent cross-validation, and agreement on code assignment was assessed using the Kappa statistic [34], with a threshold of 0.80 deemed acceptable [34]. The finalized themes were determined through consensus-building among the researchers. This approach provided a comprehensive understanding of the participants' perspectives, ensuring the analysis accurately captured their experiences and insights.

Quantitative Analysis

Quantitative data from the SUS questionnaire were analyzed with descriptive statistics, that is, mean, SD, median, IQR, showing how data spread out the central values, and 95% CIs quantifying the uncertainty in estimating the mean. Moreover, SUS 2D (average) and 3D (per patient with PsA) radar plots were estimated along with the SUS ICCs. The ICC is a foundational concept in item response theory, that is, a statistical framework used to model how individuals respond to test items

based on an underlying trait or ability [32]. An ICC is a curve that shows the probability $p(\theta)$ that a person with a given level of a latent trait θ (eg, usability perception, ability, attitude) will agree with or correctly answer a specific item. The probability $p(\theta)$ is given by a 2-Parameter Logistic (2PL) model [35] as

$$(1)p(\theta)=11+e^{-a(\theta-b)}$$

where a denotes the discrimination (slope), b the difficulty (threshold), and θ the latent trait value. A steep ICC (high a) means the item is highly sensitive to changes in usability perception, that is, ideal for detecting subtle shifts. A right-shifted ICC (high b) means the item is harder to agree with, that is, it activates only at higher usability levels. Finally, a flat ICC (low a) suggests the item does not discriminate well, that is, users across trait levels respond similarly. The whole quantitative analysis was carried out in Matlab 2025a (The MathWorks, Natick, USA).

Integration of Findings

Integration of qualitative and quantitative findings occurred through a sequential interpretive process that linked expert-derived design insights to patient-centered usability evaluation. Specifically, themes identified during the expert cocreation session—such as the need for cognitive stimulation, stress reduction, and adaptable difficulty—directly informed the development of 2 new game prototypes (Space Oddity and Four Seasons in the Results section). These prototypes operationalized therapeutic priorities and interaction mechanics discussed during the qualitative phase. In the subsequent feedback session, patients with PsA interacted with these games and provided both quantitative usability ratings via the SUS

and qualitative feedback through moderated discussion. The SUS scores offered a standardized measure of perceived usability, while the qualitative comments contextualized these scores by highlighting specific design strengths (eg, intuitive controls, calming visuals) and areas for refinement (eg, onboarding clarity, difficulty pacing). Item-level analysis using ICCs further revealed how individual SUS items aligned with patient-reported experiences, enabling cross-validation of usability dimensions such as learnability and confidence (see Results). This integration strategy ensured that the design decisions were not only expert-informed but also empirically validated through patient engagement, reinforcing the iterative and user-centered nature of the development process.

Ethical Considerations

This study involved 2 distinct data collection sessions conducted under the iPROLEPSIS Horizon Europe project. They involved members of the iPROLEPSIS Consortium participating in the 2 sessions, complying with the iPROLEPSIS-received ethical approval from the Ethics Committee of Erasmus Medical Center, Rotterdam, Netherlands (MEC-2023 - 0470). The first session was a cocreation workshop with members of the iPROLEPSIS Consortium, invited in their professional capacity. The second session was a usability feedback study involving patients with PsA. Neither session involved clinical procedures or the collection of sensitive personal health data. Participation in both sessions was entirely voluntary. For the cocreation session, verbal informed consent was obtained prior to participation. For the usability feedback session, informed consent was obtained via an embedded web-based form presented on the introductory page of the SUS questionnaire ([Multimedia Appendix 2](#)), which participants reviewed and accepted before proceeding. In both cases, participants were not compensated for their involvement. To ensure privacy and confidentiality across both phases, all audio recordings were transcribed verbatim and subsequently pseudo-anonymized. SUS survey responses were collected anonymously, with no identifying information linked to individual responses. All data were handled in accordance with General Data Protection Regulation-compliant data protection standards. No personally identifiable individuals appear in any images or supplementary materials.

Results

Qualitative Analysis

The cocreation session consisted of 14 participants, including 6 health care professionals, 4 technical experts, 3 observers, and one facilitator. The rheumatologists, all female, had a mean age of 36.3 (SD 7.2) years and over 10 years of professional experience, although most reported a basic level of technology literacy. The technical experts, all male, had a mean age of 44.8 (SD 7.9) years, with more than 15 years of experience and advanced-level technology literacy. The researchers, whose mean age was 28.5 (SD 7.0) years, included 3 females, each with over 4 years of experience and advanced-level technology literacy.

Thematic Analysis

Overview

Thematic analysis across all resolution scales, including themes and subthemes, revealed an almost perfect level of agreement between the 2 independent researchers (first and last authors), as indicated by a Kappa statistic of 0.90. In cases of uncertainty, discussions between researchers led to revisions, ultimately achieving 100% consensus across all subthemes. Data saturation was reached during the final stages of the cocreation session, as no new themes or subthemes emerged despite continued participant engagement. The structured yet open-ended nature of the sessions enabled iterative exploration of perspectives, and the consistency of responses across diverse stakeholders indicated that the thematic landscape had been sufficiently mapped. This saturation supports the robustness and completeness of the qualitative findings. From the thematic analysis, 3 primary themes emerged:

- **Therapeutic benefits (Theme 1):** This theme included subthemes such as pain distraction, memory enhancement, cognitive stimulation, stress reduction, and creative engagement, emphasizing the games' potential to help patients with PsA manage both mental and physical challenges.
- **Game difficulty (Theme 2):** This theme centered on subthemes like balancing gameplay duration, complexity, and adaptive difficulty levels to sustain engagement while avoiding fatigue or diminished therapeutic impact.
- **Accessibility and interaction (Theme 3):** This theme highlighted subthemes addressing physical limitations, optimizing touchscreen usability, and implementing inclusive design to meet the specific needs of patients with PsA.

The key findings for each identified subtheme are summarized below.

Theme 1: Therapeutic Benefits

Overview

The therapeutic value of NoPain Games was a recurring topic during the discussions, underlining their role in addressing both physical and mental health challenges. Overall, the NoPain Games were admired not only for their ability to distract patients from pain but also for enhancing cognitive functions like memory. This dual benefit positions them as a promising digital intervention, especially for patients with PsA, who often face fatigue and mental health issues.

Subtheme 1: Cognitive Stimulation, Stress Reduction, and Pain Distraction

Specifically, participants emphasized the dual benefits of NoPain Games in providing cognitive stimulation and serving as effective pain distractors. Engaging in memory-focused activities not only addresses the psychological burden of chronic illnesses but also shifts focus away from physical discomfort, ultimately promoting holistic well-being. In particular, the unique appeal of the "Lightning" NoPain Game exemplifies these combined benefits, as one clinician noted:

I'm just thinking the 'Lightning' game is really interesting in the way that besides distracting, it's also a memory game [...] because our patients are depressed, and this can be very helpful. [Rheumatologist #2]

This sentiment highlights the potential of such games to uplift patients' moods and combat depression, showing the therapeutic value of cognitive engagement. Some participants further reinforced this aspect by stating,

I like this game [the 'Lightning' game], it promotes memory skills, which can be seen as pain distractors. [Rheumatologist #3]

[...] the underlying mechanisms for this game [The Garden], is the memory, the mental stimulation in a way as part of the pain distractor. [Researcher #2]

By promoting cognitive stimulation, these serious games can also contribute indirectly to pain management, reducing the focus on discomfort. Additionally, many participants acknowledged these games as powerful tools for pain distraction and stress reduction. One participant summarized this aspect succinctly, as follows:

The NoPain Games are intended to act as a pain distractor, allowing the patient to focus on it as a means of distraction, which in turn can support pain management and stress reduction. [Researcher #1]

Overall, NoPain Games intends to integrate calming experiences and mentally stimulating activities, offering therapeutic benefits that extend beyond traditional treatments. By providing a distraction from chronic illness discomfort, these games shift patients' focus to engaging tasks that can alleviate stress, improve mood, and foster a sense of achievement. Their holistic approach addresses cognitive, emotional, and physical well-being that can enhance therapy outcomes and empower patients with PsA to actively participate in their care journey.

Subtheme 2: Promoting Creativity

This subtheme explores the integration of creativity into gameplay as a strategy for enhancing both physical and mental engagement in therapeutic contexts. Incorporating creative elements, such as drawing flowers in The Garden game, can introduce a unique combination of movement and creativity, directly supporting therapeutic objectives. As highlighted by one clinician:

Incorporating creative elements, such as drawing flowers, combines movement with creativity, stimulating the mind and physical tasks. Maybe the patient could draw the flower in this game [The Garden]. So you can combine in this way the movement and the creativity. [Rheumatologist #6]

Furthermore, another clinician proposed expanding this creative aspect, as follows:

[...] maybe there is a stage where somehow someone can use this kind of flowers to build something [...] more constructive. [Rheumatologist #5]

Incorporating creative tasks into gameplay can potentially enhance interaction and personalization, encouraging active

patient engagement in rehabilitation. Activities like drawing or constructing can stimulate fine motor skills, hand-eye coordination, cognitive focus, and emotional well-being. This creative emphasis can expand the NoPain Games' role, making them valuable tools that complement traditional treatments while combining therapeutic movement with self-expression for a more meaningful and enjoyable therapeutic experience.

Theme 2: Game Difficulty

Overview

Balancing the duration and complexity of gameplay emerged as a crucial aspect of game design during the cocreation session. Clinicians expressed concern that both extremes—prolonged sessions and overly simplistic mechanics—could undermine the games' therapeutic and engagement potential.

Subtheme 1: Managing Fatigue and Engagement

Prolonged gameplay or repetitive motions might exacerbate fatigue, reducing the games' effectiveness as therapeutic tools. One clinician pointed out the following:

For instance, it could be interesting to know how long the game sequences of the game [The 'Lightning' game] are, so how long should it last. [Rheumatologist #5]

This indicates the need for carefully tailored session lengths to prevent overstimulation. On the other hand, overly simple game mechanics might fail to captivate players or provide meaningful stimulation. In this line, one clinician also highlighted this issue, noting:

So, I'm just thinking about the complexity of the game [The 'Lightning' game], in particular the duration of the sequence and the number of repetitions to play the game. [...] have you decided or thoughts regarding the game duration? [Rheumatologist #1]

This perspective highlights the need for future work to focus on designing gameplay that achieves the optimal balance—sufficiently challenging to maintain engagement, yet not so demanding as to provoke frustration or exhaustion.

Subtheme 2: Customizable Difficulty Levels

This subtheme introduces customizable difficulty levels and adjustable session durations as effective solutions. These features enable players to adapt gameplay to their individual needs, thereby enhancing both satisfaction and therapeutic outcomes. Furthermore, the significance of gradual progression in difficulty is emphasized by one participant:

For each level of difficulty, different coordination movements can be incorporated to offer a light yet progressively challenging experience, possibly linked to timing. [Researcher #1]

This highlights how incorporating varied coordination movements fosters engagement while simultaneously supporting therapeutic goals, such as enhancing refined motor skills and cognitive reaction time. By offering challenges that escalate at an adaptable pace, players can feel a sense of growth and accomplishment, reinforcing motivation and adherence to therapeutic activities. Additionally, incorporating a time-based

element was suggested as a potential customization option. As noted by one technical expert:

Can I also ask if the game [The ‘Lightning’ game] is connected to time? Does it involve reaction time that progressively get faster? For example, if there’s a time limit, would it require completing the task within 5 seconds, then 10 seconds? [Technical expert #1]

Together, these insights can suggest that customization in difficulty and session dynamics is vital for creating a flexible and impactful game experience.

Theme 3: Accessibility and Interaction

Overview

Accessibility was identified as another pivotal theme, reflecting the need for an accessible design to accommodate the physical limitations of patients with PsA. Symptoms like joint stiffness, swelling, and reduced dexterity can significantly impact interaction with digital tools, including games.

Subtheme 1: Physical Limitations and Interface Challenges

Overall, many clinicians highlighted specific challenges related to physical symptoms. For example, one clinician commented:

I guess if people have swollen fingers, the accuracy for this game [‘The Garden’ game] might be a problem, in the sense that you’re trying to get quite a small area when you’ve got lots of flowers like this and the accuracy is potentially trickier. [Rheumatologist #6]

This underlines the need for larger touch targets and simplified game interactions to accommodate patients with swollen fingers or a limited range of motion. Similarly, another clinician noted:

Sometimes, on a small screen, it’s quite difficult to drag a marker to a particular point or to be very accurate with where we’re pressing if we’ve got bad hand arthritis. [Rheumatologist #2]

This observation points to the importance of designing user interfaces (UIs) that minimize the need for precise movements, making the games more accessible to those with severe joint issues.

Subtheme 2: Enhancing Inclusivity

To make the games more inclusive, developers could implement features such as adjustable touch sensitivity, alternative input methods (eg, voice commands), and options to magnify specific screen elements. These adaptations would enhance usability and ensure that the games can be enjoyed by a wider range of patients, regardless of their physical limitations.

These features foster a more inclusive experience for users with diverse needs, particularly those with conditions impairing their motor skills or visual acuity. For instance, adjustable touch sensitivity would allow users with reduced dexterity to interact with the interface more comfortably, while alternative input methods, such as voice commands, could provide a hands-free

option for individuals unable to perform precise touch gestures. Similarly, magnification options for specific screen game elements could support users with visual impairments, ensuring they can navigate the game environment with ease and confidence. As suggested by one participant:

In fact, the levels of difficulty and synchronization with the user profile, as suggested before using the adaptation algorithm of the games (...) includes the patient’s disease status, and eventually align with the hand or finger exhibiting better conditions in terms of symmetry and functionality. [Researcher #1]

This highlights the potential for an adaptive system that tailors the gameplay experience to the specific physical conditions of the player. By considering factors such as hand symmetry, finger functionality, or disease progression, the game could dynamically adjust its mechanics to align with each patient’s capabilities. For example, if a player has difficulty using one hand due to PsA, the game could prioritize interactions that favor the less-affected hand or allow for single-hand play. This ensures an engaging experience, minimizes frustration, and enhances therapeutic outcomes. Furthermore, a robust adaptation algorithm could learn and evolve with the player’s needs, offering more personalized solutions over time by focusing on inclusive design principles.

Summarizing, these detailed insights into each theme emphasize the multifaceted nature of NoPain Games’ design and impact. By addressing therapeutic benefits, gameplay balance, and accessibility, these games have the potential to make a meaningful difference in the lives of patients with PsA. Although the feedback from this session highlighted the clinical value of the storyboard elements and provided suggestions for design enhancements, there remains an opportunity to develop prototypes that integrate all these ideas to guide future development. The storyboards were revised based on the feedback collected, and a wider range of game mechanics was incorporated to improve the overall design and gameplay. Consequently, 2 NoPain game prototypes—Space Oddity Game and Four Seasons Game—have been conceptualized, as detailed next.

Prototyping the NoPain Games

Overview

Building on the subsequent phase of the agile approach after the ideation and storyboard design phase, 2 NoPain Games prototypes for pain relief were developed, incorporating targeted feedback gathered during the cocreation session (Development phase in [Figure 1](#)). These prototypes include Space Oddity ([Figure 5](#)), a space-themed serious game featuring 3 minigames that engage players in memory, coordination, and visual perception tasks; and Four Seasons ([Figure 5D](#)), a nature-themed game designed to promote repetitive, soothing movements. [Tables 2 and 3](#) provide an overview of the key characteristics of the proposed NoPain Game prototypes, including their game concepts, clinical value, visual design, game mechanics, and difficulty progression.

Figure 5. Refined storyboards for the proposed NoPain Games: (A) Space Oddity coordination challenge, (B) Space Oddity memory challenge, (C) Space Oddity visual perception challenge, and (D) Four Seasons.

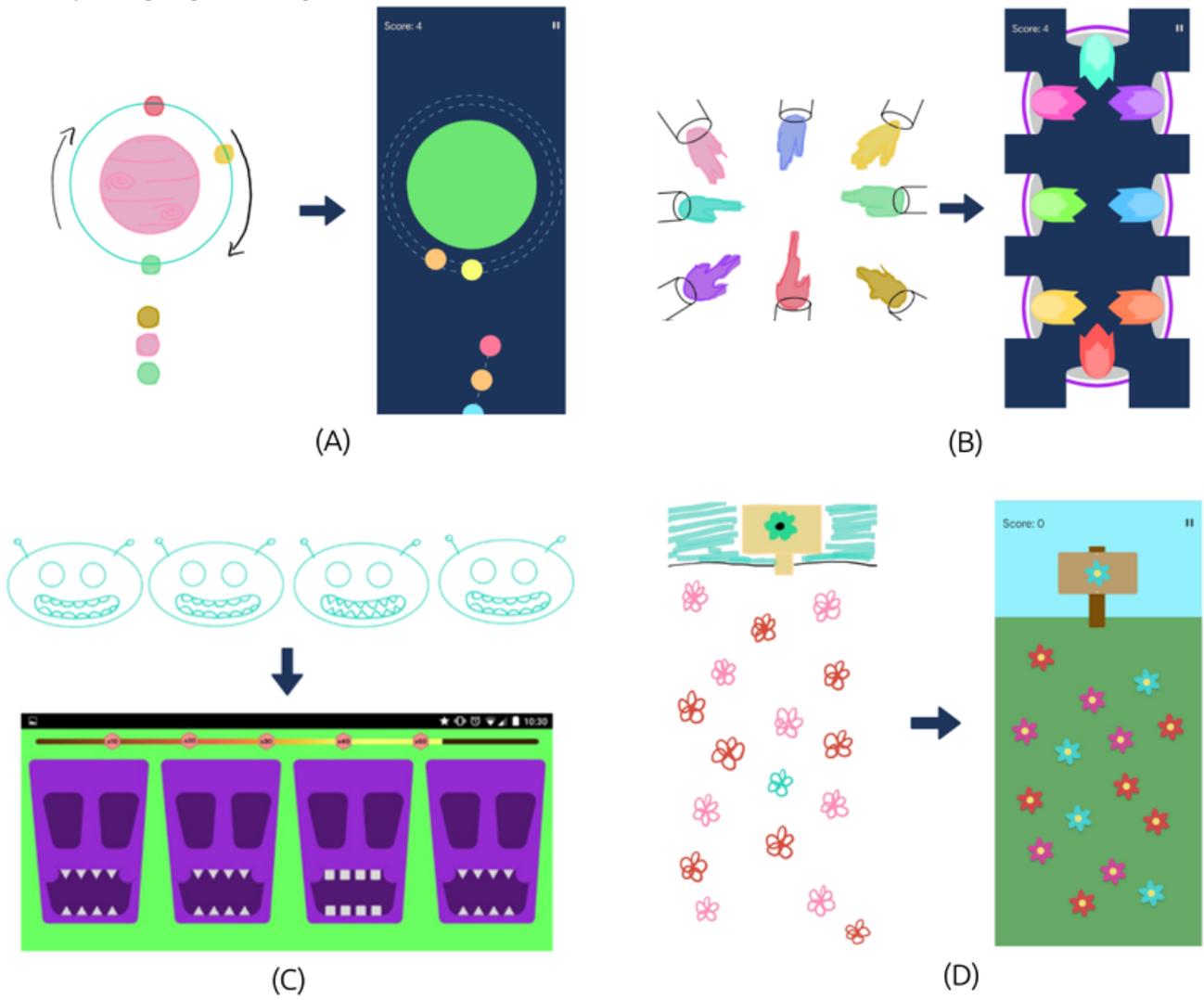


Table . Main characteristics of the proposed Space Oddity game prototypes, featuring 3 distinct mini-game-related areas, that is, memory-based challenge, coordination skills, and visual perception, each designed to target specific cognitive and motor functions associated with psoriatic arthritis. For each area, the game concept, intended clinical value, visual design elements, core gameplay mechanics, and difficulty progression are described.

	Memory-based challenge	Coordination skills	Visual perception
Game concept	Players memorize and reproduce a sequence of ignited spaceships.	Players launch asteroids into orbit while avoiding collisions.	Players identify the odd creature among a group.
Clinical value	Strengthening cognitive and memory sequencing skills.	Developing motor coordination by improving reaction timing and precision.	Improving focus, visual perception, and cognitive processing speed.
Visual design	Embracing the game's space esthetic, the spaceships' ignition includes vibrant and distinct hue colors to strengthen visual memory cues during gameplay, with clear contrast against the background for easy recognition.	A central, colorful planet with orbiting asteroids stands out against a calm-toned universe, with new asteroids emerging from below to ensure an intuitive and structured gameplay flow.	Aligned with the futuristic theme, alien robot-like creatures were chosen as the main characters, since various features (such as eye color, teeth shape, and accessories) can be easily modified for a dynamic experience.
Game mechanics	At the start of each round, a sequence of rockets lights up. The player must replicate the sequence by tapping the rockets in the correct order. The game verifies the input, granting progression if correct. If incorrect, the sequence repeats for the player to try again.	The game consists of multiple stages, each with a set number of asteroids. Players must launch asteroids into orbit while ensuring they do not collide. When all asteroids are placed successfully, a new planet appears. If 2 asteroids collide, the game ends.	A set of creatures is displayed, with one differing in subtle ways. The player must identify the intruder. Feedback is given through color and audio cues.
Difficulty progression	The sequence length increases as the game progresses, making it more challenging. An easier mode maintains the sequence and adds only one new element per round, while a harder mode randomizes the sequence each time, increasing difficulty.	Initially, asteroids move in a straight path toward the orbit. At higher difficulty levels, additional asteroids move dynamically, requiring more precise timing. Some areas on the planet may be restricted (eg, mountains), and color-coded zones may require matching asteroids.	At lower difficulties, the creatures' differences are more apparent. At higher levels, the distinctions become more subtle, requiring greater attention to detail.

Table . Main characteristics of the proposed Four Seasons Game prototypes, targeting 2 distinct areas, that is, visual perception and coordination skills, each designed to support stress reduction, focus, and motor precision in individuals with psoriatic arthritis. For each area, the game concept, intended clinical value, visual design elements, core gameplay mechanics, and difficulty progression are described.

	Visual perception	Coordination skills
Game concept	Players identify specific flower species on a garden field.	The player has to select fruits to fall from a tree, while ensuring they fall into a basket and not on the floor.
Clinical value	Stress reduction, improved focus, and coordination.	Reaction time and precision.
Visual design	A relaxing nature-based theme in the form of a garden. Various colorful plants and flowers populate the garden, giving it a vibrant atmosphere.	A nature-based theme with fruit trees in the background full of colorful fruits.
Game mechanics	The player is instructed to find and tap a chosen flower species, in a garden scenario full of various plants. Feedback is provided to let the player know if they selected the correct flower. As the game progresses, the seasons change, creating a dynamic environment and increasing the level of challenge.	The player has to select fruits to fall from a tree, as a rolling bucket passes underneath. The player gains points if they correctly calculate the timing of the falling fruit, making them land in the basket. They lose points if the fruits miss the basket instead.
Difficulty progression	As the game progresses, some elements become more challenging. For example, the flowers are no longer static; they start moving or become more similar to each other.	As difficulty rises, the number of fruits increases, the basket moves faster, and the basket becomes smaller.

Space Oddity

The Space Oddity NoPain Game was inspired by the space theme expressed in [Figure 3B](#) and incorporated a memory-based challenge with a patterns-matching mechanic similar to that in [Figure 3A](#). Based on clinician feedback provided at the second design phase, other game tasks related to coordination and visual perception were added, while maintaining the overall space thematic and base mechanics. [Table 2](#) links the specific game tasks with the themes highlighted at the second cocreation session. Throughout the game, players should experience all these tasks in the form of minigames. Therefore, as the players enter Space Oddity and select a difficulty level, 3 minigames appear in a different order to avoid monotony and ensure that the game will be played various times without getting repetitive.

The first minigame is designed to train coordination skills ([Figure 5A](#)). A main planet stands at the center of the screen, while being surrounded by a fixed-speed orbit of asteroids. As new asteroids emerge from below, the player/patient must launch them into orbit one at a time, ensuring they do not collide with existing asteroids. Success depends on precise timing, requiring the player/patient to tap the screen at the right moments to avoid collisions. After a certain number of asteroids are launched, a new planet emerges with new challenges. Additional features to add diversity to the game include areas on the planet with mountains where asteroids cannot be placed, or color-coded zones to correspond to asteroids of the same color ([Figure 5A](#)). The second minigame, depicted in [Figure 5B](#), features 8 spaceships scattered across the screen, all starting with their engines turned off. It operates, like the Simon game [36], in rounds where players watch as a sequence of rockets ignite and then are required to replicate the sequence by selecting the rockets in the correct order. The first round begins with a single randomly ignited rocket. In each subsequent round, the previous sequence repeats in the same order, with one additional rocket added. The minigame continues until a certain number of rounds is successfully completed. To enhance memorization, unique colors and musical notes can be attributed to each rocket, providing both visual and auditory stimuli. [Figure 5B](#) illustrates the core prototype elements and an example screen in which all rockets are ignited with their distinct colors. The third minigame focuses on memory, concentration, and visual perception, where a player/patient performs several rounds of identifying from a set of 4 creatures, the one that differs from the others (intruder) ([Figure 5C](#)). Overall, various features may be modified to create

distinctions, such as the creatures' shape or variations in their elements.

Four Seasons

The Four Seasons NoPain Game builds upon the concepts illustrated in the storyboard of [Figure 3C](#). Similarly, [Tables 2](#) and [3](#) link the specific game tasks with the themes highlighted at the second cocreation session. During gameplay, the players/patients explore a garden filled with various flowers and are asked to identify a specific one ([Figure 5D](#)). As the game progresses, the seasons change, affecting both the environment's aesthetic and the challenges presented. Each season introduces unique difficulties and advantages, adding variety to the game. The Four Seasons game was designed to distract players from pain by offering 2 distinct modes to suit individual preferences. The Relax Mode allows players to peacefully search for flowers in a garden across different seasons, while the Party Mode introduces added excitement with minigames at the end of each season for a more dynamic experience. Overall, this game, designed to reflect the unique characteristics of each season, has the potential to offer a refreshing gameplay variety while introducing an element of dynamism for players/patients seeking a more immersive experience.

Feedback From Patients With PsA

Quantitative Analysis

The patients with PsA feedback group included 2 women and 3 men, with ages ranging from 25 to 64 years, split into age groups of 25 - 34 years (1), 35 - 44 years (1), 45 - 54 years (2), and 55 - 64 years (1). All patients had been diagnosed with PsA for more than 8 years; had an average experience in gaming, and good experience in smartphone usage. During the session, 4 participants (1-4) reported experiencing a PsA flare.

The session first focused on "Space Oddity" through a playable version developed in Unity3D ([Figure 4A](#)). Participants played the game and then evaluated its usability by completing the SUS questionnaire. [Table 4](#) tabulates the derived descriptives from the SUS data, that is, mean (SD), 95% CI, and median (IQR). From the latter, it is seen that an average SUS score of 79 (SD 10.4; 95% CI 69.89 - 88.11) was reached, suggesting good overall usability, with some variation in individual user perceptions.

Table . Descriptives of the System Usability Scale (SUS).

Item definitions	Mean (SD)	95% CI	Median (IQR)
Frequency Use (1)	3.80 (0.84)	3.07 - 4.53	4 (3-4)
Complexity (2)	2.40 (1.95)	0.69 - 4.11	1 (1-4)
Ease of Use (3)	4.60 (0.55)	4.12 - 5.08	5 (4-5)
Support Needs (4)	1.00 (0.0)	1.00 - 1.00	1 (1-1)
Functionality Integration (5)	3.80 (0.84)	3.07 - 4.53	4 (3-4)
Inconsistency (6)	2.00 (1.0)	1.12 - 2.88	2 (1-3)
Confidence (7)	4.40 (0.55)	3.92 - 4.88	4 (4-5)
Cumbersomeness (8)	1.60 (0.90)	0.82 - 2.38	1 (1-2)
Learnability (9)	4.20 (0.84)	3.47 - 4.93	4 (4-5)
Learning Curve (10)	2.20 (1.64)	0.76 - 3.64	2 (1-2)
Total SUS	79 (10.4)	69.89 - 88.11	75 (75 - 82.5)

Moreover, [Figure 6](#) displays a 2D radar plot of average SUS responses across all participants, with each axis representing a usability factor mapped to its corresponding SUS item number (eg, Frequency Use (1), Cumbersomeness (8), etc). The radial layout allows for intuitive comparison of agreement levels across items, with values ranging from 1 (“Strongly Disagree”) to 5 (“Strongly Agree”). Peaks in the plot are observed around positively worded items such as Ease of Use (3), Confidence (7), and Learnability (9), indicating strong agreement and direct contributions to higher SUS scores. Conversely, troughs appear around negatively worded items like Support Needs (4), Cumbersomeness (8), and Learning Curve (10), where low agreement is desirable and reflects positive usability perceptions. The overall shape of the radar plot reveals a balanced usability profile, with high agreement on core interaction elements and consistent disagreement with statements implying complexity or poor integration.

In addition, [Figure 6B](#) presents a 3D radar surface plot where axes radiate outward from the center, each labeled with a usability factor and its corresponding SUS item number, for example, Consistency (6) and Learning Curve (10). The surface height represents participant index, while the color gradient (blue to yellow) encodes response values from 1 to 5. The plot reveals structural patterns in usability perception, with elevated

regions around Ease of Use (3), Confidence (7), and Learnability (9), positively worded items where high agreement directly contributes to higher SUS scores. Conversely, lower scores on negatively worded items such as Complexity (2), Support Needs (4), Inconsistency (6), Cumbersomeness (8), and Learning Curve (10) are desirable, as they reflect disagreement with statements implying poor usability. Notably, Support Needs (4) received a uniform score of 1 across all participants, indicating unanimous disagreement with the notion that external support was required to use the system, an encouraging signal of intuitive design and self-sufficiency. Participant-level differences are also evident. In particular, Participants 1 - 4, who reported experiencing PsA flare episodes during evaluation, showed slightly more variability in items such as Functionality Integration (5) and Inconsistency (6), suggesting that symptom severity may influence perceptions of system responsiveness and coherence. In contrast, Participant 5, who did not report a flare, exhibited consistently high agreement on positively worded items and low agreement on negatively worded ones, resulting in a smoother and more elevated usability profile across the radar surface. Together, [Figure 6A and B](#) offer complementary perspectives on usability perception, highlighting both structural consistency and individual-level nuances.

Figure 6. (A) Radar chart of mean SUS item scores arranged clockwise, highlighting usability strengths and weaknesses. (B) 3D radar surface plot of individual SUS responses, showing participant-level variation across usability dimensions. SUS: System Usability Scale.

Furthermore, Figure 7 illustrates a detailed visualization of item-level behavior across the latent usability trait spectrum. The left panel presents 2D ICCs for selected SUS items, each labeled by item index. These curves illustrate the probability of agreement $p(\theta)$ (see (1)) as a function of latent usability (θ), revealing distinct patterns of discrimination and difficulty. The right panel complements this view with a 3D ICC ribbon plot, where each SUS item is vertically separated and labeled by its full descriptor. This format enhances interpretability by exposing how response probabilities vary not only across θ but also across items. Moreover, Table 5 includes the corresponding parameters (a,b) (see (1)) of the ICCs depicted in Figure 7. Analytically, from Figure 7 and Table 5, the following observations can be derived per SUS item, revealing their heterogeneity:

- Frequency of Use (1)-positive wording: With $a = 1.29$ and $b = -1.10$, this item shows moderate discrimination and early activation. The curve is gently sloped and elevated across the trait spectrum, indicating consistent agreement even at low usability levels. This suggests it reflects habitual or contextual familiarity rather than sensitivity to usability changes. Its contribution is stable but diagnostically limited.
- Complexity (2)-negative wording: With $a = 0.67$ and $b = 0.49$, the curve rises gradually and activates in the mid-range of θ . Users increasingly disagree as usability improves, aligning with expected polarity. However, the low a value indicates weak discrimination. The ribbon is broad and shallow, offering general friction-related information but lacking precision.
- Ease of Use (3)-positive wording: Despite its conceptual importance, $a = 1.73 \times 10^{13}$ and $b = -5.93 \times 10^{14}$ suggest a modeling anomaly, likely due to uniform agreement across participants. The curve appears flat and saturated, with no meaningful slope. While users universally endorse this item, its statistical discrimination is negligible, making it confirmatory rather than diagnostic.
- Support Needs (4)-negative wording: With $a = -8.72 \times 10^{14}$ and $b = -1.18 \times 10^1$, this item exhibits modeling instability. The curve is flat and extremely right-shifted, indicating that nearly all users disagreed regardless of usability level. In fact, all users disagreed. This uniform rejection aligns with high usability but limits the item's ability to differentiate experiences. It may be conceptually relevant but statistically redundant.
- Functionality Integration (5)-positive wording: With $a = -0.11$ and $b = 3.48$, the curve is shallow. The negative a suggests an inverse slope, possibly due to inconsistent response patterns. The ribbon is low and delayed, indicating poor discrimination and limited responsiveness.
- Inconsistency (6)-negative wording: Like item 4, $a = -8.72 \times 10^{14}$ and $b = -1.18 \times 10^1$ reflect modeling collapse. The curve is flat and nondiscriminative, with uniform disagreement across users. While this aligns with high usability, the lack of slope or variability renders it statistically inert.
- Confidence (7)-positive wording: With $a = 1.73 \times 10^{13}$ and $b = -5.93 \times 10^{14}$, this item shows extreme early activation and saturation. The curve is flat and elevated, indicating universal agreement. Like *Ease of Use (3)*, it functions as a strong confirmatory item but lacks statistical discrimination due to the ceiling effect.
- Cumbersomeness (8)-negative wording: Again, $a = -8.72 \times 10^{14}$ and $b = -1.18 \times 10^1$ suggest modeling failure. The curve is flat and low, with near-universal disagreement. While this reflects low friction, the lack of variability limits its diagnostic value.
- Learnability (9)-positive wording: With $a = 59.00$ and $b = -2.28$, this item shows extremely high discrimination and early activation. The curve is steep and sharply contoured, making it the most responsive item in the set. The ribbon is narrow and elevated, indicating strong sensitivity to usability perception across the trait range. This item is statistically and conceptually robust.
- Learning Curve (10)-negative wording: With $a = 0.33$ and $b = 4.33$, this item exhibits low discrimination but great difficulty. The curve rises gradually and activates only at the upper end of the usability trait spectrum. The ribbon in Figure 7 is delayed but upward-sloping, confirming that the item captures late-stage usability clarity, particularly in systems that excel in onboarding. While its responsiveness is limited at lower θ levels, it still contributes meaningful coverage of the high-usability tail, making it useful for identifying systems with strong learnability.

From the aforementioned, we can identify four interpretive categories: (1) Statistically robust item: "Learnability (9)" shows exceptional discrimination and early activation, making it the most diagnostically powerful item. (2) Confirmatory but saturated items: "Ease of Use (3)" and "Confidence (7)" are universally endorsed but exhibit near-zero discrimination due to ceiling effects. They confirm high usability but offer limited differentiation. (3) Low-performing or inert items: "Support Needs (4)," "Inconsistency (6)," and "Cumbersomeness (8)" show flat, nonresponsive curves and modeling collapse, limiting their diagnostic utility. (4) Mid-range and coverage items: "Frequency of Use (1)," "Complexity (2)," and "Learning Curve (10)" contribute moderate or delayed responsiveness, offering coverage across the trait spectrum but with limited precision. This item-level interpretation supports evaluation of the SUS, guiding decisions on item weighting and further scrutiny for potential refinement.

Figure 7. 2D Item characteristic curves (left) showing agreement probability $p(\theta)$ across latent usability trait (θ), with each curve labeled by item index, along with the 3D item characteristic curve ribbons (right) for all System Usability Scale items, visualizing response probabilities across θ and item index, with full item labels on the horizontal axis.

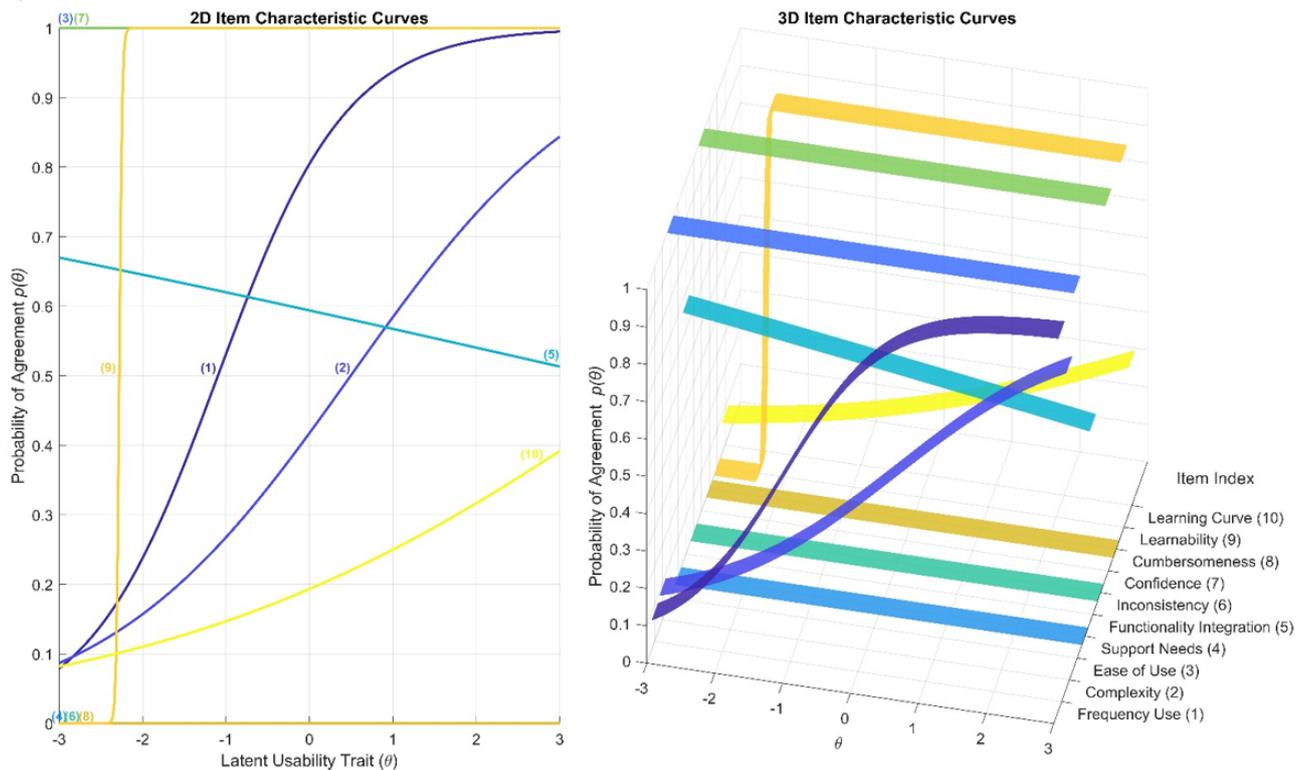


Table . Estimated parameters (a, b) of the corresponding item characteristic curves of Figure 7 for each System Usability Scale (SUS) item.

SUS item	Alpha (a)	Beta (b)
Frequency Use (1)	1.29	-1.10
Complexity (2)	0.67	0.49
Ease of Use (3)	1.73×10^{-13}	-5.93×10^{14}
Support Needs (4)	-8.72×10^{-14}	-1.18×10^{15}
Functionality Integration (5)	-0.11	3.48
Inconsistency (6)	-8.72×10^{-14}	-1.18×10^{15}
Confidence (7)	1.73×10^{-13}	-5.93×10^{14}
Cumbersomeness (8)	-8.72×10^{-14}	-1.18×10^{15}
Learnability (9)	59.00	-2.28
Learning Curve (10)	0.33	4.33

Qualitative Analysis

During the patients with PsA feedback session, the group was guided through a design prototype demonstration of the Four Seasons (Figure 4B), covering its tutorial, practice session, navigation, gameplay, bonus level, and scoring. Feedback was generally positive and constructive, with discussions focusing on several key areas across the 2 games. In particular, they proposed:

- resolving iOS compatibility issues to improve accessibility,
- modifying difficulty levels and incorporating adaptive artificial intelligence (AI) in “Space Oddity” to tailor challenges to individual player profiles,

- transitioning “The Four Seasons” from prototype to full interactive development with seasonal mechanics and high-quality 3D rendering, and
- introducing smartwatch integration to track stress levels before and after gameplay in both games.

From a combined perspective, the derived average SUS score of 79 reflects a generally positive user experience, placing the NoPain Games within the “good” usability range according to established benchmarks. This score suggests that core interaction elements, such as navigation, responsiveness, and interface clarity, were well-received across a diverse group of patients with PsA, even during flare episodes. To better understand how this overall score reflects item-level behavior, we examined the

individual ICC profiles. The accompanying ICC analysis revealed several positive outcomes. These item-level patterns help contextualize the overall SUS score of the NoPain Games, revealing how specific usability dimensions contributed to the positive user experience. This is combined with the feedback provided during the session, which revealed that usability alone does not fully capture user expectations in therapeutic contexts. Participants articulated a clear desire for functional enhancements and personalization features that extend beyond baseline usability. Their suggestions, ranging from adaptive AI and platform compatibility to immersive rendering and biometric integration, indicate that patients are not merely evaluating ease of use, but actively envisioning how the system could evolve into a more responsive and clinically meaningful tool. This feedback reveals the importance of interpreting usability scores in tandem with qualitative insights, especially when designing interventions for chronic conditions where therapeutic relevance and emotional engagement are critical.

Discussion

Main Findings

This study aimed to explore expert perspectives on the design of 2 serious games, that is, NoPain Games, intended to alleviate chronic pain and improve quality of life for individuals with PsA. The cocreation session revealed 3 key thematic areas: therapeutic benefits, game difficulty, and accessibility and interaction. These themes directly informed modifications to the game prototypes, aligning with the study's objective to develop personalized, user-centered digital interventions for chronic pain management. Moreover, the usability assessments of patients with PsA using SUS and feedback further supported the design approach, indicating generally positive perceptions of the game prototypes and reinforcing their potential as accessible and engaging tools for digital pain management. The extended quantitative analysis of SUS responses, including item-level discrimination and difficulty modeling via ICCs, provided deeper insight into usability perceptions. These findings offer item-level insights that can guide targeted enhancements to interface design and interaction mechanics. This quantitative granularity complements the thematic findings, particularly those related to therapeutic intent and user engagement.

Therapeutic benefits were a major focus, with the games designed to provide pain distraction and cognitive engagement. For example, the "Lightning" game focuses on memory enhancement, which can help distract from pain and improve cognitive function [3]. Additionally, the games incorporate soothing and repetitive mechanics, such as in the "Flashing Stars" game, which promotes relaxation and stress reduction [6]. Moreover, previous studies had also suggested a link between serious games and improvements in working memory [37], prompting further research on this topic.

Game difficulty was another critical theme, with feedback emphasizing the need to balance game duration and complexity to maintain engagement without causing fatigue. This led to the development of adaptive difficulty levels that adjust to the player's abilities [38]. The games were designed to be

challenging enough to keep players engaged but not so difficult that they become frustrating. This balance is crucial for maintaining the therapeutic benefits of the games [39]. In fact, the challenge of balancing game difficulty has been identified in previous studies [38,39], yet it remains complex due to the need to provide therapeutic benefits while maintaining player engagement. When designing serious games, it is essential to balance activity levels, as over-exercise can cause fatigue and frustration, while insufficient activity may reduce therapeutic benefits. Adaptive difficulty levels aligning with players' abilities and goals could ensure engagement, gradual improvement, and overall well-being.

Accessibility and interaction were also key considerations. Overall, the game design considered the physical limitations of patients with PsA, such as joint stiffness and swelling. This led to the optimization of touchscreen usability and the inclusion of larger touch targets to accommodate swollen fingers [40]. Ensuring that the games are accessible to all patients with PsA, regardless of their physical abilities, was a key focus. This included designing intuitive and easy-to-use interfaces [41]. Moreover, smartphone accessibility has been analyzed in various domains [40,41] due to the fine motor skills required for small touchscreen interactions. However, its impact on patients with PsA could be investigated further.

The alterations in game design during the cocreation workshops align with findings from existing literature on serious games and digital health interventions. Studies have shown that serious games can reduce stress, enhance cognitive function, and provide pain relief through immersive and engaging experiences [3,6]. The NoPain Games leverage these benefits by incorporating memory and relaxation activities. The importance of balancing game difficulty to sustain engagement and therapeutic efficacy is well-documented. Adaptive difficulty levels, planned for implementation in the NoPain Games, are crucial for maintaining player interest and ensuring the games' effectiveness [38,39]. Research highlights the need for accessible game design, especially for individuals with physical limitations. The NoPain Games' focus on optimizing touchscreen usability and inclusive design reflects these findings [40,41]. Furthermore, the conceptual framework (2D-ME) for explaining self-first and self-third person views of prototyping dynamics in serious games design highlights the importance of iterative feedback and dynamic constructs in the game design process [42]. This framework supports the iterative refinement seen in the NoPain Games development, ensuring that the games are both engaging and therapeutic.

Furthermore, the design of NoPain Games leverages key insights into how the human brain pays attention to various aspects of game design. By understanding and applying principles related to perception, memory, attention, and emotional engagement, the developers can create games that are both therapeutic and engaging.

For example, perception plays a crucial role in how players interact with NoPain Games. The brain processes visual and auditory information to understand and navigate the game world. For instance, the "Lightning" NoPain game uses effective visual cues and sound effects to guide players' attention and enhance

their immersion [43]. The UI design is also critical, ensuring that players can easily access information and controls without being distracted by confusing or cluttered interfaces. This is particularly important for patients with PsA, who may have physical limitations affecting their game interaction [40].

Memory is another essential aspect of the proposed NoPain Games. Players rely on short-term memory to remember recent actions, objectives, and game mechanics. This is particularly important in fast-paced games where quick decision-making is required. For example, the “Lightning” NoPain Game focuses on memory enhancement, helping players to distract from pain and improve cognitive function [3]. Long-term memory is engaged when players learn and remember game rules, storylines, and strategies. By reinforcing learning through repetition and rewards, NoPain Games can enhance player retention and engagement [43].

Attention is also vital for maintaining focus and engagement in the proposed NoPain Games. The brain’s ability to maintain focus is critical for sustained engagement, and the developers have used various techniques to capture and hold players’ attention. This includes compelling narratives and dynamic gameplay [44]. Managing cognitive load is essential to prevent player fatigue. NoPain Games balance complexity and simplicity to keep players engaged without overwhelming them. This involves designing intuitive controls, clear objectives, and manageable challenges, ensuring that the games are accessible and enjoyable for patients with PsA.

Emotional and motivational factors significantly influence the brain’s attention to NoPain Games. Emotions play a significant role in how players experience the games. Engaging storylines, relatable characters, and emotional rewards enhance player immersion and satisfaction. The brain responds to rewards and achievements, which can motivate players to continue playing. NoPain Games use various reward systems, such as points and progression systems, to keep players motivated and engaged [43].

Achieving a flow state is another critical aspect of NoPain Games. The concept of “flow” refers to a state of deep focus and immersion where players lose track of time [45]. Achieving flow involves balancing challenge and skill, providing clear goals, and offering immediate feedback. NoPain Games successfully induce flow by creating a balance between challenge and skill, ensuring that players remain deeply engaged and enjoy the therapeutic benefits of the games [43].

Approaching the aforementioned from the lens of the 3 themes identified, that is, therapeutic benefits, game difficulty, and accessibility and interaction, a deep interwovenness can be identified with sustained engagement, player attention, cognitive load, player fatigue, emotions, and flow state in the context of NoPain Games for patients with PsA. In particular:

1. **Therapeutic benefits:** NoPain Games aim to alleviate pain and improve emotional well-being by providing cognitive stimulation, stress reduction, and memory enhancement. These benefits are crucial for fostering sustained engagement as players find therapeutic value in their activity, which keeps their attention fixed on the games.

The soothing and predictable mechanics of games like “Flashing Stars” reduce *cognitive load* by presenting clear objectives and repetitive tasks, allowing players to focus without being overwhelmed. This emotional engagement also helps maintain the ideal flow state, where players are immersed in gameplay and distracted from their chronic pain.

2. **Game difficulty:** Balancing and adapting difficulty levels directly supports sustained engagement by keeping tasks manageable yet stimulating. Customizable difficulty prevents player fatigue, ensuring sessions remain enjoyable rather than exhausting. A dynamically adjusted challenge ensures that attention is maintained without inducing frustration, helping to optimize both players’ attention and cognitive effort. Gradual progression of difficulty, such as increasing complexity in the “Lightning” game, enhances emotions of achievement and satisfaction, which is key for reinforcing engagement and maintaining the flow state.
3. **Accessibility and interaction:** The design considerations in this theme, such as larger touch targets and intuitive interfaces, ensure that physical limitations do not hinder participation. This inclusivity supports sustained engagement by enabling patients to interact easily and prevent frustration. These accommodations reduce the cognitive load on players, allowing them to focus on the tasks rather than overcoming interface challenges. This theme also reduces players’ fatigue by minimizing physical strain and frustration, helping to maintain a relaxed state essential for emotional well-being and immersion into a flow state.

Together, these themes create a synergistic experience where accessibility removes barriers, therapeutic benefits uplift emotional and cognitive states, and adaptive difficulty sustains attention and motivation, all of which harmonize to engage patients with PsA while mitigating pain effectively.

The iterative optimization process in developing NoPain Games prototypes involved refining visual cues, simplifying the UI, balancing cognitive load, and incorporating emotional and motivational elements. By understanding how the brain pays attention to different aspects of game design, the designers/developers were able to create games that are both effective and enjoyable for patients with PsA. This approach ensured the potential for the proposed NoPain Games to provide therapeutic benefits while maintaining high engagement and satisfaction levels.

Overall, the cocreation themes guided the iterative refinement of the NoPain Games, confirming they are both engaging and therapeutic for patients with PsA. The integration of feedback from the cocreation session with insights from the relevant literature review resulted in 2 game prototypes that are well-suited to the needs of their target audience, as expressed via the feedback from patients with PsA. In fact, the patient feedback session extended the thematic framework by introducing concrete implementation priorities that reflect lived experience and technical expectations. The mean effective SUS score (79) confirmed good usability, yet the qualitative suggestions revealed areas for refinement that map directly onto the 3 core themes.

First, the recommendation to resolve iOS compatibility issues reinforces theme 3 (accessibility and interaction). While the expert session emphasized touchscreen optimization and inclusive design, patients highlighted platform-specific barriers that could limit access. This aligns with Beltran-Alacreu et al [16], who emphasized the importance of device-level accessibility in digital interventions for chronic pain populations. Addressing cross-platform compatibility is essential to ensure equitable access, particularly for older adults or those with limited digital literacy.

Second, the proposal to incorporate adaptive AI into Space Oddity to tailor difficulty to individual profiles directly advances theme 2 (game difficulty). This feedback moves beyond static balancing and introduces personalization as a therapeutic strategy. It resonates with Tuah et al [20], who identified progression systems and adaptive mechanics as key to sustaining engagement in rehabilitation games. In the context of PsA, where symptom severity and cognitive capacity fluctuate, AI-driven difficulty modulation could enhance both usability and therapeutic relevance.

Third, the suggestion to transition the Four Seasons from prototype to full interactive development with seasonal mechanics and high-quality 3D rendering reflects an evolution of theme 1 (therapeutic benefits). While experts emphasized cognitive stimulation and stress reduction, patients implicitly called for deeper immersion and aesthetic refinement. This aligns with Gromala et al [6], who argued that realistic environments and sensory richness contribute to pain relief. The seasonal metaphor may also support emotional regulation by anchoring gameplay in familiar, cyclical rhythms.

Finally, the idea of integrating smartwatch-based stress tracking introduces a novel extension to Theme 1, bridging subjective experience with physiological feedback. This aligns with recent findings by Pinge et al [46], who systematically reviewed wearable-based stress detection and highlighted the clinical potential of physiological signals such as heart rate variability and electrodermal activity for real-time stress monitoring. By capturing pre- and post-game stress levels, future iterations of NoPain Games could offer biofeedback-informed personalization, enhancing therapeutic precision and enabling longitudinal tracking within digital care ecosystems.

Overall, these patient-driven suggestions not only validate the thematic structure but also push its boundaries toward real-world deployment. They emphasize the importance of technical adaptability, personalized challenge design, and physiological integration in therapeutic game development. More broadly, they demonstrate how participatory feedback can transform conceptual frameworks into actionable design trajectories, ensuring that digital health tools remain grounded in both clinical insight and patient reality.

Limitations

Although this study offers important insights, it is essential to acknowledge certain limitations. The study engaged a limited sample of 14 experts from 3 European countries in one session, followed by 5 patients with PsA from 4 European countries in a subsequent session, including SUS-based evaluation. While

the results offer initial insights into usability perceptions, the limited sample size may have constrained variability, leading to some of the ICC curves being saturated, reducing the ability to detect item-level differentiation. Future research can build on this by involving a broader and more diverse participant pool to deepen understanding of the varied needs and experiences of individuals living with PsA. Particularly, including more testing with patients is necessary, which is essential to the proper design and effectiveness of the games despite the recruiting challenges associated with this task. Disparities in technology literacy among participants, especially health care professionals, also posed a challenge. Providing training or educational materials on the technologies used in the serious games could bridge this gap and improve feedback quality. Moreover, as the study is still at the prototype stage, further testing and clinical validation with patients with PsA through pilot studies and trials will be essential to refine the games based on real-world usability data. Finally, the proposed game designs may not fully consider the physical limitations of patients with PsA, such as joint stiffness and swelling, which can affect usability. Future iterations should incorporate adaptive game design features like customizable controls to improve accessibility.

Implications

The research on co-designing serious mobile games for patients with PsA presented here has several important implications across various domains.

From a managerial standpoint, the study underlines the importance of interdisciplinary collaboration in developing effective digital health interventions. Managers in health care and technology sectors should foster partnerships between clinicians, researchers, and game developers to leverage their combined expertise. This collaborative approach can lead to more innovative and user-centered solutions [47]. Additionally, the agile co-design methodology highlighted in the study can be adopted by managers to ensure continuous feedback and iterative improvements, enhancing the overall quality and relevance of the developed digital solutions [48]. By promoting a culture of collaboration and iterative development, managers can drive the successful implementation of digital health tools that meet the specific needs of patients [48,49].

From a financial perspective, the development and implementation of NoPain Games for patients with PsA could lead to significant cost savings in health care. By providing an engaging and effective tool for pain management and cognitive stimulation, these games can reduce the reliance on more expensive treatments and interventions [50]. Improved patient adherence to therapeutic activities can also lead to better health outcomes, potentially decreasing the frequency of hospital visits and the need for additional medical support. Furthermore, the scalability of mobile games means that they can be distributed widely at a relatively low cost, making them an economically viable option for health care providers and patients alike [51]. Investing in the development of such digital health solutions can yield long-term financial benefits for both health care systems and patients.

From the clinical point of view, while the NoPain Games presented here are still in the prototype stage and have not yet

been used in clinical practice, their potential therapeutic benefits are promising. The proposed serious games are designed to provide pain distraction, cognitive stimulation, and stress reduction, which could significantly improve the quality of life for patients with PsA. Future clinical trials and real-world testing will be crucial to validate these benefits and refine the game mechanics based on patient feedback [52]. If proven effective, these serious games could be integrated into treatment plans as complementary therapy, offering a novel approach to managing chronic pain and enhancing cognitive function. This research also highlights the importance of adaptive difficulty levels and personalized gameplay experiences, which could be further explored in clinical settings to tailor the interventions to individual patient needs [53].

Finally, from a societal perspective, the emphasis on accessibility and inclusive design in the NoPain Games highlights the importance of creating digital health solutions that are usable by individuals with varying physical abilities. By addressing the specific needs of patients with PsA, the games promote greater equity in health care access and support for chronic pain management. This inclusive approach can serve as a model for developing other digital health interventions that address the needs of diverse patient populations [54]. Moreover, the widespread adoption of such games can raise awareness about the potential of serious games in health care, encouraging further innovation and investment in this field [51]. By making effective pain management tools accessible to a broader audience, the research can contribute to improved health outcomes and quality of life for many individuals living with chronic conditions.

In contrast to prior studies that focus on general chronic pain, this research is the first to co-design mobile serious games specifically for patients with PsA, integrating both expert and patient perspectives. It contributes to the field by demonstrating how agile, participatory design can yield tailored, accessible, and therapeutically meaningful digital interventions for a rheumatic population largely overlooked in serious game development.

Conclusions

The proposed NoPain Games hold significant potential for alleviating chronic pain in individuals with PsA by providing engaging and immersive gameplay experiences aimed at reducing discomfort. This study highlights the collaborative efforts of researchers and clinical/technical experts in designing NoPain Games as cocreators. It introduces 2 serious game prototypes developed using agile methodology and co-design principles, incorporating expert feedback. The study explores the cocreation process, presenting initial findings through storyboards, game visualizations, and prototypes informed by the collected input. Future work includes obtaining further patient feedback on the prototypes and conducting real-world testing to evaluate their feasibility, acceptability, and overall user satisfaction. The findings highlight the critical role of managing game difficulty, which can be addressed through the integration of a dynamic difficulty adjustment system [38,39] in developing the game to customize challenges based on each patient's condition. Moreover, potential additional features could include integrating smartwatch-based biometric data, such as stress levels and heart rate, to personalize gameplay experiences and provide valuable clinical insights.

Acknowledgments

The authors would like to thank and acknowledge all members of the iPROLEPSIS Consortium for their valuable contributions. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or European Health and Digital Executive Agency. Neither the European Union nor the European Health and Digital Executive Agency can be held responsible for them. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Funding

This work received funding from the European Union under the Horizon Europe Grant agreement number 101095697 (iPROLEPSIS: Psoriatic Arthritis Inflammation Explained through Multi-Source Data Analysis guiding a Novel Personalized Digital Care Ecosystem).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy restrictions and consortium policies on data sharing, but are available from the corresponding author on reasonable request.

Authors' Contributions

BR and SBD contributed to the study conceptualization and design; BR and SBD collected the data; BR, S Gomes, and SBD contributed to project administration. BR, S Gomes, MSV, FM, RGC, and SBD contributed to data curation and formal analysis; LH and SBD contributed to funding acquisition; BR, S Gomes, MSV, FM, RGC, LH, and SBD contributed to the methodology; BR, S Gomes, and SBD contributed to writing the original draft; BR, S Gomes, MSV, FM, RGC, S Gama, LH, and SBD contributed to reviewing and editing the paper; all authors reviewed and approved the final version of the manuscript.

We thank the collaborators from iPROLEPSIS Consortium: Leontios J. Hadjileontiadis, Vasileios Charisis, Georgios Apostolidis, Nikos Athanasopoulos, Eleni Vasileiou, Despina Petsini, Evdokimos Konstantinidis, Konstantina Tsimpita, Panagiotis Bamidis,

Jolanda Luime, Patty de Groot, Ilja Tchetverikov, Batoul Hojeij, Andreas Raptopoulos, Dimitrios Karamitros, Sofia Balula Dias, Samuel Gomes, Bárbara Ramalho, Marta Vicente, Filomena Carnide, Fátima Baptista, Kosmas Dimitropoulos, Dimitris Konstantinidis, Laura Coates, Cátia Gonçalves, Ana Rodrigues, Francesca Levi-Schaffer, Prince Ofori, Silvia Reis, Hugo Silva, Daniel Osório, Rodrigo Braga, Nikos Melanitis, Thanos Vidakis, Alex Bensenousi, Kristina Leipuviene, Ioannis Drivas, Amalia Ntemou, and Nikoleta Tsampanaki.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Semistructured guide script for the cocreation session.

[[DOCX File, 2785 KB - games_v14i1e75072_app1.docx](#)]

Multimedia Appendix 2

Online consent form.

[[PNG File, 306 KB - games_v14i1e75072_app2.png](#)]

References

1. Eichenberg C, Grabmayer G, Green N. Acceptance of serious games in psychotherapy: an inquiry into the stance of therapists and patients. *Telemed J E Health* 2016 Nov;22(11):945-951. [doi: [10.1089/tmj.2016.0001](#)] [Medline: [27045452](#)]
2. Brauner P, Ziefle M. Social acceptance of serious games for physical and cognitive training in older adults residing in ambient assisted living environments. *J Public Health (Berl)* 2022 Jan;30(1):63-75. [doi: [10.1007/s10389-021-01524-y](#)]
3. 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](#)]
4. Neri SG, Cardoso JR, Cruz L, et al. Do virtual reality games improve mobility skills and balance measurements in community-dwelling older adults? systematic review and meta-analysis. *Clin Rehabil* 2017 Oct;31(10):1292-1304. [doi: [10.1177/0269215517694677](#)] [Medline: [28933612](#)]
5. Pacheco TBF, de Medeiros CSP, de Oliveira VHB, Vieira ER, de Cavalcanti FAC. Effectiveness of exergames for improving mobility and balance in older adults: a systematic review and meta-analysis. *Syst Rev* 2020 Jul 18;9(1):163. [doi: [10.1186/s13643-020-01421-7](#)] [Medline: [32682439](#)]
6. Gromala D, Tong X, Shaw C, Amin A, Ulas S, Ramsay G. Mobius Floe: an immersive virtual reality game for pain distraction. *Electronic Imaging* 2016;28:1-5. [doi: [10.2352/ISSN.2470-1173.2016.4.ERVR-413](#)]
7. Saragih ID, Suarilah I, Saragih IS, Lin YK, Lin CJ. Efficacy of serious games for chronic pain management in older adults: a systematic review and meta-analysis. *J Clin Nurs* 2024 Mar;33(3):1185-1194. [doi: [10.1111/jocn.17012](#)] [Medline: [38291564](#)]
8. Vivekanantham A, McGagh D, Coates LC. Current treatments and recommendations for Psoriatic Arthritis. *Best Pract Res Clin Rheumatol* 2021 Jun;35(2):101680. [doi: [10.1016/j.berh.2021.101680](#)]
9. Hetland ML. Psoriatic arthritis: still room for improvement. *Lancet* 2020 May;395(10235):1463-1465. [doi: [10.1016/S0140-6736\(20\)30800-X](#)]
10. Fox S, Brown LJE, Antrobus S, et al. Co-design of a smartphone app for people living with dementia by applying agile, iterative co-design principles: development and usability study. *JMIR Mhealth Uhealth* 2022 Jan 14;10(1):e24483. [doi: [10.2196/24483](#)] [Medline: [35029539](#)]
11. Martin A, Caon M, Adorni F, et al. A mobile phone intervention to improve obesity-related health behaviors of adolescents across Europe: iterative co-design and feasibility study. *JMIR Mhealth Uhealth* 2020 Mar 2;8(3):e14118. [doi: [10.2196/14118](#)] [Medline: [32130179](#)]
12. Vasileiou E, Dias SB, Hadjimitriou S, et al. Novel digital biomarkers for fine motor skills assessment in psoriatic arthritis: The DaktylAct touch-based serious game approach. *IEEE J Biomed Health Inform* 2025 Jan;29(1):128-141. [doi: [10.1109/JBHI.2024.3487785](#)] [Medline: [39471112](#)]
13. Gold JI, Belmont KA, Thomas DA. The neurobiology of virtual reality pain attenuation. *Cyberpsychol Behav* 2007 Aug;10(4):536-544. [doi: [10.1089/cpb.2007.9993](#)]
14. Gunst M, De Meyere I, Willems H, Schoenmakers B. Effect of exergaming on wellbeing of residents in a nursing home: a single blinded intervention study. *Aging Clin Exp Res* 2022 Jan;34(1):151-157. [doi: [10.1007/s40520-021-01903-1](#)] [Medline: [34156650](#)]
15. Hoffman HG, Richards TL, Van Oostrom T, et al. The analgesic effects of opioids and immersive virtual reality distraction: evidence from subjective and functional brain imaging assessments. *Anesth Analg* 2007 Dec;105(6):1776-1783. [doi: [10.1213/01.ane.0000270205.45146.db](#)] [Medline: [18042882](#)]

16. Beltran-Alacreu H, Navarro-Fernández G, Godia-Lledó D, et al. A serious game for performing task-oriented cervical exercises among older adult patients with chronic neck pain: development, suitability, and crossover pilot study. *JMIR Serious Games* 2022 Feb 1;10(1):e31404. [doi: [10.2196/31404](https://doi.org/10.2196/31404)] [Medline: [35103608](https://pubmed.ncbi.nlm.nih.gov/35103608/)]
17. Peña J, Koebner I, Weisman W. Using digital art and attachment priming in a web-based serious game to reduce pain and social disconnection in individuals with chronic pain and loneliness: randomized controlled trial. *JMIR Serious Games* 2024 Nov 27;12:e52294. [doi: [10.2196/52294](https://doi.org/10.2196/52294)] [Medline: [39602564](https://pubmed.ncbi.nlm.nih.gov/39602564/)]
18. Maskeliūnas R, Damaševičius R, Kulikajevas A, Marley J, Larsson C. Evaluation of MyRelief serious game for better self-management of health behaviour strategies on chronic low-back pain. *Informatics (MDPI)* 2022;9(2):40. [doi: [10.3390/informatics9020040](https://doi.org/10.3390/informatics9020040)]
19. Ijaz K, Wang Y, Milne D, Calvo RA. Competitive vs affiliative design of immersive VR exergames. : Springer; 2016 Presented at: Serious Games: Second Joint International Conference, JCSG 2016; Sep 26-27, 2016 p. 140-150. [doi: [10.1007/978-3-319-45841-0_13](https://doi.org/10.1007/978-3-319-45841-0_13)]
20. Tuah NM, Ahmedy F, Gani A, Yong LN. A survey on gamification for health rehabilitation care: applications, opportunities, and open challenges. *Information* 2021;12(2):91. [doi: [10.3390/info12020091](https://doi.org/10.3390/info12020091)]
21. Tessarolo F, Petsani D, Conotter V, et al. Developing ambient assisted living technologies exploiting potential of user-centred co-creation and agile methodology: the CAPTAIN project experience. *J Ambient Intell Human Comput* 2024 May;15(5):2661-2676. [doi: [10.1007/s12652-021-03649-0](https://doi.org/10.1007/s12652-021-03649-0)]
22. iPROLEPSIS. URL: <https://www.iprolepsis.eu> [accessed 2025-03-09]
23. Hadjileontiadis LJ, Charisis V, Hadjidimitriou S, et al. European advances in digital rheumatology: explainable insights and personalized digital health tools for psoriatic arthritis. *EClinicalMedicine* 2025 Jun;84:103243. [doi: [10.1016/j.eclinm.2025.103243](https://doi.org/10.1016/j.eclinm.2025.103243)] [Medline: [40496877](https://pubmed.ncbi.nlm.nih.gov/40496877/)]
24. Ramalho B, Vicente M, Escobar H, et al. Developing exergames for psoriatic arthritis using agile storyboarding and game design processes. 2024 Jun 26 Presented at: PETRA '24. [doi: [10.1145/3652037.3663939](https://doi.org/10.1145/3652037.3663939)]
25. Ramalho B, Vicente M, Escobar H, et al. Developing sensorimotor art games for psoriatic arthritis using agile storyboarding and game co-design processes. 2024 Nov 13 Presented at: DSAI 2024. [doi: [10.1145/3696593.3696603](https://doi.org/10.1145/3696593.3696603)]
26. Ramalho B, Vicente M, Escobar H, et al. “The kite” breathing serious game: agile co-design for psoriatic arthritis. 2024 Presented at: 2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON). [doi: [10.1109/MELECON56669.2024.10608739](https://doi.org/10.1109/MELECON56669.2024.10608739)]
27. Knapp, J, Zeratsky J, Kowitz B. *Sprint: How to Solve Big Problems and Test New Ideas in Just Five Days*: Simon and Schuster; 2016.
28. Craft B, Cairns P. *Sketching sketching: outlines of a collaborative design method*. : BCS Learning and Development; 2009.
29. JW C, J DC. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*: SAGE; 2018.
30. Brooke J. *SUS: a quick and dirty usability scale*. : Digital Equipment Corporation; 1986.
31. IPROLEPSIS NoPain Games SUS questionnaire. Google Docs. URL: <https://forms.gle/hXbaMKPeaRQEfAEJ9> [accessed 2026-01-12]
32. Reise S, Moore T. Item response theory. In: Cooper H, Coutanche M, McMullen L, Panter A, Rindskopf D, Sher K, editors. *Handbook of Research Methods in Psychology: Foundations, Planning, Measures, and Psychometrics*: American Psychological Association; 2023:809-835. [doi: [10.1037/0000318-037](https://doi.org/10.1037/0000318-037)]
33. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
34. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* 2012;22(3):276-282. [doi: [10.11613/BM.2012.031](https://doi.org/10.11613/BM.2012.031)] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
35. Birnbaum A. Some latent trait models and their use in inferring an examinee's ability. In: Lord F, Novick M, editors. *Statistical Theories of Mental Test Scores* 1968:397-479.
36. 40 years of Simon, the electronic game that never stops reinventing itself. *Fast Company*. URL: <https://www.fastcompany.com/40491956> [accessed 2025-02-10]
37. Abd-Alrazaq A, Alhuwail D, Al-Jafar E, et al. The effectiveness of serious games in improving memory among older adults with cognitive impairment: systematic review and meta-analysis. *JMIR Serious Games* 2022 Aug 9;10(3):e35202. [doi: [10.2196/35202](https://doi.org/10.2196/35202)] [Medline: [35943792](https://pubmed.ncbi.nlm.nih.gov/35943792/)]
38. Lopes JC, Lopes RP. A review of dynamic difficulty adjustment methods for serious games. Presented at: Optimization, Learning Algorithms and Applications; Oct 24-25, 2022. [doi: [10.1007/978-3-031-23236-7_11](https://doi.org/10.1007/978-3-031-23236-7_11)]
39. Gobel S, Hardy S, Wendel V, Mehm F, Steinmetz R. Serious games for health: personalized exergames. Presented at: MM '10: ACM Multimedia Conference; Oct 25-29, 2010. [doi: [10.1145/1873951.1874316](https://doi.org/10.1145/1873951.1874316)]
40. Nunes F, Rato Grego P, Araújo R, Silva PA. Self-report user interfaces for patients with rheumatic and musculoskeletal diseases: app review and usability experiments with mobile user interface components. *Comput Graph* 2023 Dec;117:61-72. [doi: [10.1016/j.cag.2023.10.009](https://doi.org/10.1016/j.cag.2023.10.009)]
41. Armstrong-Wood R, Messiou C, Kite A, et al. Smartphone accessibility: understanding the lived experience of users with cervical spinal cord injuries. *Disabil Rehabil Assist Technol* 2024 May;19(4):1434-1445. [doi: [10.1080/17483107.2023.2192246](https://doi.org/10.1080/17483107.2023.2192246)] [Medline: [37010939](https://pubmed.ncbi.nlm.nih.gov/37010939/)]

42. Hadjileontiadou S, Dias SB, Hadjileontiadis L. A conceptual framework (2D-ME) for explaining self-first and self-third person views of prototyping dynamics in serious games design: experimental case study. *JMIR Serious Games* 2023 Apr 24;11:e41824. [doi: [10.2196/41824](https://doi.org/10.2196/41824)] [Medline: [37093627](https://pubmed.ncbi.nlm.nih.gov/37093627/)]
43. Hodent C. *The Gamer's Brain: How Neuroscience and UX Can Impact Video Game Design*: CRC Press; 2017. [doi: [10.1201/9781315154725](https://doi.org/10.1201/9781315154725)]
44. Smith ET, Basak C. A game-factors approach to cognitive benefits from video-game training: a meta-analysis. *PLOS ONE* 2023;18(8):e0285925. [doi: [10.1371/journal.pone.0285925](https://doi.org/10.1371/journal.pone.0285925)] [Medline: [37531408](https://pubmed.ncbi.nlm.nih.gov/37531408/)]
45. Csikszentmihalyi M, Csikszentmihalyi IS. *Optimal experience: psychological studies of flow in consciousness*. : Cambridge University Press; 1992.
46. Pinge A, Gad V, Jaisighani D, Ghosh S, Sen S. Detection and monitoring of stress using wearables: a systematic review. *Front Comput Sci* 2024;6. [doi: [10.3389/fcomp.2024.1478851](https://doi.org/10.3389/fcomp.2024.1478851)]
47. Krause-Jüttler G, Weitz J, Bork U. Interdisciplinary collaborations in digital health research: mixed methods case study. *JMIR Hum Factors* 2022 May 4;9(2):e36579. [doi: [10.2196/36579](https://doi.org/10.2196/36579)] [Medline: [35507400](https://pubmed.ncbi.nlm.nih.gov/35507400/)]
48. Kilfoy A, Hsu TCC, Stockton-Powdrell C, Whelan P, Chu CH, Jibb L. An umbrella review on how digital health intervention co-design is conducted and described. *NPJ Digit Med* 2024 Dec 23;7(1):374. [doi: [10.1038/s41746-024-01385-1](https://doi.org/10.1038/s41746-024-01385-1)] [Medline: [39715947](https://pubmed.ncbi.nlm.nih.gov/39715947/)]
49. Oham C, Ejike OG. Creativity and collaboration in creative industries: proposing a conceptual model for enhanced team dynamics. *Magna Sci Adv Res Rev* 2024;12(1):185-194. [doi: [10.30574/msarr.2024.12.1.0172](https://doi.org/10.30574/msarr.2024.12.1.0172)]
50. Vugts MA, Zedlitz AM, Joosen MC, Vrijhoef HJ. Serious gaming during multidisciplinary rehabilitation for patients with chronic pain or fatigue symptoms: mixed methods design of a realist process evaluation. *J Med Internet Res* 2020 Mar 9;22(3):e14766. [doi: [10.2196/14766](https://doi.org/10.2196/14766)] [Medline: [32149720](https://pubmed.ncbi.nlm.nih.gov/32149720/)]
51. Damaševičius R, Maskeliūnas R, Blažauskas T. Serious games and gamification in healthcare: a meta-review. *Information* 2023;14(2):105. [doi: [10.3390/info14020105](https://doi.org/10.3390/info14020105)]
52. Lane K, Majkowski R, Gruber J, et al. Using gamification to enhance clinical trial start-up activities. *J Clin Trans Sci* 2022;6(1):e75. [doi: [10.1017/cts.2022.405](https://doi.org/10.1017/cts.2022.405)]
53. Fisher N, Kulshreshtha AK. Exploring dynamic difficulty adjustment methods for video games. *Virtual Worlds* ;3(2):230-255. [doi: [10.3390/virtualworlds3020012](https://doi.org/10.3390/virtualworlds3020012)]
54. Waller S, Bradley M, Hosking I, Clarkson PJ. Making the case for inclusive design. *Appl Ergon* 2015 Jan;46:297-303. [doi: [10.1016/j.apergo.2013.03.012](https://doi.org/10.1016/j.apergo.2013.03.012)] [Medline: [23538129](https://pubmed.ncbi.nlm.nih.gov/23538129/)]

Abbreviations

- AI:** artificial intelligence
AI-PGS: Artificial Intelligence-Personalized Game Suite
ICC: item characteristic curve
PsA: psoriatic arthritis
SUS: System Usability Scale
UI: user interface
VR: virtual reality

Edited by S Brini; submitted 28.Mar.2025; peer-reviewed by E Benitez-Guerrero, T Adeel; accepted 11.Nov.2025; published 30.Jan.2026.

Please cite as:

Ramalho B, Gomes S, Silva Vicente M, Magalhães F, Gonçalves Costa R, Gama S, Charisis V, Hadjileontiadis L, B Dias S, iPROLEPSIS Consortium

Co-Designing Mobile Serious Games to Support Patients With Psoriatic Arthritis and Chronic Pain: Mixed Methods Study
JMIR Serious Games 2026;14:e75072

URL: <https://games.jmir.org/2026/1/e75072>

doi: [10.2196/75072](https://doi.org/10.2196/75072)

© Bárbara Ramalho, Samuel Gomes, Marta Silva Vicente, Filipa Magalhães, Rodolfo Gonçalves Costa, Sandra Gama, Vasileios Charisis, Leontios Hadjileontiadis, Sofia B Dias, iPROLEPSIS Consortium. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 30.Jan.2026. 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.

Original Paper

Efficacy and Safety of a Video Game–Like Digital Therapy Intervention for Chinese Children With Attention-Deficit/Hyperactivity Disorder: Single-Arm, Open-Label Pre-Post Study

Mengyi Bao^{1*}, MM; Shang Feng^{2*}, PhD; Jiaheng Wang³, PhD; Jia Ye⁴, MM; Jiangping Wang¹, MD; Wenyu Li², BEng; Kewen Jiang¹, MD; Lin Yao^{3,5,6}, PhD

¹Children's Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

²SDO Digital Therapeutics Innovation Center, Shanghai, Shanghai, China

³the MOE Frontiers Science Center for Brain and Brain-Machine Integration, the College of Computer Science, and the College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, Zhejiang, China

⁴Shengzhou People's Hospital (the First Affiliated Hospital of Zhejiang University Shengzhou Branch), Shengzhou, Zhejiang, China

⁵Nanhu Brain-Computer Interface Institute, Hangzhou, Zhejiang, China

⁶Affiliated Mental Health Center and Hangzhou Seventh People's Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

*these authors contributed equally

Corresponding Author:

Lin Yao, PhD

the MOE Frontiers Science Center for Brain and Brain-Machine Integration, the College of Computer Science, and the College of Biomedical Engineering and Instrument Science

Zhejiang University

No. 866, Yuhangtang Road, Xihu District (Post code 310000)

Hangzhou, Zhejiang,

China

Phone: 86 0571 87952838

Email: lin.yao@zju.edu.cn

Abstract

Background: The digital therapy of attention-deficit/hyperactivity disorder (ADHD) based on a “self-adaptive multitasking training paradigm” has been developed to improve the cognitive functional impairments and attention deficits of children with ADHD. However, the efficacy and safety of such treatment for Chinese patients remain untested.

Objective: This study aimed to preliminarily evaluate the actual intervention effects of a video game–like training software (ADHD-DTx) for children with ADHD aged 6–12 years as the first nationally certified digital therapeutics medical device for ADHD in China. We performed a single-arm, open-label efficacy and safety study.

Methods: This is a single-arm, open-label, pre-post efficacy and safety study. A total of 97 participants were included in the analysis. Participants received digital therapy (ADHD-DTx) and basic behavioral parent training for 4 weeks (25 min/day, ≥ 5 times/week) without medication. The efficacy outcomes included the Test of Variables of Attention (TOVA), Swanson, Nolan, and Pelham Questionnaire, version 4 (SNAP-IV), Weiss Functional Impairment Rating Scale (WFIRS), and Conner's Parent Symptom Questionnaire (PSQ). Safety-related events were monitored during and after the trial.

Results: From day 0 (baseline) to day 28, the population TOVA Attention Performance Index exhibited statistically significant improvement (from mean -4.15 , SE of the mean [SEM] 0.32 to mean -1.70 , SEM 0.30 ; $t_{94}=-8.78$; $n=95$; $P<.001$); the population total, inattention (AD), hyperactivity/impulsivity (HD), and oppositional defiant disorder (ODD) scores of SNAP-IV all significantly improved (total: from mean 1.33 , SEM 0.05 to mean 1.09 , SEM 0.05 ; $t_{96}=5.32$; $P<.001$; AD: from mean 1.71 , SEM 0.06 to mean 1.44 , SEM 0.06 ; $t_{96}=4.44$; $P<.001$; HD: from mean 1.38 , SEM 0.07 to mean 1.05 , SEM 0.06 ; $t_{96}=5.96$; $P<.001$; ODD: mean 0.84 , SEM 0.05 to mean 0.75 , SEM 0.05 ; $Z=2.47$; $P=.03$; $n=97$); for WFIRS results, domains of “family” and “social activities” showed significant population improvement (family: from mean 0.75 , SEM 0.05 to mean 0.65 , SEM 0.04 ; $Z=2.80$; $P=.01$; social activities: from mean 0.56 , SEM 0.05 to mean 0.45 , SEM 0.05 ; $Z=2.91$; $P=.01$; $n=97$); for PSQ results, domains of “learning problem,”

“psychosomatic problem,” “impulsivity-hyperactivity,” and “hyperactivity index” showed significant improvement (learning problem: from mean 1.72, SEM 0.06 to mean 1.57, SEM 0.06; $Z=2.42$; $P=.03$; psychosomatic problem: from mean 0.40, SEM 0.03 to mean 0.32, SEM 0.03; $Z=2.66$; $P=.02$; impulsivity-hyperactivity: from mean 0.94, SEM 0.06 to mean 0.80, SEM 0.06; $Z=2.49$; $P=.03$; hyperactivity index: from mean 1.06, SEM 0.05 to mean 0.92, SEM 0.05; $Z=2.90$; $P=.01$; $n=97$). No device-related adverse event or severe adverse event was observed or reported during or after the intervention.

Conclusions: This study preliminarily suggested the significant improvements of ADHD symptoms and attention function after 4 weeks of ADHD-DTx digital therapy combining basic behavioral parent training with satisfying safety outcomes.

(*JMIR Serious Games 2026;14:e76114*) doi:[10.2196/76114](https://doi.org/10.2196/76114)

KEYWORDS

ADHD; attention; attention-deficit/hyperactivity disorder; cognitive training; digital therapy; rehabilitation

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder that commonly occurs in childhood and is characterized by inattention, hyperactivity, and impulsivity. The primary treatment for ADHD is pharmacotherapy, using methylphenidate, dexamphetamine, or atomoxetine as the first-line drugs [1]. However, pharmacotherapy for ADHD could cause adverse drug reactions such as digestive system issues, irritability, palpitations, and headaches, and more importantly, pharmacotherapy does not target the core functional deficits of patients with ADHD [2]. A recent study also suggested that beginning treatment with behavioral intervention may produce better outcomes overall than beginning treatment with medication [3]. Therefore, many patients with ADHD and their parents are looking forward to alternative interventions [4] like behavior therapy, neurofeedback, counseling, and, more recently, digital therapy [5].

The digital therapy of ADHD has been developed to improve the cognitive functional impairments and attention-control deficits of children with ADHD. In 2020, the US Food and Drug Administration (FDA) approved EndeavorRx (AKL-T01), the first video game-like training software for children with ADHD aged 8-12 years. Growing evidence has suggested that digital therapy could provide a safe and effective intervention to improve functional performance (such as attention, working memory, etc) and problematic behavior of children with ADHD, with minimal risk of adverse events (AEs) compared to pharmacotherapy [6-9]. In 2023, the Chinese National Medical Products Administration (NMPA) approved the “Attention Enhancement Training Software” (ADHD-DTx), a video game-like training software for children with ADHD aged 6-12 years, as the first nationally certified digital therapeutics medical device for ADHD in China.

The neurophysiological mechanisms of EndeavorRx and ADHD-DTx are both based on the self-adaptive multitasking training paradigm (NeuroRacer), which consists of 2 tasks: the “driving task” (sustained attention) and the go/no-go “sign task” (signal detection, attention-shifting, and inhibition control). Multiple studies have revealed that this training paradigm could improve participants’ cognitive control abilities (enhanced sustained attention and working memory) after 1-month training [6-8,10], as well as increase the frontal midline theta (FM) power, the neuromarker of sustained attention and cognitive

control [9,10]. Therefore, this paradigm has been used for the intervention of ADHD in a video game-like form to improve compliance in children. Besides the “driving task” and “sign task,” ADHD-DTx also included a third task: the “digit cancellation task” (a widely used attention assessment and training method in clinical practice) [11] to further enhance the training effect of attention function.

To preliminarily evaluate the actual intervention effects of ADHD-DTx in Chinese children with ADHD aged 6-12 years, we performed a single-arm, open-label efficacy and safety study in the Children’s Hospital of Zhejiang University School of Medicine (Hangzhou, China) in 2021. This study provided valuable efficacy and safety data of ADHD-DTx (and ADHD digital therapy based on the self-adaptive multitasking training paradigm) for the first time in Chinese children with ADHD. The efficacy data suggested by this pilot study provided a critical contribution to the design of the following randomized, double-blinded, parallel-controlled clinical trials, which were conducted during 2022 and 2023 (Feng S, PhD, unpublished data, 2025). The Chinese NMPA approval of ADHD-DTx (as the first Chinese ADHD digital therapy medical device) was based on the results of a key Good Clinical Practice clinical trial conducted during 2022-2023 (Feng S, PhD, unpublished data, 2025).

All participants had to be off any ADHD medication and without other significant comorbid psychiatric diagnoses. Included participants were treated with ADHD-DTx therapy and basic behavioral parent training (BPT; positive reinforcement training at home, required according to the ethical consideration of the regulatory agency) [12,13]. Efficacy outcomes included the computerized attention test: Test of Variables of Attention (TOVA) [14,15] and classic scales: (1) the Swanson, Nolan, and Pelham Questionnaire, version 4 (SNAP-IV) [16-18]; (2) the Weiss Functional Impairment Rating Scale (WFIRS) [19]; and (3) the Conner’s Parent Symptom Questionnaire (PSQ) [20,21]. Safety outcomes included the proportions of device-related AEs or severe AEs.

Methods

Overview

This study was a single-arm, open-label study in children (aged 6-12 years) with a confirmed diagnosis of ADHD (as per the DSM-5 [*Diagnostic and Statistical Manual of Mental Disorders* {Fifth Edition}]) in the Children’s Hospital of Zhejiang

University School of Medicine (Hangzhou, China) in 2021. Participants had to be off any ADHD medication and not present other significant comorbid psychiatric diagnoses. All participants had an IQ score of ≥ 80 (per the Wechsler Intelligence Test) and an Attention-Deficit/Hyperactivity Disorder Rating Scale IV (ADHD-RS-IV) total score of > 28 .

Study Design

The study planned to enroll about 100 participants with a confirmed diagnosis of ADHD. Participants received digital therapy (ADHD-DTx, a video game-like training software running on an Android tablet) and basic BPT (positive reinforcement training at home) for 4 weeks (25 min/day, ≥ 5 times/week). Efficacy outcomes were measured on day 0 (baseline visit) and day 28 (after 4-week treatment). Safety-related events were monitored during and after the trial.

Participants

Eligible patients were male or female children aged 6-12 years with a confirmed diagnosis of ADHD (as per the DSM-V). Participants had to be off any ADHD medication (for at least 4 weeks before the baseline visit) and not present other significant comorbid psychiatric diagnoses. All participants had an IQ score of ≥ 80 (per the Wechsler Intelligence Test) and an ADHD-RS-IV total score of ≥ 28 . Complete inclusion and exclusion criteria are in [Multimedia Appendix 1](#).

Procedures

The ADHD-DTx intervention was preinstalled on Huawei MatePad tablets (Huawei). At the baseline visit, eligible patients were instructed to use ADHD-DTx for about 10 minutes while a study coordinator monitored the session to ensure that patients could follow the rules of ADHD-DTx. Patients were further assessed by ADHD-related scales (symptoms and impairments) and TOVA (attention functioning) at the baseline visit.

Afterward, patients received the ADHD-DTx digital treatment and basic BPT at home for 4 weeks (25 min/day, ≥ 5 times/week). Basic BPT was required according to the ethical considerations of the regulatory agency. The daily training task consisted of five 4- to 5-minute multitasking missions (total time on task was about 25 minutes). Compliance was monitored remotely using the network by investigators, and daily training reminders were sent to patients' caregivers manually or automatically. In accordance with a strict, predefined protocol, any participant who failed to complete the required training (25 min/day, ≥ 5 times/week) and persistently ignored reminders was classified as noncompliant and withdrawn from the trial.

ADHD-DTx is a digital therapeutic that uses a proprietary algorithm designed to improve attention control by training interference management (multitasking). ADHD-DTx mechanisms have been described previously. In brief, users multitask by responding to a perceptual discrimination targeting task and a simultaneous sensory motor navigation task. Users advance by reducing interference costs (closing the performance gap between multitasking and single-tasking), and real-time and periodic recalibration occurs to maintain an optimal difficulty level.

BPT consisted of parent-training courses (focusing on teaching parents specialized child management techniques primarily involving contingency management, such as behavior management principles, parental attending skills and home token system), aerobic exercises (such as jogging, swimming, or rope-skipping 40 min/day, ≥ 4 times/week), listening-retelling training (the child was asked to accurately retell sentences presented orally by parent, and the sentences were gradually made longer and more detailed to train the child's auditory attention and memory span), and reading-aloud training (the child was asked to read texts aloud to maintain focus and reduce mind-wandering, thereby promoting the child's sustained attention) [22-24]. BPT is a regular basic treatment for all children with ADHD in clinical practice.

Safety-related events such as vomiting, dizziness, headache, palpitation, addiction, frustration, eye discomfort, and other similar symptoms were monitored during and after the trial, as described: (1) during the 4-week intervention period, all participants' parents were contacted daily by a research assistant using online social software to collect safety-related events; and (2) after the intervention period, all participants' parents were contacted monthly by a research assistant using the telephone for 3 months to collect long-term safety-related events.

Outcomes

To effectively evaluate the impact of treatment on the core symptoms, problematic behaviors, and functional deficits of children with ADHD, and also ensure the reliability, validity, and objectivity of data, we chose both the computerized attention test and classic scales as the efficacy outcomes.

The primary end point was the improvement of attentional functioning as measured by TOVA from baseline to day 28. TOVA is a computerized, objective test of attentional functioning and has been globally used in clinical and academic institutions [22-24]. The outcomes of TOVA could objectively reflect the functional training effect of ADHD-DTx and provide essential information about the efficacy of the intervention.

Secondary end points included the improvements of classic scales that assess ADHD-related symptoms, functional impairments, and problematic behaviors from baseline to day 28. The SNAP-IV is widely used as the key assessment of ADHD core symptoms, including 3 subsets: (1) inattention (AD), (2) hyperactivity/impulsivity (HD), and (3) oppositional defiant disorder (ODD) [16,17]. WFIRS is a multidimensional, ADHD-specific, functional impairment assessment scale, including 6 domains: family, school, life skills, child's self-concept, social activities, and risky activities [19]. The PSQ has been widely used to assess problematic behaviors related to ADHD, including 6 domains: conduct problem, learning problem, psychosomatic problem, impulsivity-hyperactivity, anxiety, and hyperactivity index [20,21]. The selection of classic, widely validated scales could provide rich information about the behavioral symptoms, and importantly, the outcomes of these scales could reflect the influence of daily life by the training of ADHD-DTx, which is crucial for the rehabilitation of children with ADHD. More detailed descriptions of the outcome measurements can be found in [Multimedia Appendix 2](#).

Statistical Analysis

All analyses were performed according to a prespecified statistical analysis plan. Unless otherwise indicated, statistical comparisons used a 2-tailed significance test evaluated at the 95% level of confidence. All analyses were conducted using a complete case analysis. In no situation were missing data to be imputed. Student *t* test was performed only if the analyzed data passed the normality test (Shapiro-Wilk test); otherwise, the Wilcoxon nonparametric test would be used. The chi-square test was used for the statistical inference for counting data. Multiple comparisons corrections (using the false discovery rate method) were performed for all the secondary efficacy end points [25].

The primary efficacy end point for each participant was the change in the TOVA Attention Performance Index (API) from baseline to day 28, defined as the score on day 28 minus the score at baseline. Missing data were not imputed. Participants with missing data either on day 0 or day 28 would be excluded from the paired significance analysis. Unless otherwise indicated, the results of the efficacy analysis were summarized as mean (SE of the mean [SEM]). Significance was assessed with a 2-sided paired *t* test evaluated at the 95% level of confidence.

The following secondary efficacy end points were tested using the same technique outlined for the primary efficacy analysis: (1) change in SNAP-IV from baseline to day 28, (2) change in WFIRS from baseline to day 28, and (3) change in PSQ from baseline to day 28.

Ethical Considerations

The study was conducted in accordance with the International Conference on Harmonisation Regulations and was approved by the Institutional Review Board of Children's Hospital of Zhejiang University School of Medicine. All participants and their caregivers provided written informed consent prior to any study activities being conducted. The data of all participants were anonymous throughout the study. All participants received

compensation for the examination fees. The study was subject to independent supervision by regulatory agencies (the Institutional Review Board and the Office of Clinical Trial Institution of Children's Hospital of Zhejiang University School of Medicine) throughout the entire process. Authors who were employed by SDO Digital Therapeutics (the developer of ADHD-DTx) did not participate in any data collection process, and the data analysis results were checked and confirmed by authors from the Children's Hospital of Zhejiang University School of Medicine. To safeguard the design details and clinical parameters of the investigational product (ADHD-DTx) and mitigate potential commercial risks, the producer requested a delay in the clinical trial registration until after market approval was obtained from the NMPA of China. Following the product's market approval in 2023, the trial was submitted for registration with the Chinese Clinical Trial Registry, where it is currently pending review.

Results

Participants

A total of 114 participants were screened for inclusion in this study, with 110 meeting eligibility criteria (Figure 1). The sample size was determined comprehensively, referring to previous research using similar digital therapy for children with ADHD [8], the power of statistics, and the amount of available resources. The mean age of included participants was 7.78 (SD 1.14) years, and 90% (99/110) were male. The vast majority (106/110, 96.36%) of included participants were of the Han ethnic group, and other involved ethnic groups included She (2/110, 1.82%), Korean (1/110, 0.91%), and Tujia (1/110, 0.91%). The educational status of participants' parents: postgraduate (6/110, 5.45%), graduate or junior college (62/110, 56.36%), high school or vocational school (28/110, 25.45%), middle school (12/110, 10.91%), primary school, and below (2/110, 1.82%). Demographic characteristics of included participants are listed in Table 1.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of the single-arm, open-label study of children with attention-deficit/hyperactivity disorder in China during 2021. TOVA: Test of Variables of Attention.

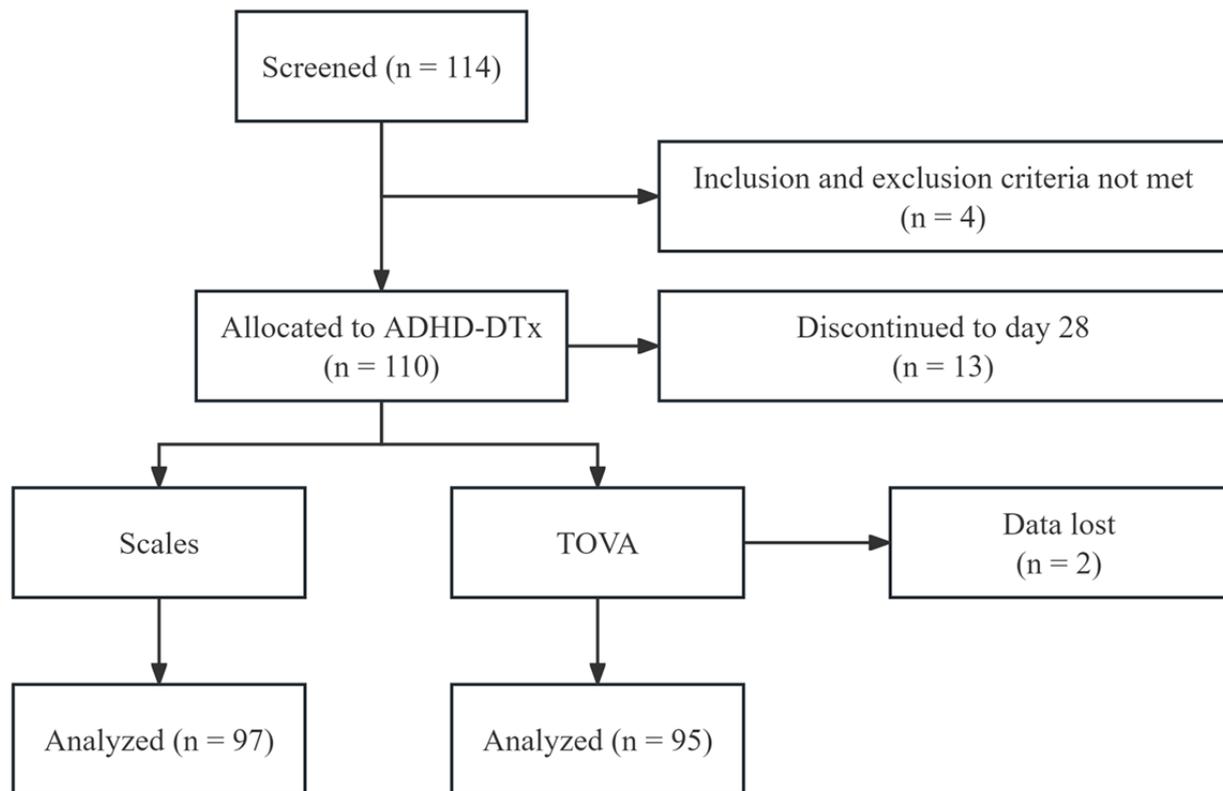


Table 1. Demographic characteristics of included participants.

Demographic	Value (n=110)
Age (years), mean (SD)	7.78 (1.14)
Gender, n (%)	
Male	99 (90)
Female	11 (10)
Ethnicity, n (%)	
Han	106 (96.36)
She	2 (1.82)
Korean	1 (0.91)
Tujia	1 (0.91)
Parental education level, n (%)	
Postgraduate	6 (5.45)
Graduate or junior college	62 (56.36)
High school or vocational school	28 (25.45)
Middle school	12 (10.91)
Primary school and below	2 (1.82)

Overall, 97/110 (88%) participants completed the study, 13 participants did not complete the study (lost to follow-up or noncompliance), and the TOVA data of 2 participants were lost due to device failure. All 97 participants who completed the study adhered strictly to the training protocol (25 min/day, ≥ 5

times/week for 4 weeks), and the average number of actual training days completed per participant was 20. Therefore, we finally collected and analyzed 97 participants' scales data, as well as 95 participants' TOVA data. All participants received

ADHD-DTx digital therapy (combined with basic BPT) in 2021 and remained off medication during the treatment period.

Efficacy Outcomes

In this study, efficacy outcomes were measured using the computerized attention test (TOVA) and classic scales:

SNAP-IV, WFIRS, and PSQ. After a 4-week intervention, we found that the treatment significantly improved ADHD-related symptoms and impairments (all detailed statistical results are present in [Table 2](#)).

Table 2. Summary of efficacy outcomes during the 4-week treatment.

Efficacy outcome	Participants, n	Day 0, mean (SEM)	Day 28, mean (SEM)	P value	Test statistic
TOVA^a-API^b	95	-4.15 (0.32)	-1.70 (0.30)	<.001	-8.78 (94) ^c
Younger group	69	-4.15 (0.35)	-1.59 (0.35)	<.001	-7.90 (68) ^c
Older group	26	-4.15 (0.73)	-2.00 (0.59)	.001	-3.90 (25) ^c
Male group	84	-4.18 (0.34)	-1.74 (0.31)	<.001	-8.11 (83) ^c
Female group	11	-3.89 (1.08)	-1.38 (1.19)	.008	-3.31 (10) ^c
SNAP-IV^d total	97	1.33 (0.05)	1.09 (0.05)	<.001	5.32 (96) ^c
Younger group	70	1.30 (0.05)	1.09 (0.06)	<.001	4.44 (69) ^c
Older group	27	1.39 (0.11)	1.11 (0.09)	.007	2.91 (26) ^c
Male group	86	1.35 (0.05)	1.12 (0.05)	<.001	4.78 (85) ^c
Female group	11	1.14 (0.08)	0.87 (0.16)	.03	2.59 (10) ^c
SNAP-IV AD^e	97	1.71 (0.06)	1.44 (0.06)	<.001	4.44 (96) ^c
Younger group	70	1.72 (0.07)	1.42 (0.07)	<.001	4.77 (69) ^c
Older group	27	1.68 (0.11)	1.51 (0.12)	.23	1.22 (26) ^c
Male group	86	1.73 (0.06)	1.48 (0.07)	<.001	3.97 (85) ^c
Female group	11	1.57 (0.11)	1.20 (0.19)	.07	2.04 (10) ^c
SNAP-IV HD^f	97	1.38 (0.07)	1.05 (0.06)	<.001	5.96 (96) ^c
Younger group	70	1.34 (0.07)	1.07 (0.08)	<.001	4.53 (69) ^c
Older group	27	1.49 (0.14)	1.01 (0.09)	<.001	3.98 (26) ^c
Male group	86	1.42 (0.07)	1.07 (0.06)	<.001	5.69 (85) ^c
Female group	11	1.09 (0.19)	0.89 (0.25)	.09	1.91 (10) ^c
SNAP-IV ODD^g	97	0.84 (0.05)	0.75 (0.05)	.03	2.47 ^h
Younger group	70	0.80 (0.06)	0.74 (0.06)	.17	1.38 (69) ^c
Older group	27	0.95 (0.13)	0.76 (0.11)	.01	2.50 (26) ^c
Male group	86	0.86 (0.06)	0.78 (0.06)	.06	1.88 (85) ^c
Female group	11	0.70 (0.10)	0.47 (0.08)	.04	2.02 (10) ^c
WFIRSⁱ					
Family	97	0.75 (0.05)	0.65 (0.04)	.01	2.80 ^h
School	97	0.78 (0.05)	0.72 (0.04)	.14	1.55 ^h
Life skills	97	0.72 (0.04)	0.65 (0.04)	.09	1.85 ^h
Child's self-concept	97	0.65 (0.06)	0.59 (0.05)	.34	0.96 ^h
Social activities	97	0.56 (0.05)	0.45 (0.05)	.01	2.91 ^h
Risky activities	97	0.27 (0.03)	0.23 (0.03)	.14	1.54 ^h
PSQ^j					
Conduct problem	97	0.82 (0.05)	0.74 (0.05)	.06	2.04 ^h
Learning problem	97	1.72 (0.06)	1.57 (0.06)	.03	2.42 ^h
Psychosomatic problem	97	0.40 (0.03)	0.32 (0.03)	.02	2.66 ^h

Efficacy outcome	Participants, n	Day 0, mean (SEM)	Day 28, mean (SEM)	P value	Test statistic
Anxiety	97	0.36 (0.04)	0.34 (0.04)	.17	1.40 ^h
Impulsivity-hyperactivity	97	0.94 (0.06)	0.80 (0.06)	.03	2.49 ^h
Hyperactivity index	97	1.06 (0.05)	0.92 (0.05)	.01	2.90 ^h

^aTOVA: Test of Variables of Attention.

^bAPI: Attention Performance Index.

^ct test (df).

^dSNAP-IV: Swanson, Nolan, and Pelham Questionnaire, version 4.

^eAD: inattention subscale of SNAP-IV.

^fHD: hyperactivity/impulsivity subscale of SNAP-IV.

^gODD: oppositional defiant disorder subscale of SNAP-IV.

^hZ value of Wilcoxon nonparametric test.

ⁱWFIRS: Weiss Functional Impairment Rating Scale.

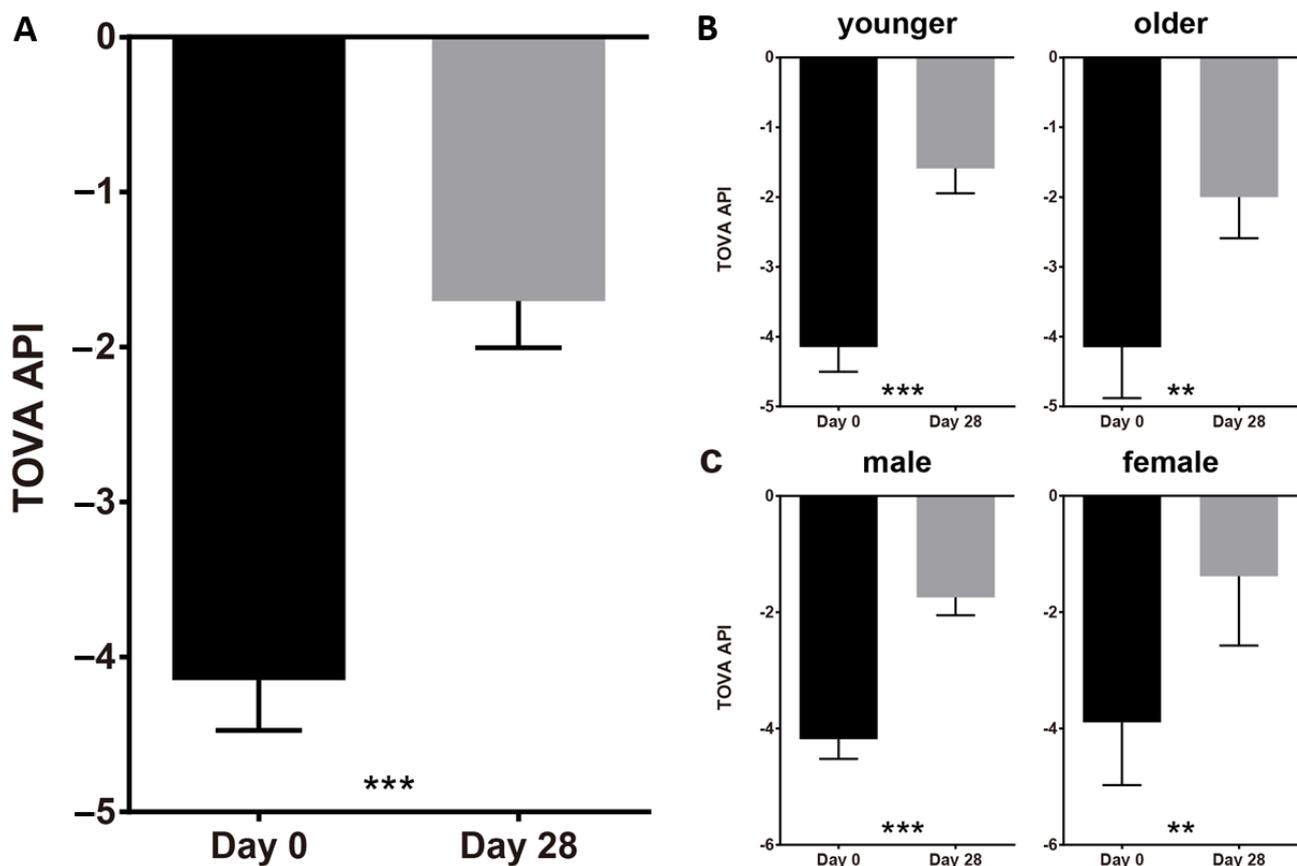
^jPSQ: Conner's Parent Symptom Questionnaire.

Primary End Point: TOVA-API

We chose TOVA (an objective measurement of attention function) as the primary measurement of intervention efficacy,

and the TOVA-API was chosen as the primary end point of this study (Figure 2).

Figure 2. Statistical results of the Test of Variables of Attention (TOVA) suggested significant improvement after 4 weeks of intervention. (A) TOVA Attention Performance Index (API) outcomes from day 0 to day 28; (B) TOVA-API of younger (aged 6-8 years) and older (aged 9-12 years) groups from day 0 to day 28; (C) TOVA-API of male and female groups from day 0 to day 28. Error bars indicate SEM. ** $P < .01$, *** $P < .001$.



The baseline data of TOVA-API passed the normality test ($n=95$; $P=.36$). From day 0 (baseline) to day 28, the population TOVA-API exhibited statistically significant improvement (Figure 2A, from mean -4.15 , SEM 0.32 to mean -1.70 , SEM 0.30 ; $T=-8.78$; $n=95$; $P<.001$), suggesting the efficacy of objective functional improvement of attention.

We conducted subgroup analysis of TOVA-API on age (Figure 2B; younger: 6-8 years; older: 9-12 years) and gender (Figure 2C; male and female) to investigate more detailed efficacy characteristics. We found that from day 0 (baseline) to day 28, both the younger and older groups exhibited statistically significant improvement (younger group: from mean -4.15 ,

SEM 0.35 to mean -1.59 , SEM 0.35; $T=-7.90$; $n=69$; $P<.001$; older group: from mean -4.15 , SEM 0.73 to mean -2.00 , SEM 0.59; $T=-3.90$; $n=26$; $P=.001$). The improvement of the younger group (pre-post difference= 2.56) slightly exceeded that of the older group (pre-post difference= 2.15), but no statistical difference was found ($P=.53$). As for the gender analysis, from day 0 (baseline) to day 28, both the male and female groups exhibited statistically significant improvement (male group: from mean -4.18 , SEM 0.34 to mean -1.74 , SEM 0.31; $T=-8.11$; $n=84$; $P<.001$; female group: from mean -3.89 , SEM 1.08 to mean -1.38 , SEM 1.19; $T=-3.31$; $n=11$; $P=.008$). The improvements of both groups showed no statistical difference ($P=.93$). We also analyzed the potential influence of parental education level (college degree or above: $n=56$; below college level: $n=39$) on TOVA-API and found no statistically significant difference ($P=.78$).

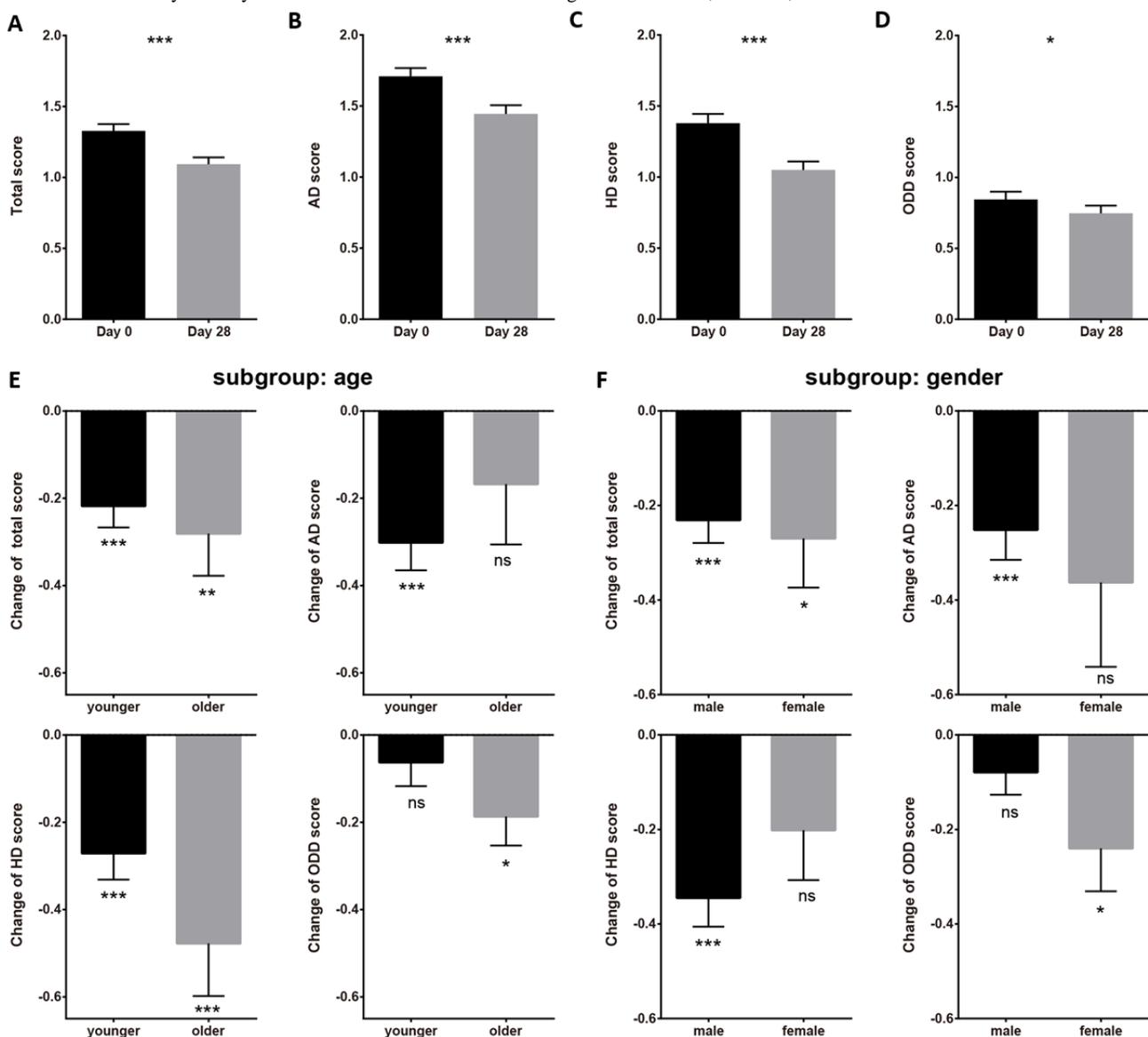
The above findings suggested that the 4-week intervention could significantly improve the attention function of children with ADHD (regardless of age or gender), as measured by TOVA-API.

Secondary End Points: SNAP-IV, WFIRS, and PSQ

We chose the results of classic clinical scales (SNAP-IV, WFIRS, and PSQ) as the secondary end points. Multiple comparisons corrections were performed for the statistical analysis results of the scale scores (see Methods section).

Using the SNAP-IV scale, we measured the treatment effect of ADHD core symptoms (Figure 3). The baseline data of the SNAP-IV total score, AD subscale, and HD subscale passed the normality test ($n=97$; $P_{\text{total}}=.19$; $P_{\text{AD}}=.09$; $P_{\text{HD}}=.25$), while that of the ODD subscale did not ($n=97$; $P_{\text{ODD}}<.001$). Therefore, we used the Student *t* test for total score, AD, and HD data, and the Wilcoxon nonparametric test for ODD data.

Figure 3. Statistical results of the Swanson, Nolan, and Pelham Questionnaire, version 4 (SNAP-IV), suggested significant improvement after 4 weeks of intervention. (A) SNAP-IV total score from day 0 to day 28; (B) SNAP-IV inattention (AD) score from day 0 to day 28; (C) SNAP-IV hyperactivity/impulsivity (HD) score from day 0 to day 28. (D) SNAP-IV oppositional defiant disorder (ODD) score from day 0 to day 28. (E) Subgroup analysis of age (younger: 6-8 years; older: 9-12 years) of SNAP-IV scores from day 0 to day 28. (F) Subgroup analysis of gender (male and female) of SNAP-IV scores from day 0 to day 28. Error bars indicate SEM. ns: nonsignificant. * $P<.05$, ** $P<.01$, *** $P<.001$.



From day 0 to day 28, the population total score of SNAP-IV significantly improved (descended) from mean 1.33, SEM 0.05 to mean 1.09, SEM 0.05 (Figure 3A; $T=5.32$; $n=97$; $P<.001$); the population AD score of SNAP-IV significantly improved (descended) from mean 1.71, SEM 0.06 to mean 1.44, SEM 0.06 (Figure 3B; $T=4.44$; $n=97$; $P<.001$); the population HD score of SNAP-IV significantly improved (descended) from mean 1.38, SEM 0.07 to mean 1.05, SEM 0.06 (Figure 3C; $T=5.96$; $n=97$; $P<.001$); the population ODD score of SNAP-IV significantly improved (descended) from mean 0.84, SEM 0.05 to mean 0.75, SEM 0.05 (Figure 3D; $Z=2.47$; $n=97$; $P=.03$). The above results suggested that the treatment could significantly improve the core symptoms (attention deficit and hyperactivity) of children with ADHD.

We also conducted subgroup analysis of SNAP-IV on age (Figure 3E; younger: 6-8 years; older: 9-12 years) and gender (Figure 3F; male and female). It was notable that due to the small sample size of the female group ($n=11$) and older group ($n=27$), the efficacy of statistical inference might be limited.

For the SNAP-IV total score, all the subgroups exhibited statistically significant improvement (younger group: from mean 1.30, SEM 0.05 to mean 1.09, SEM 0.06; $T=4.44$; $n=70$; $P<.001$; older group: from mean 1.39, SEM 0.11 to mean 1.11, SEM 0.09; $T=2.91$; $n=27$; $P=.007$; male group: from mean 1.35, SEM 0.05 to mean 1.12, SEM 0.05; $T=4.78$; $n=86$; $P<.001$; female group: from mean 1.14, SEM 0.08 to mean 0.87, SEM 0.16; $T=2.59$; $n=11$; $P=.03$ paired t test). No statistical difference was found between different age and gender groups (age: $P=.56$; gender: $P=.74$).

For the SNAP-IV AD score, only the younger group and male group exhibited statistically significant improvement (younger group: from mean 1.72, SEM 0.07 to mean 1.42, SEM 0.07; $T=4.77$; $n=70$; $P<.001$; male group: from mean 1.73, SEM 0.06 to mean 1.48, SEM 0.07; $T=3.97$; $n=86$; $P<.001$), while the older group and female group did not (older group: from mean

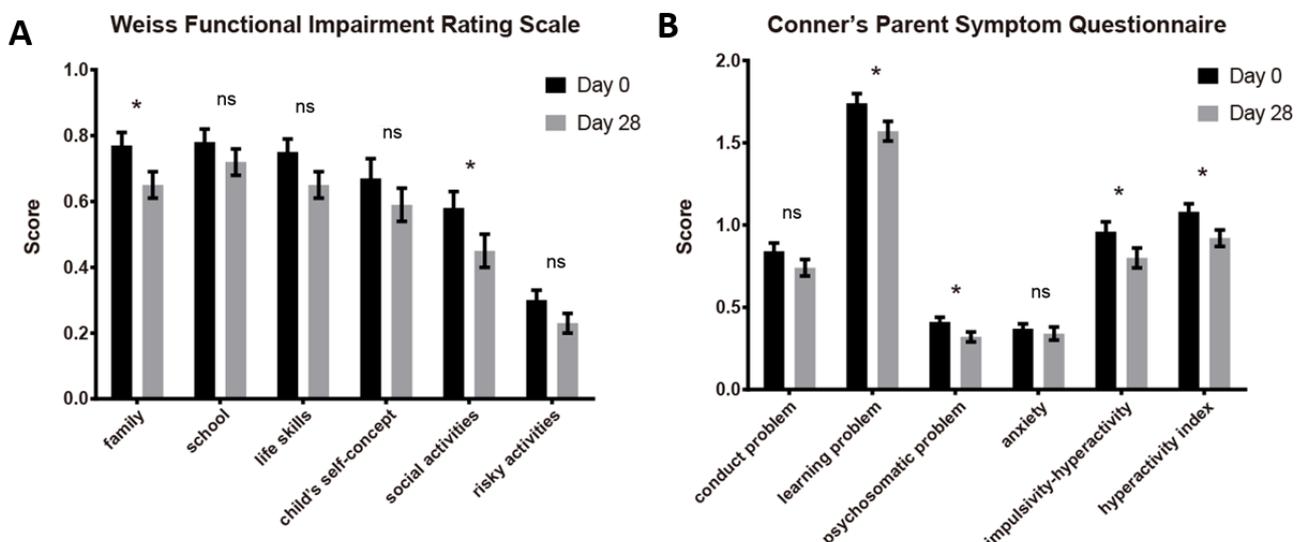
1.68, SEM 0.11 to mean 1.51, SEM 0.12; $T=1.22$; $n=27$; $P=.23$; female group: from mean 1.57, SEM 0.11 to mean 1.20, SEM 0.19; $T=2.04$; $n=11$; $P=.07$). No statistical difference was found between different age and gender groups (age: $P=.39$; gender: $P=.57$).

For the SNAP-IV HD score, the younger, older, and male groups exhibited statistically significant improvement (younger group: from mean 1.34, SEM 0.07 to mean 1.07, SEM 0.08; $T=4.53$; $n=70$; $P<.001$; older group: from mean 1.49, SEM 0.14 to mean 1.01, SEM 0.09; $T=3.98$; $n=27$; $P<.001$; male group: from mean 1.42, SEM 0.07 to mean 1.07, SEM 0.06; $T=5.69$; $n=86$; $P<.001$), while the female group did not ($n=11$; $P=.09$). No statistical difference was found between different age and gender groups (age: $P=.94$; gender: $P=.26$).

For the SNAP-IV ODD score, only the older group and female group exhibited statistically significant improvement (older group: from mean 0.95, SEM 0.13 to mean 0.76, SEM 0.11; $T=2.50$; $n=27$; $P=.01$; female group: from mean 0.70, SEM 0.10 to mean 0.47, SEM 0.08; $T=2.02$; $n=11$; $P=.04$), while the younger group and male group did not (younger group: $T=1.38$; $n=70$; $P=.17$; male group: $T=1.88$; $n=86$; $P=.06$). No statistical difference was found between different age and gender groups (age: $P=.17$; gender: $P=.22$).

The functional impairments of participants were measured using WFIRS (Figure 4A). WFIRS data did not pass the normality test; therefore, the Wilcoxon signed-rank test was used. We found that after 4 weeks of treatment, the impairments of the “family” and “social activities” domains were statistically significantly improved (family: from mean 0.75, SEM 0.05 to mean 0.65, SEM 0.04; $Z=2.80$; $P=.01$; social activities: from mean 0.56, SEM 0.05 to mean 0.45, SEM 0.05; $Z=2.91$; $n=97$; $P=.01$), while the domains of “school,” “life skills,” “child’s self-concept,” and “risky activities” exhibited nonsignificant improvement (school: $P=.14$; life skills: $P=.09$; child’s self-concept: $P=.34$; risky activities: $P=.14$; $n=97$).

Figure 4. Statistical results of the Weiss Functional Impairment Rating Scale and Conner’s Parent Symptom Questionnaire suggested significant improvement after 4 weeks of intervention. (A) Weiss Functional Impairment Rating Scale outcomes (scores) from day 0 to day 28. (B) Conner’s Parent Symptom Questionnaire outcomes (scores) from day 0 to day 28. Error bars indicate SEM. ns: nonsignificant. * $P<.05$, ** $P<.01$, *** $P<.001$.



We also analyzed the number of items scored ≥ 2 in WFIRS to investigate the treatment effect on severe functional impairments using the chi-square test. We found that the correlation between the intervention and the number of items scored ≥ 2 of “child’s self-concept” was statistically significant (Pearson $\chi^2_1=7.249$; $P=.007$). There were also almost significant correlations between the intervention and the number of items scored ≥ 2 of “family” and “life skills” (family: Pearson $\chi^2_1=3.183$; $P=.07$; life skills: Pearson $\chi^2_1=3.071$; $P=.08$). While there were no significant correlations between the intervention and the number of items scored ≥ 2 of “school,” “social activities,” and “risky activities” (school: Pearson $\chi^2_1=0.949$; $P=.33$; social activities: Pearson $\chi^2_1=0.899$; $P=.34$; risky activities: Pearson $\chi^2_1=0.519$; $P=.47$).

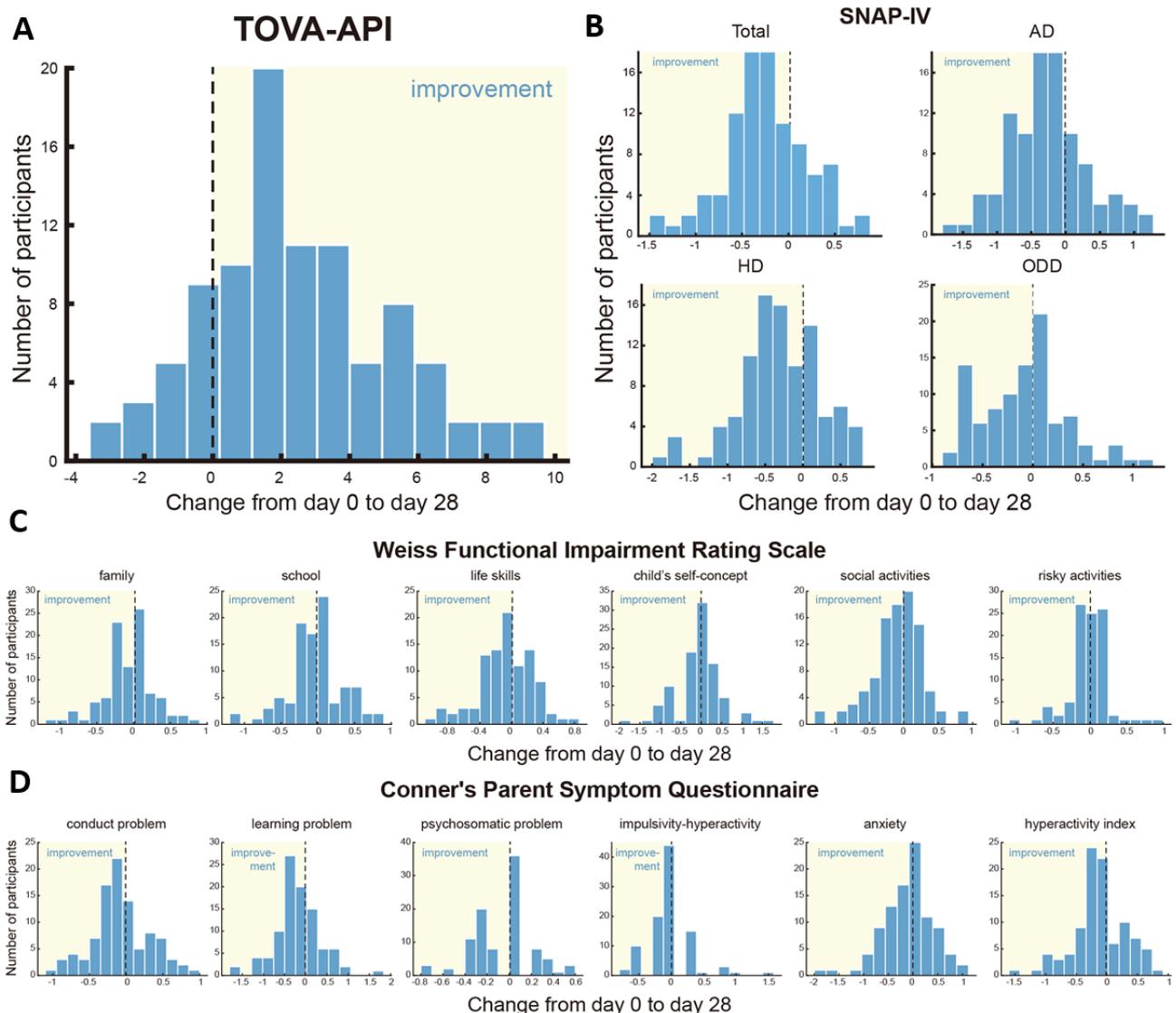
The problematic behaviors of participants were measured using PSQ (Figure 4B). PSQ data did not pass the normality test; therefore, the Wilcoxon signed-rank test was used. After 4 weeks treatment, the domains of “learning problem,” “psychosomatic problem,” “impulsivity-hyperactivity,” and “hyperactivity index” were statistically significantly improved (learning problem: from mean 1.72, SEM 0.06 to mean 1.57, SEM 0.06; $Z=2.42$; $P=.03$; psychosomatic problem: from mean 0.40, SEM 0.03 to mean 0.32, SEM 0.03; $Z=2.66$; $P=.02$; impulsivity-hyperactivity: from mean 0.94, SEM 0.06 to mean 0.80, SEM 0.06; $Z=2.49$; $P=.03$; hyperactivity index: from mean

1.06, SEM 0.05 to mean 0.92, SEM 0.05; $Z=2.90$; $P=.01$; $n=97$). While the domains of “conduct problem” and “anxiety” exhibited nonsignificant improvement (conduct problem: $P=.06$; anxiety: $P=.17$).

The above findings suggested that the 4-week intervention could significantly improve the core symptoms of children with ADHD (measured by SNAP-IV), as well as a few domains related to functional impairments and problematic behaviors (measured by WFIRS and PSQ).

Table 2 summarizes the statistical results of efficacy outcomes during the 4-week treatment. The distributions of efficacy-outcome improvements (from day 0 to day 28) among the participants are shown in Figure 5. The proportions of participants whose efficacy outcomes showed improvements were 80% (76/95 in TOVA), 71.13% (69/97 in SNAP-IV total), 70.1% (68/97 in SNAP-IV AD), 70.1% (68/97 in SNAP-IV HD), 55.67% (54/97 in SNAP-IV ODD), 54.64% (53/97 in WFIRS-family), 52.58% (51/97 in WFIRS-school), 52.58% (51/97 in WFIRS-life skills), 38.14% (37/97 in WFIRS-child’s self-concept), 54.64% (53/97 in WFIRS-social activities), 42.27% (41/97 in WFIRS-risky activities), 58.76% (57/97 in PSQ-conduct problem), 48.45% (47/97 in PSQ-learning problem), 45.36% (44/97 in PSQ-psychosomatic problem), 32.99% (32/97 in PSQ-anxiety), 47.42% (46/97 in PSQ-impulsivity-hyperactivity), and 57.73% (56/97 in PSQ-hyperactivity index).

Figure 5. Distributions of efficacy-outcome improvements (from day 0 to day 28) among the participants suggest significant improvement of symptoms. (A) Distribution of Test of Variables of Attention (TOVA) Attention Performance Index (API) changes from day 0 to day 28; (B) Distributions of Swanson, Nolan, and Pelham Questionnaire, version 4 (SNAP-IV) score changes from day 0 to day 28; (C) Distributions of Weiss Functional Impairment Rating Scale score changes from day 0 to day 28; (D) Distributions of Conner's Parent Symptom Questionnaire score changes from day 0 to day 28. The ranges corresponding to improvement are highlighted. AD: inattention; HD: hyperactivity/impulsivity; ODD: oppositional defiant disorder.



Safety Results

Parents and participants were informed to report any safety concerns and discomfort during the study. Any concerns and discomfort reported by parents and participants during the study were recorded. Overall, no device-related AE or severe AE was observed or reported during the 4-week intervention and the 3-month follow-up period, suggesting the safety of treatment.

Discussion

Principal Findings

This paper reports the first Chinese clinical trial outcomes of an adaptive multitasking training paradigm for the treatment of children with ADHD without medication. This study preliminarily suggested the significant improvements of ADHD symptoms and attention function after 1-month digital therapy (combined with basic BPT, without medication) using the “Attention Enhancement Training Software” (ADHD-DTx), a

video game-like training software for children with ADHD aged 6-12 years, with satisfying safety outcomes.

Comparison to Prior Work

The efficacy results are consistent with previous studies using an adaptive multitasking training paradigm [7,8,10], indicating the potential application value (as an individual or adjuvant treatment) of video game-like digital therapy for children with ADHD. The potential contribution of game-like cognitive training on attention has been reported both in children and adults [26-30], which suggests a new digital solution for attention-related deficits (such as ADHD, mild cognitive impairment, etc) across age groups [31,32].

The BPT has been extensively studied, and the overall effect size was estimated by previous researchers. Gubbels and colleagues [33] suggested that the overall effect of parent training was 0.416 (Cohen *d*) in their meta-analysis research, while Zwi and colleagues [23] suggested that the effect size

might be “between small and medium” in their review. In comparison, we calculated the effect size (Cohen *d*) of the intervention in this study, based on TOVA-API and SNAP-IV data. We found that the effect size of TOVA-API was 0.80, while that of SNAP-IV total score was 0.50. These results exceeded the effect sizes of BPT estimated by previous researchers, suggesting that the ADHD-DTx digital therapeutic, when integrated with BPT, may yield superior outcomes to BPT alone. However, such cross-study comparisons are informal and cannot independently establish the efficacy of ADHD-DTx. More rigorously controlled trials—directly comparing the combined intervention against BPT alone—are needed to draw definitive conclusions regarding the stand-alone contribution of the digital therapy.

Compared with the FDA-approved EndeavorRx (AKL-T01) software, the “Attention Enhancement Training Software” (ADHD-DTx) added the “digit cancellation task” (a widely used attention assessment and training method in clinical practice [11]) to further enhance the training effect on attention function. The “digit cancellation task” was simplified from the clinical version and consisted of 5 different types of tasks: (1) choose a specific number (eg, “choose 3”), (2) choose the adjacent (left or right) number of a specific number (eg, “choose the adjacent number on the left of 3”), (3) choose a specific number adjacent (left or right) to a specific number (eg, “choose 2 on the left of 3”), (4) choose the number in the middle position of 2 specific numbers (eg, “choose the number in the middle position of 2 and 3”), and (5) choose a specific type (even or odd) of number in the middle position of 2 specific numbers (eg, “choose the even number in the middle position of 2 and 3”). These tasks targeted multiple dimensions of attention (such as pointing, shifting, selection, span, and allocation), aiming to activate extensive attention-related brain regions and achieve the training effect of attention function.

Limitations

The key objective of this study was to explore the clinical safety and efficacy of ADHD-DTx digital therapy. However, considering that this study was single-armed and did not use a control group, it would be meaningful to attempt to exclude the influence of basic BPT from the final therapeutic effect. We suggest that further research should use a randomized, double-blinded, parallel-controlled design, which could provide the most reliable data reflecting the actual effect of ADHD-DTx.

In this study, several potential confounding factors should be considered. First, the subjective evaluation bias was caused by the open-label design, the subjective parent-reported scales for all secondary outcomes, and the large research time span. Open-label design and subjective scales could introduce a significant risk of reporter bias or placebo effect, while the 1-year time span of this study might further enhance the potential subjective bias due to the fluctuating behavior patterns of children between school and vacation periods. These confounding factors could not be separated from the intervention’s true effect, and future studies should adopt a double-blind, randomized controlled design to mitigate these biases.

Second, heterogeneity in the effectiveness of BPT may have contributed to variability in outcomes. The actual effect of BPT largely depended on the performance of parents when providing behavioral guidance to children with ADHD. However, due to differences in educational level, professional skills, and training quality, the actual effectiveness may vary among different parents, resulting in population heterogeneity. During the study, all parents were asked to daily report their home-based training progress, and all participants exhibited high adherence to both ADHD-DTx and BPT. And the subgroup analysis focusing on parental education level did not show a significant difference in efficacy outcome. Therefore, the potential influence of adherence and parental education level may be limited. However, the potential impact of training quality differences between parents still existed.

Third, the potential contribution of BPT was not excluded in this single-arm study design, which might amplify the effectiveness results. As a standard foundational treatment for ADHD, BPT was mandated by the regulatory agency for all participants due to ethical considerations. While its efficacy was well-established, albeit generally mild, its concurrent implementation with the investigational digital therapy (ADHD-DTx) introduced a considerable confounding effect that may inflate the perceived effectiveness of the intervention. Improved study design should include a control group (providing BPT equally to both test and control groups, while only the test group is treated with ADHD-DTx) so as to effectively control the confounding influence of BPT, allowing for a more precise assessment of the benefit of ADHD-DTx.

Fourth, the generalizability of the results was constrained due to limited sample diversity. The included participants were predominantly male (99/110, 90%) and of Han ethnicity (106/110, 96.36%), which weakened the interpretability of outcomes for females and other ethnicities. Future research should prioritize enrolling more representative samples with balanced gender distribution and greater ethnic diversity to ascertain the efficacy of the intervention across the broader population with ADHD.

Fifth, the rigorously monitored and controlled conditions of a clinical trial may not perfectly predict the intervention’s effectiveness in real-world clinical practice. The high adherence observed in this study was facilitated by intensive monitoring and support, which limited the direct translation of these efficacy results into real-world practice, where effectiveness might be weaker. Consequently, future investigation should involve pragmatic trials to generate robust evidence on the actual efficacy and implementation of the intervention.

Finally, the subgroup analyses for age and gender were likely underpowered due to the small sample sizes in the female and older groups. Consequently, any observed differences (or lack thereof) between these subgroups should be considered preliminary and interpreted with caution. Further research specifically designed to investigate these demographic groups is needed to advance our understanding of digital therapeutics for ADHD.

Future Directions

Further improvements could still be carried out in future research. In this single-arm, open-label study, we preliminarily revealed the safety and efficacy of ADHD-DTx digital therapy. While future research should use a randomized, double-blinded, parallel controlled design so as to prevent potential placebo effect and rule out the effect of basic BPT at home, it will achieve a better evaluation of the actual effect size of ADHD-DTx. Further study could also investigate the effect of a prolonged treatment period (>4 weeks) to study the long-term influence of digital therapy. Longer safety monitoring after the intervention period is also needed to reveal potential effects on visual acuity, daily activities, sleep quality, screen usage time, and so on. The independent ethical oversight should continue to be valued and implemented to avoid potential conflicts of interest. Future studies should include comparisons with well-established interventions like pharmacotherapy, psychological behavioral therapy, and so on, to obtain more informative evidence for better clinical practice.

There is psychological and neurophysiological evidence suggesting the potential mechanisms of ADHD treatment by the adaptive multitasking paradigm. According to electroencephalogram studies focusing on the neural mechanisms of the adaptive multitasking paradigm, after 4 weeks of training, participants' FM significantly enhanced [9,10]. The FM is a well-established neural marker of attention control [34-38], and its enhancement is related to the suppression of a key node of the "default mode network" [10,39,40], leading to a reduction of the susceptibility to internal distraction, resulting in better task performance and sustained attention. To explore potential neural mechanisms of digital therapy, physiological measurements of brain function (such as

electroencephalogram, functional magnetic resonance imaging, functional near-infrared spectroscopy, functional ultrasound, etc) could provide irreplaceable insight and should be considered in future studies.

The evaluation methods used in this study were limited, and further study could include other powerful tools such as other neuropsychological tests (eg, the Integrated Visual and Auditory Continuous Performance Test, the Wisconsin Card Sorting Test, the Cambridge Neuropsychological Test Automatic Battery, etc), classic scales (eg, the Execution Function Parent Questionnaire, the Vanderbilt ADHD Parent Rating Scale, etc).

Due to the limited data on female patients, we suggest that further research could include a higher proportion of female patients so as to make the results more representative. The inclusion and exclusion criteria set strict restrictions on comorbidity, which, on the one hand, could reduce potential confounding factors; however, on the other hand, it would limit our knowledge of real-world situations. We suggest that subsequent research can specifically investigate the effect of ADHD digital therapy involving comorbidities. We also suggest including the examination of visual acuity in further study to investigate the potential impact of long-term electronic device training on vision.

Conclusions

According to our pilot study on the Attention Enhancement Training Software (ADHD-DTx), we suggest that it could be used as a daily home-training tool for children with ADHD as an adjunct therapy to medication or other behavioral therapies, thereby further improving the intervention efficacy and reducing potential adverse drug reactions.

Acknowledgments

No artificial intelligence-generated text was used in this paper.

Funding

We thank all volunteers for their participation in the study. This work was supported in part by STI 2030-Major Projects (grant 2021ZD0200400), in part by the National Natural Science Foundation of China (grant 62336007), in part by the Key Research and Development Program of Zhejiang (grant 2023C03003), in part by the Key R&D Program of Zhejiang (2024SSYS0016), in part by the Starry Night Science Fund of the Zhejiang University Shanghai Institute for Advanced Study (grant SN-ZJU-SIAS-002), in part by the Fundamental Research Funds for the Central Universities, in part by ZJU-GENSCI Children's Health Research & Development Center (ZJU-GENSCI2024YB003), in part by the Project for Hangzhou Medical Disciplines of Excellence, and in part by the Key Project for Hangzhou Medical Disciplines.

Data Availability

The data that support the findings of this study are available on request from the corresponding author upon reasonable request.

Authors' Contributions

LY, KJ, and WL designed the study. MB, JY, Jiangping Wang, and KJ collected the data. Jiaheng Wang and SF contributed unpublished analytic scripts. SF analyzed the data and drafted the manuscript.

Conflicts of Interest

SF and WL are employed by SDO Digital Therapeutics (a provider of advanced digital and neurological therapies for children) at the time of submission, while SF did not have any financial or nonfinancial competing interests during the clinical trial period

(2021). SDO Digital Therapeutics did not participate in the data collection of this study. All other authors declare no financial or nonfinancial competing interests. The views and opinions expressed within this manuscript are those of all the authors. Our adherence to the policies on data availability is not altered.

Multimedia Appendix 1

Complete inclusion and exclusion criteria.

[[DOCX File , 17 KB - games_v14i1e76114_app1.docx](#)]

Multimedia Appendix 2

Description of outcome measurements.

[[DOCX File , 14 KB - games_v14i1e76114_app2.docx](#)]

References

1. Faraone SV, Banaschewski T, Coghill D, Zheng Y, Biederman J, Bellgrove MA, et al. The World federation of ADHD international consensus statement: 208 evidence-based conclusions about the disorder. *Neurosci Biobehav Rev* 2021;128:789-818 [[FREE Full text](#)] [doi: [10.1016/j.neubiorev.2021.01.022](https://doi.org/10.1016/j.neubiorev.2021.01.022)] [Medline: [33549739](#)]
2. Dalsgaard S, Mortensen PB, Frydenberg M, Thomsen PH. ADHD, stimulant treatment in childhood and subsequent substance abuse in adulthood - a naturalistic long-term follow-up study. *Addict Behav* 2014;39(1):325-328. [doi: [10.1016/j.addbeh.2013.09.002](https://doi.org/10.1016/j.addbeh.2013.09.002)] [Medline: [24090624](#)]
3. Pelham WE, Fabiano GA, Waxmonsky JG, Greiner AR, Gnagy EM, Pelham WE, et al. Treatment sequencing for childhood ADHD: a multiple-randomization study of adaptive medication and behavioral interventions. *J Clin Child Adolesc Psychol* 2016;45(4):396-415 [[FREE Full text](#)] [doi: [10.1080/15374416.2015.1105138](https://doi.org/10.1080/15374416.2015.1105138)] [Medline: [26882332](#)]
4. Qiu H, Liang X, Wang P, Zhang H, Shum DH. Efficacy of non-pharmacological interventions on executive functions in children and adolescents with ADHD: a systematic review and meta-analysis. *Asian J Psychiatr* 2023;87:103692 [[FREE Full text](#)] [doi: [10.1016/j.ajp.2023.103692](https://doi.org/10.1016/j.ajp.2023.103692)] [Medline: [37450981](#)]
5. He F, Qi Y, Zhou Y, Cao A, Yue X, Fang S, et al. Meta-analysis of the efficacy of digital therapies in children with attention-deficit hyperactivity disorder. *Front Psychiatry* 2023;14:1054831 [[FREE Full text](#)] [doi: [10.3389/fpsy.2023.1054831](https://doi.org/10.3389/fpsy.2023.1054831)] [Medline: [37260755](#)]
6. Davis NO, Bower J, Kollins SH. Proof-of-concept study of an at-home, engaging, digital intervention for pediatric ADHD. *PLoS One* 2018;13(1):e0189749 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0189749](https://doi.org/10.1371/journal.pone.0189749)] [Medline: [29324745](#)]
7. Kollins SH, DeLoss DJ, Cañadas E, Lutz J, Findling RL, Keefe R, et al. A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): a randomised controlled trial. *Lancet Digit Health* 2020;2(4):e168-e178 [[FREE Full text](#)] [doi: [10.1016/S2589-7500\(20\)30017-0](https://doi.org/10.1016/S2589-7500(20)30017-0)] [Medline: [33334505](#)]
8. Kollins SH, Childress A, Heusser AC, Lutz J. Effectiveness of a digital therapeutic as adjunct to treatment with medication in pediatric ADHD. *NPJ Digit Med* 2021;4(1):58 [[FREE Full text](#)] [doi: [10.1038/s41746-021-00429-0](https://doi.org/10.1038/s41746-021-00429-0)] [Medline: [33772095](#)]
9. Gallen CL, Anguera JA, Gerdes MR, Simon AJ, Cañadas E, Marco EJ. Enhancing neural markers of attention in children with ADHD using a digital therapeutic. *PLoS ONE* 2021;16(12):e0261981. [doi: [10.1371/journal.pone.0261981](https://doi.org/10.1371/journal.pone.0261981)] [Medline: [34972140](#)]
10. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. *Nature* 2013;501(7465):97-101. [doi: [10.1038/nature12486](https://doi.org/10.1038/nature12486)] [Medline: [24005416](#)]
11. Hatta T, Yoshizaki K, Ito Y, Mase M, Kabasawa H. Reliability and validity of the digit cancellation test, a brief screen of attention. *Int J Psychol Sci* 2012;55(4):246-256. [doi: [10.2117/psysoc.2012.246](https://doi.org/10.2117/psysoc.2012.246)]
12. van den Hoofdakker BJ, van der Veen-Mulders L, Sytema S, Emmelkamp PMG, Minderaa B, Nauta MH. Effectiveness of behavioral parent training for children with ADHD in routine clinical practice: a randomized controlled study. *J Am Acad Child Adolesc Psychiatry* 2007;46(10):1263-1271. [doi: [10.1097/chi.0b013e3181354bc2](https://doi.org/10.1097/chi.0b013e3181354bc2)] [Medline: [17885567](#)]
13. Chacko A, Jensen SA, Lowry LS, Cornwell M, Chimklis A, Chan E, et al. Engagement in behavioral parent training: review of the literature and implications for practice. *Clin Child Fam Psychol Rev* 2016;19(3):204-215. [doi: [10.1007/s10567-016-0205-2](https://doi.org/10.1007/s10567-016-0205-2)] [Medline: [27311693](#)]
14. Forbes GB. Clinical utility of the test of variables of attention (TOVA) in the diagnosis of attention-deficit/hyperactivity disorder. *J Clin Psychol* 1998;54(4):461-476. [doi: [10.1002/\(sici\)1097-4679\(199806\)54:4<461::aid-jclp8>3.0.co;2-q](https://doi.org/10.1002/(sici)1097-4679(199806)54:4<461::aid-jclp8>3.0.co;2-q)]
15. Greenberg LM, Waldman ID. Developmental normative data on the test of variables of attention (T.O.V.A.). *J Child Psychol Psychiatry* 1993;34(6):1019-1030. [doi: [10.1111/j.1469-7610.1993.tb01105.x](https://doi.org/10.1111/j.1469-7610.1993.tb01105.x)] [Medline: [8408366](#)]
16. Swanson J. *School-Based Assessments and Interventions for ADD Students*. Frederick, Maryland, USA: KC publishing; 1992.
17. Swanson JM, Kraemer HC, Hinshaw SP, Arnold LE, Conners CK, Abikoff HB, et al. Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *J Am Acad Child Adolesc Psychiatry* 2001;40(2):168-179. [doi: [10.1097/00004583-200102000-00011](https://doi.org/10.1097/00004583-200102000-00011)] [Medline: [11211365](#)]

18. Swanson JM, Sandman CA, Deutsch C, Baren M. Methylphenidate hydrochloride given with or before breakfast: I. behavioral, cognitive, and electrophysiologic effects. *Pediatrics* 1983;72(1):49-55. [Medline: [6866591](#)]
19. Weiss MD, McBride NM, Craig S, Jensen P. Conceptual review of measuring functional impairment: findings from the Weiss Functional Impairment Rating Scale. *Evid Based Ment Health* 2018;21(4):155-164 [FREE Full text] [doi: [10.1136/ebmental-2018-300025](#)] [Medline: [30314990](#)]
20. Fan J, Du Y, Wang L. The norm and reliability of the Conners Parent Symptom Questionnaire in Chinese urban children. *Shanghai Arch Psychiatry* 2005;17(6):321-323.
21. Conners CK, Sitarenios G, Parker JDA, Epstein JN. The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *J Abnorm Child Psychol* 1998;26(4):257-268. [doi: [10.1023/a:1022602400621](#)] [Medline: [9700518](#)]
22. van der Oord S, Tripp G. How to improve behavioral parent and teacher training for children with ADHD: integrating empirical research on learning and motivation into treatment. *Clin Child Fam Psychol Rev* 2020;23(4):577-604 [FREE Full text] [doi: [10.1007/s10567-020-00327-z](#)] [Medline: [32968886](#)]
23. Zwi M, Jones H, Thorgaard C, York A, Dennis JA. Parent training interventions for attention deficit hyperactivity disorder (ADHD) in children aged 5 to 18 years. *Cochrane Database Syst Rev* 2011;2011(12):CD003018 [FREE Full text] [doi: [10.1002/14651858.CD003018.pub3](#)] [Medline: [22161373](#)]
24. Hornstra R, Groenman AP, van der Oord S, Luman M, Dekkers TJ, van der Veen-Mulders L, et al. Review: which components of behavioral parent and teacher training work for children with ADHD? - a meta-regression analysis on child behavioral outcomes. *Child Adolesc Ment Health* 2023;28(2):258-268. [doi: [10.1111/camh.12561](#)] [Medline: [35417075](#)]
25. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* 1995;57(1):289-300. [doi: [10.1111/j.2517-6161.1995.tb02031.x](#)]
26. Anguera JA, Schachtner JN, Simon AJ, Volponi J, Javed S, Gallen CL, et al. Long-term maintenance of multitasking abilities following video game training in older adults. *Neurobiol Aging* 2021;103:22-30. [doi: [10.1016/j.neurobiolaging.2021.02.023](#)] [Medline: [33789209](#)]
27. Jurigova BG, Gerdes MR, Anguera JA, Marco EJ. Sustained benefits of cognitive training in children with inattention, three-year follow-up. *PLoS One* 2021;16(2):e0246449 [FREE Full text] [doi: [10.1371/journal.pone.0246449](#)] [Medline: [33539468](#)]
28. Anguera JA, Volponi JJ, Simon AJ, Gallen CL, Rolle CE, Anguera-Singla R, et al. Integrated cognitive and physical fitness training enhances attention abilities in older adults. *NPJ Aging* 2022;8(1):12 [FREE Full text] [doi: [10.1038/s41514-022-00093-y](#)] [Medline: [36042247](#)]
29. Gallen CL, Schachtner JN, Anguera-Singla R, Anguera JA, Gazzaley A. Influence of game features on attention in adults. *Front Psychol* 2023;14:1123306 [FREE Full text] [doi: [10.3389/fpsyg.2023.1123306](#)] [Medline: [37228349](#)]
30. Ziegler DA, Simon AJ, Gallen CL, Skinner S, Janowich JR, Volponi JJ, et al. Closed-loop digital meditation improves sustained attention in young adults. *Nat Hum Behav* 2019;3(7):746-757 [FREE Full text] [doi: [10.1038/s41562-019-0611-9](#)] [Medline: [31160812](#)]
31. Mishra J, Anguera J, Gazzaley A. Video games for neuro-cognitive optimization. *Neuron* 2016;90(2):214-218 [FREE Full text] [doi: [10.1016/j.neuron.2016.04.010](#)] [Medline: [27100194](#)]
32. Anguera JA, Gazzaley A. Video games, cognitive exercises, and the enhancement of cognitive abilities. *Curr Opin Behav Sci* 2015;4:160-165. [doi: [10.1016/j.cobeha.2015.06.002](#)]
33. Gubbels J, van der Put CE, Assink M. The effectiveness of parent training programs for child maltreatment and their components: a meta-analysis. *Int J Environ Res Public Health* 2019;16(13):2404 [FREE Full text] [doi: [10.3390/ijerph16132404](#)] [Medline: [31284575](#)]
34. Cavanagh JF, Frank MJ. Frontal theta as a mechanism for cognitive control. *Trends Cogn Sci* 2014;18(8):414-421 [FREE Full text] [doi: [10.1016/j.tics.2014.04.012](#)] [Medline: [24835663](#)]
35. Clayton MS, Yeung N, Cohen Kadosh R. The roles of cortical oscillations in sustained attention. *Trends Cogn Sci* 2015;19(4):188-195. [doi: [10.1016/j.tics.2015.02.004](#)] [Medline: [25765608](#)]
36. Brandmeyer T, Delorme A. Closed-loop frontal midline θ neurofeedback: a novel approach for training focused-attention meditation. *Front Hum Neurosci* 2020;14:246 [FREE Full text] [doi: [10.3389/fnhum.2020.00246](#)] [Medline: [32714171](#)]
37. Messel MS, Raud L, Hoff PK, Stubberud J, Huster RJ. Frontal-midline theta reflects different mechanisms associated with proactive and reactive control of inhibition. *Neuroimage* 2021;241:118400 [FREE Full text] [doi: [10.1016/j.neuroimage.2021.118400](#)] [Medline: [34311382](#)]
38. Cowley BU, Juurmaa K, Palomäki J. Reduced power in fronto-parietal theta EEG linked to impaired attention-sampling in adult ADHD. *eNeuro* 2022;9(1) [FREE Full text] [doi: [10.1523/ENEURO.0028-21.2021](#)] [Medline: [34893505](#)]
39. Tu W, Ma Z, Ma Y, Dopfel D, Zhang N. Suppressing anterior cingulate cortex modulates default mode network and behavior in awake rats. *Cereb Cortex* 2021;31(1):312-323 [FREE Full text] [doi: [10.1093/cercor/bhaa227](#)] [Medline: [32820327](#)]
40. Chadick JZ, Zanto TP, Gazzaley A. Structural and functional differences in medial prefrontal cortex underlie distractibility and suppression deficits in ageing. *Nat Commun* 2014;5:4223 [FREE Full text] [doi: [10.1038/ncomms5223](#)] [Medline: [24979364](#)]

Abbreviations

AD: inattention subscale of SNAP-IV
ADHD: attention-deficit/hyperactivity disorder
ADHD-RS-IV: Attention-Deficit/Hyperactivity Disorder Rating Scale IV
AE: adverse event
API: Attention Performance Index
BPT: behavioral parent training
DSM-V: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
FDA: US Food and Drug Administration
FM : frontal midline theta
HD: hyperactivity/impulsivity subscale of SNAP-IV
NMPA: National Medical Products Administration
ODD: oppositional defiant disorder subscale of SNAP-IV
PSQ: Conner's Parent Symptom Questionnaire
SEM: SE of the mean
SNAP-IV: Swanson, Nolan, and Pelham Questionnaire, version 4
TOVA: Test of Variables of Attention
WFIRS: Weiss Functional Impairment Rating Scale

Edited by A Coristine; submitted 16.Apr.2025; peer-reviewed by S Mehan, S Triscari; comments to author 27.Jun.2025; accepted 22.Oct.2025; published 05.Jan.2026.

Please cite as:

Bao M, Feng S, Wang J, Ye J, Wang J, Li W, Jiang K, Yao L

Efficacy and Safety of a Video Game-Like Digital Therapy Intervention for Chinese Children With Attention-Deficit/Hyperactivity Disorder: Single-Arm, Open-Label Pre-Post Study

JMIR Serious Games 2026;14:e76114

URL: <https://games.jmir.org/2026/1/e76114>

doi: [10.2196/76114](https://doi.org/10.2196/76114)

PMID:

©Mengyi Bao, Shang Feng, Jiaheng Wang, Jia Ye, Jiangping Wang, Wenyu Li, Kewen Jiang, Lin Yao. Originally published in JMIR Serious Games (<https://games.jmir.org>), 05.Jan.2026. 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.

Original Paper

Effects of Virtual Nature Embodiment on Compassion, Empathy, Nature Connectedness, and Symptom Burden in Patients With Psychosis or Depression: An Explorative Clinical Study

Alva Lütt^{1,2,3,4}, Dr med; Kristina Lütgens⁵, BSc; Christiane Montag^{1,2}, Prof Dr Med; Stefan Gutwinski^{1,2}, Prof Dr Med; Georg Felix Reuth⁶, MEd; Steve Nebel⁶, Prof Dr; Felix BERPohl^{1,2}, Prof Dr Med; Pia Spangenberg⁶, PhD

¹Department of Psychiatry and Neurosciences, Charité Campus Mitte, Charité – Universitätsmedizin Berlin, Berlin, Germany

²Department of Psychiatry and Psychotherapy, Charité at St. Hedwig Hospital, Charité – Universitätsmedizin Berlin, Berlin, Germany

³German Center for Mental Health (DZPG), partner site Berlin, Berlin, Germany

⁴Berlin Institute of Health at Charité - Universitätsmedizin Berlin, BIH Biomedical Innovation Academy, BIH Charité Junior Digital Clinician Scientist Program, Berlin, Germany

⁵Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany

⁶Department of Education, University of Potsdam, Potsdam, Germany

Corresponding Author:

Alva Lütt, Dr med

Department of Psychiatry and Neurosciences, Charité Campus Mitte, Charité – Universitätsmedizin Berlin

Charitéplatz 1

Berlin, 10117

Germany

Phone: 49 2311 2969

Email: alva.luettt@charite.de

Abstract

Background: Nature experiences may have a positive impact on mental health. Innovative alternatives, such as immersive virtual reality (iVR), can have similar effects. In previous studies on embodying a tree in virtual reality, connectedness to nature has been induced in healthy participants and shown to be influenced by compassion. Compassion and empathy, however, can be altered in psychiatric disorders, leading to impaired relationships with fellow human beings. The potential effect of nature experiences in iVR on compassion, empathy, and nature connectedness in mental health disorders has not yet been investigated.

Objective: This study aims to examine the development of nature connectedness, compassion, empathy, and individual symptom load in patients with depression (n=20), schizophrenia (n=20), and healthy controls (n=20), measured pre and post 1 session of virtual embodying a rainforest tree.

Methods: We conducted an explorative clinical trial, comparing 3 groups (depression, schizophrenia, and healthy control), using repeated measures ANOVA and multiple regression analysis. Effect sizes (η_p^2) and 95% CIs were reported where applicable. We assessed the impact of iVR-exposure on nature connectedness, empathy, compassion, individual symptoms, spirituality, cybersickness, presence, and virtual body ownership. Electrodermal activity was measured to capture physiological correlates of emotional arousal.

Results: Individual symptom load decreased significantly through the experience of embodying a growing tree in iVR in both patient groups ($F_{1,38}=40.93$, $\eta_p^2=0.52$, 95% CI 0.29-0.67; $P<.001$). All groups benefited equally from the iVR experience regarding a change in nature connectedness ($F_{1,57}=100.12$, $\eta_p^2=0.637$, 95% CI 0.48-0.74; $P<.001$) and compassion ($F_{1,57}=12.86$, $\eta_p^2=0.18$, 95% CI 0.04-0.36; $P<.001$). The change in empathy did not differ significantly between the 3 groups. The analysis of electrodermal activity during iVR showed significantly higher nonspecific skin conductance response numbers than at the beginning of the iVR (2-tailed $t_{53}=4.40$, mean difference=2.38, 95% CI 1.30-3.46; $P<.001$). Changes in compassion were predicted by spirituality ($\beta=-0.35$, 2-tailed $t_{59}=-2.89$, $\eta_p^2=0.13$, 95% CI 0.01-0.30; $P=.005$) and virtual body ownership ($\beta=0.29$, 2-tailed $t_{59}=2.15$, $\eta_p^2=0.08$, 95% CI 0.00-0.23; $P=.04$) but not by the feeling of presence.

Conclusions: Our results indicate that virtually embodying a tree seems to have a positive effect on individual symptoms. Patients with depression and schizophrenia were able to receive an iVR-based intervention and benefited from the experience.

Our results provide a deeper understanding of the phenomenon of virtual body ownership of nature, a specific affordance of iVR, thereby laying the groundwork for future innovative body psychotherapy interventions for individuals with mental health disorders.

Trial Registration: ClinicalTrials.gov NCT06446856; <https://tinyurl.com/nhjar26p>

(*JMIR Serious Games* 2026;14:e74337) doi:[10.2196/74337](https://doi.org/10.2196/74337)

KEYWORDS

immersive virtual reality; nature connectedness; psychosis; depression; compassion; empathy; symptom burden; virtual body ownership

Introduction

Background

Positive effects of nature experiences on mental health have been shown by numerous studies and meta-analyses, especially in depressive and anxiety disorders [1-3]. However, a recent systematic review by Paredes-Céspedes et al [4] from 2024 concluded that outcomes of nature-based therapy studies were inconclusive in regard to definitive effects on stress, anxiety, or depression, for example, due to varying interventions, sample sizes, and methods, thereby highlighting the need for further studies on this topic [4]. Regarding the potential therapeutic effects of experiencing nature, the practicability of potential interventions in everyday clinical routines for patients, which sometimes take place far away from nature (eg, in hospitals), has to be considered. Additionally, urbanization means that fewer and fewer people are able to experience nature directly [5].

Virtual reality (VR)-based applications are becoming increasingly important in the development of innovative, easily accessible, therapeutic procedures in mental health care [6,7]. Patients can immerse themselves in various computer-generated environments with the help of a head-mounted display (HMD), ranging from, for example, exposure scenarios for anxiety disorders or substance use disorders to avatar-based interventions for psychotic disorders [6]. Exploring these 3D scenarios, a feeling of “presence,” of being in this virtual world, can arise, the subjective correlate of “immersion” as the computer system’s capacity to deliver a vivid experience [8]. Related to the sense of presence, “embodiment” is another key concept in immersive virtual reality (iVR), with one of its subcomponents being “virtual body ownership” [9]. Although an experience in iVR cannot replace real experiences in nature [10], a recent study in healthy volunteers compared specific emotions, presence, and immersion in iVR with real natural environments and found that only some aesthetic emotions, state anxiety, and presence, but not immersion, were significantly more intense in real situations than in iVR environments [11]. Studies show that the level of immersion and realism is highly relevant for affective responses to nature scenarios, underlining the importance of iVR over, for example, videos [12]. Especially, virtual body ownership seems to be a promising approach to foster affective and cognitive processes relevant to the feeling of nature connectedness [13,14]. Hence, it is not surprising that interest in iVR-based nature exposure is also increasing in the field of mental health.

Ilioudi et al [15] described the patient experiences of 20 patients with bipolar and unipolar depression who used an iVR calm room with relaxing nature scenarios in an inpatient psychiatric setting. Qualitative interviews showed an increase in awareness, calmness, and well-being as well as patients’ gratefulness about these nonpharmacological alternatives for anxiety relief [15].

Another pilot study on the feasibility and acceptability of an iVR relaxation intervention in an acute psychiatric setting showed a statistically significant increase in relaxation, happiness, and connectedness to nature and a decrease in stress, anxiety, and sadness, all measured by visual analog scales (VASs) [16]. Even effects on violence were shown—according to the authors, violent incidents and restrictive practices decreased during the implementation of this iVR-application. The intervention consisted of 1 session, maximum 1 hour, with several possible scenarios open for exploration (eg, a beach and scuba diving with dolphins [16]).

Despite these promising examples, the induction of affective processes and their effects on different psychiatric disorders of such iVR-based nature experiences remain relatively unexplored. In a preceding study with healthy students embodying a tree in a virtual rain forest, previous studies could show that compassion is a crucial influencing factor for the development of nature connectedness [17]. However, deficits in compassion and empathy are described in several psychiatric disorders, whereas patients with schizophrenia, examined with the Multifaceted Empathy Test (MET) and Interpersonal Reactivity Index, showed impairments in cognitive empathy but not emotional empathy [18]; patients with persistent or recurrent depression completing the same tests had deficits in emotional but not cognitive empathy [19]. The importance of these abilities for psychosocial functioning, or on the other side, isolation and social withdrawal is obvious [18,19]. Whether these impairments also influence connectedness to nature in patients with psychotic or depressive disorders remains unclear.

In addition to the development of nature connectedness, iVR-based interventions, especially the experience of virtually embodying a tree, could lead to an enhancement of body experiences, which are also the focus of body psychotherapy interventions [20]. The question arises whether a single iVR intervention is already associated with changes in affect and symptom load, thereby indicating a potential therapeutic effectiveness of such interventions, which could then be investigated in further studies.

Objective of the Study

This explorative study aims to examine the effect of 1 session of an iVR-based nature experience in 2 different patient groups

with highly prevalent serious mental illnesses, patients with depressive disorders, and patients with schizophrenia, compared with a healthy control (HC) group. We test our hypotheses that patients with a depressive or psychotic disorder and HCs show differential effects of the iVR intervention on (1) nature connectedness, (2) empathy, and (3) compassion. In addition, we hypothesize that (4) patients with psychotic or depressive disorder have a lower symptom burden after the iVR experience than before the iVR experience. On an explorative basis, we test the hypothesis that (5) the measurement of electrodermal activity (EDA) before, during, and after the iVR application allows emotional arousal to be objectified. Additionally, we test the explorative hypothesis (H6) that spirituality, presence and virtual body ownership have an influence on nature connectedness, empathy and compassion in all 3 groups.

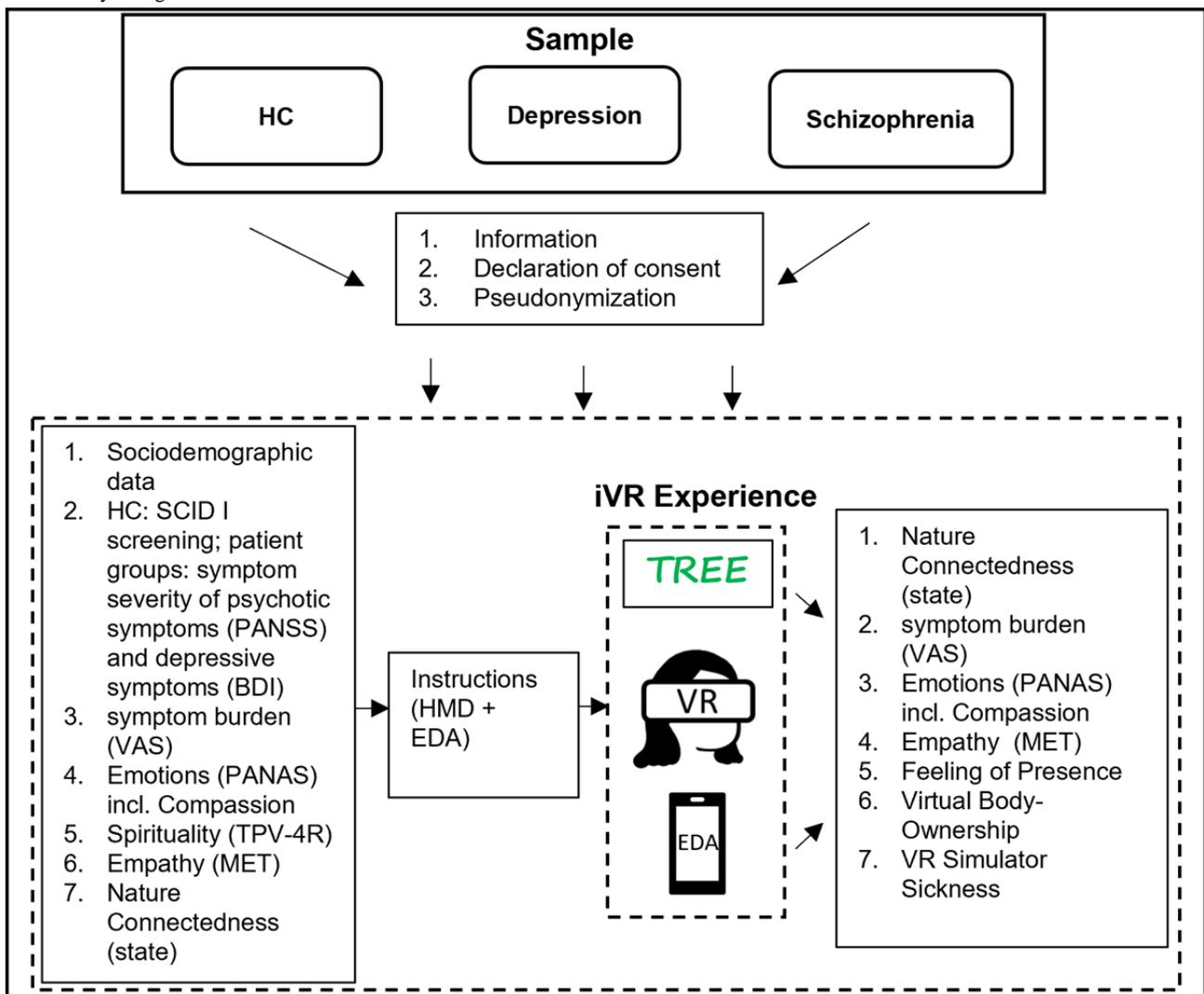
Methods

Study Design

This clinical trial study aimed to evaluate symptom load, empathy, compassion, and nature connectedness before and after an iVR experience in patients with depressive disorder, patients with schizophrenia, and HC participants (participants between 18 and 65 years of age).

The study was designed as follows. Before participating in the study, participants were informed about the nature of the experiment and declared consent (Figure 1). Symptom severity was assessed with the Beck Depression Inventory (BDI) for depressive symptoms and the Positive and Negative Syndrome Scale (PANSS) for psychotic symptoms. HCs were screened with the Structured Clinical Interview for *DSM-IV* (*Diagnostic and Statistical Manual of Mental Disorders* [Fourth Edition]; SCID I). Being a pre-post intervention, participants filled out an online questionnaire before the VR-exposure, containing questions on sociodemographics, symptom burden perceived in the moment with a VAS, as well as perceived nature connectedness (state), spirituality, and emotions, including compassion (Positive and Negative Affect Schedule [PANAS]) and empathy (MET). After filling out the questionnaire, participants put on the VR headset and were equipped with an EDA measurement device attached to the arm and fingers. After experiencing the VR application for 4 minutes and 20 seconds, participants filled out the post-VR questionnaire, containing items on symptom burden (VAS), nature connectedness (state), emotions including compassion (PANAS), feeling of presence, virtual body ownership, VR simulator sickness, and empathy (MET).

Figure 1. Study Design.



Setting

The experimental setting included a notebook for questionnaire administration and the VR hardware (refer to Material section), arranged in a setting that provided sufficient space for participants. Data collection took place at the Psychiatric University Hospital Charité at St. Hedwig Hospital, Berlin, Germany, over a 4-week period (May 23-June 19, 2024), during which participants were recruited, and data were assessed. Participants were reimbursed with €15 (approximately US \$17).

Material

The iVR application “Tree” was used, developed by a group of researchers from the Massachusetts Institute of Technology and filmmakers [21,22]. This application has been used in previous studies to examine its impact on nature connectedness within healthy individuals, observing a positive effect on nature connectedness [13,14,17]. In the application, users explore the growth of a rainforest tree, starting from the viewpoint of a seed under the earth to a fully-grown tree reaching the sky (refer to Figure 2 for a screenshot from the VR application). The illusion of virtual body ownership was induced by synchronous

movements of, for example, the arms of participants and the corresponding branch of the tree. While users embody the growing tree, they can watch various animals (eg, ants, birds, and monkeys) around them, as well as other trees in the jungle. In the original version, a forest fire occurs at the end of the experience. However, in this study, we used only the nonthreatening part of the iVR experience, stopping the application after 4 minutes and 20 seconds. The VR sessions were conducted by a psychiatrist (AL) and a second trained experimenter (KL) with experience in VR application and treatment of patients with acute psychiatric conditions. The hardware used for the VR exposure consisted of an HMD (HTC Vive Focus 3, including controllers, field of view 120°, resolution of 2448×2448 pixels per eye, 6DoF, 90Hz refresh rate), a notebook, an Android smartphone, and an EDA device. Patients were standing throughout the exposure and were able to move. The HMD was connected to an Alienware notebook during the entire experiment (Intel Core i9-8 8950HK processor, 16 GB RAM, NVIDIA GeForce GTX 1080, Windows 10 Home). The smartphone was a Samsung S10e running on Android 12, connected to the Mindfield eSense Skin Response Meter (recording at 10Hz).

Figure 2. Own Screenshot of the VR-application Tree purchased from VIVEPORT developed by New Reality Co. (<http://newreality.co>) [21, 22]. Used under quotation exception for research purposes.



Participants

Participants were inpatients or outpatients and HCs recruited via an open call (purposive sampling) at the Psychiatric University Hospital Charité at St. Hedwig Hospital, Berlin,

Germany. Interested parties were invited to a brief information session (approximately 5-10 minutes), during which it was also assessed whether the inclusion criteria were met. The following inclusion and exclusion criteria (Textbox 1) were binding for participating in the study.

Textbox 1. Inclusion and exclusion criteria.**Inclusion criteria**

- Age: 18 to 65 years
- Inpatients or outpatients of the Psychiatric University Hospital Charité at St. Hedwig Hospital
- Diagnosis of schizophrenia (F20; schizophrenia) or unipolar depressive disorder according to the *ICD-10 (International Statistical Classification of Diseases, Tenth Revision)*; F32, F33; depression) or healthy controls between 18 and 65 years
- Ability to provide written informed consent after receiving study information

Exclusion criteria

- Acute suicidality or endangerment of others
- Primary eating disorder requiring treatment
- Acute dermatological condition affecting the hands that could interfere with electrodermal activity measurement
- Healthy control group: Diagnosis of a psychiatric or psychosomatic disorder; additional exclusion criteria corresponding to those of the patient group

Assessments

All participants were exposed to the virtual body ownership experience as described in the previous sections. Hypotheses of this study can be found in the Introduction section. Outcomes and methods of assessment are described in detail in the Data Sources section.

Data Sources**Nature Connectedness**

Measurement of nature connectedness as a state variable with the “Inclusion of Nature in Self” (INS) scale [23] before and after iVR exposure, based on the German version by Spangenberg et al [13]. The item consisted of a slider based on percentages (0%=zero overlap of “I” and “nature” and 100%=total overlap). Calculating the test-retest reliability revealed a satisfying Pearson correlation of $r=0.64$.

Empathy

Evaluation of cognitive empathy and emotional empathy explicit with the MET [24], before and after the iVR exposure. Questions on emotional empathy implicit (levels of arousal) were excluded. The questionnaire was divided into 2 equal parts, which the patients completed in pseudorandomized order before and after the iVR exposure. Internal reliability measures were conducted for both pseudorandomized orders and both subscales. The cognitive empathy scale revealed an acceptable internal reliability for both pseudorandomized orders (Order A: pre iVR McDonald $\omega=0.64$, post iVR McDonald $\omega=0.74$; and Order B: pre iVR McDonald $\omega=0.75$, post iVR McDonald $\omega=0.73$). The emotional empathy scale revealed an excellent internal reliability for both pseudorandomized orders (Order A: pre iVR McDonald $\omega=0.97$, post VR McDonald $\omega=0.97$; and Order B: pre iVR McDonald $\omega=0.95$, post iVR McDonald $\omega=0.97$).

Compassion

Evaluation of compassion (state compassion) pre and post the iVR exposure with the PANAS and 5 integrated items measuring compassion based on Pfattheicher et al [25]. The compassion scale revealed a good internal reliability for pre (McDonald $\omega=0.85$) and post iVR exposure (McDonald $\omega=0.85$).

Symptom Burden

Symptom burden was assessed with the VAS (0-10) measuring the burden of the individual 3 main symptoms, previously specified during assessment of symptom severity (PANSS based on Kay et al [26]) and BDI based on Beck et al [27]. Calculating the test-retest, the VAS reliability revealed a satisfying Pearson correlation of $r=0.62$.

Explorative EDA

EDA was analyzed using the counted number of nonspecific skin conductance responses (NS-SCRs) provided by the Mindfield eSense Skin Response device during iVR exposure.

Structured Clinical Interview for DSM-IV (SCID I)

The screening with SCID I was conducted before the iVR exposure to screen for psychiatric symptoms in the HC group based on First and Gibbon [28]. The SCID I revealed an excellent internal reliability (McDonald $\omega=0.90$; item 3 excluded due to an SD 0).

VR Simulator Sickness

Motion sickness as a potential side effect of iVR, measured with the Virtual Reality Sickness Questionnaire based on Kim et al [29], after exposure. This questionnaire revealed an acceptable internal reliability (McDonald $\omega=0.70$).

Definition of Covariates

Embodiment was assessed as the feeling of virtual body ownership with an adapted and translated subscale of the Embodiment scale by Spangenberg et al [13], based on Ahn et al [30] and Slater et al [31], consisting of 5 items. This questionnaire revealed an excellent internal reliability (McDonald $\omega=0.95$).

Spirituality was assessed using the 4-item short version of the questionnaire “Transpersonales Vertrauen” based on Hampel et al [32] before iVR exposure. This questionnaire revealed a good internal reliability (McDonald $\omega=0.81$).

Feeling of presence in iVR: Scale according to Ahn et al [30], adapted from the “Spatial Presence scale” by Bailenson et al [33], German translation by Spangenberg et al [13], adapted

to the virtual Amazon forest area after exposure. The internal reliability of this questionnaire was acceptable (McDonald $\omega=0.79$).

PANSS was based on the study by Kay et al [26]. Furthermore, 5 subscores according to van der Gaag et al [34] were calculated to better differentiate between positive symptoms, negative symptoms, disorganization, excitement, and emotional distress. The PANSS total internal reliability was excellent (McDonald $\omega=0.93$).

BDI was based on the study by Beck et al [27]. The BDI revealed a good internal reliability (McDonald $\omega=0.87$).

Study Size

Due to the exploratory nature of this study and the lack of comparable studies, the number of cases is not based on a sample size calculation but based on a consideration of feasibility (given the technically complex preparation and implementation), and is intended to provide a basis for subsequent follow-up studies. Effect sizes in the 3 groups could be used in a follow-up study for power analyses and thus for calculating sample size.

Data Analysis

All outcome variables were checked for outliers and extreme values before conducting the analyses. The assumptions for analysis of variance were also examined [35]. The results are presented in the [Multimedia Appendix 1](#).

Baseline characteristics between the 3 groups were compared using either the chi-square test or, if cell counts were less than 5, the Fisher exact test.

To analyze hypothesis 1 (differences in nature connectedness), hypothesis 2 (differences in cognitive empathy and emotional empathy explicit), hypothesis 3 (differences in compassion), and hypothesis 4 (differences in mean VAS scores pre- vs post-iVR exposure), we calculated a repeated measures ANOVA and post hoc tests (Tukey [36]). Additionally, we conducted ANCOVA with depressive (BDI sum score) and psychotic symptom severity (PANSS subscales according to the 5-factor solution by van der Gaag et al [34]) as covariates and delta VAS as the dependent variable.

To test for physical emotional arousal (hypothesis 5), EDA was recorded throughout the experience using the eSense skin-response device from Mindfield, capturing data in μS at a 5 Hz sampling rate. Then, 2 electrodes were attached to the

participants' index and middle fingers on the palm side and connected by cable to a smartphone that used the eSense App to measure EDA through skin conductance and NS-SCR. The number of NS-SCRs at the end of the first full minute was set as the baseline. Paired-samples *t* tests were conducted to examine the difference between the mean of the number of NS-SCRs per minute during the iVR and the baseline.

We tested the direct effects postulated in hypothesis 6 with 4 multiple linear regression analyses, examining the influence of each independent variable as a predictor (spirituality, presence, and virtual body ownership) and the combination of them on the dependent variables (nature connectedness, empathy, and compassion).

Ethical Considerations

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The study was approved by the Ethics Committee of Charité - Universitätsmedizin Berlin Institutional Review Board (EA2/04/24; April 26, 2024) and preregistered on ClinicalTrials.gov. All participants gave written informed consent after a previous verbal explanatory discussion and the opportunity to ask questions. They received compensation of €15 (approximately US \$17) for their participation. All study data were pseudonymized by sequential numbering (without using initials or dates of birth), and no identification through images or supplementary material is possible.

Results

Descriptive Statistics

Baseline descriptive characteristics of the study participants are displayed in [Table 1](#). One participant stopped the application early due to general discomfort. All other participants completed the full VR application. Mean Virtual Reality Sickness Questionnaire scores were 1.40 (SD 0.301, range 1.00-2.22) for the patient groups, including 3 outliers, and 1.26 (SD 0.173, range 1.00-1.56) for the HC group. Patients showed a mean value of the virtual body ownership items of 4.36 (SD 1.36, range 2.00-7.00), and HCs showed a mean value of 5.35 (SD 1.37, range 1.60-7.00). The presence questionnaire post VR resulted in a mean value of 3.65 (SD 0.708, range 2.40-5.00) for patients and a mean value of 3.60 (SD 0.676, range 2.40-5.00) for HCs.

Table 1. Sample characteristics.

Characteristics and categories	HC ^a (n=20)	Schizophrenia (n=20)	Depression (n=20)	Statistical value, <i>F</i> test (<i>df</i>) or chi-square (<i>df</i>)	<i>P</i> value
Demographics					
Age (y), mean (SD)	37.75 (9.94)	43.75 (10.36)	45.05 (15.42)	2.05 (2,57) ^b	.14
Sex, n (%)				0.93 (2) ^c , N=60	.63
Male	10 (50)	12 (60)	9 (45)		
Female	10 (50)	8 (40)	11 (55)		
Treatment setting, n (%)				— ^d	.43 ^e
In-patient	—	13 (65)	12 (60)		
Day clinic	—	1 (5)	4 (20)		
Out-patient	—	6 (30)	4 (20)		
Years of education, mean (SD)	18.90 (3.39)	15.80 (6.38)	14.60 (6.11)	9.86 (2) ^{c,f} , N=60	.007 ^f
BMI (kg/m ²), mean (SD)	24.05 (5.08)	28.24 (7.59)	27.62 (5.45)	6.83 (2) ^{c,f} , N=60	.03 ^f
Living situation, n (%)				—	<.001 ^e
Living alone	7 (35)	17 (85)	4 (20)		
Parents	1 (5)	1 (5)	1 (5)		
Partner	5 (25)	0 (0)	8 (40)		
Sibling or other relatives	4 (20)	0 (0)	2 (10)		
Shared apartment or residential home	3 (15)	2 (10)	5 (25)		
Netto income after deduction of fix expenses, mean (SD)	1080.00 (616.95)	503.65 (424.77)	703.50 (405.08)	12.56 (2) ^{c,f} , N=60	.002 ^f
Marriage status^g, n (%)					
Single	12 (60)	18 (90)	10 (50)	—	.02 ^e
Married	4 (20)	0 (0)	6 (30)	—	.03 ^e
Divorced	2 (10)	2 (10)	3 (15)	—	>.99 ^e
Living separately	1 (5)	0 (0)	0 (0)	—	>.99 ^e
Widow or widower	0 (0)	0 (0)	0 (0)	—	—
Monogamous	1 (5)	0 (0)	2 (10)	—	.77 ^e
Polygamous	1 (5)	0 (0)	0 (0)	—	>.99 ^e
Comorbidities, n (%)				10.58 (2) ^c , N=60	.005
No	15 (75)	8 (40)	5 (25)		
Yes	5 (25)	12 (60)	15 (75)		
Somatic comorbidity^h, n (%)				5.76 (2) ^c , N=60	.06
No	15 (75)	9 (45)	8 (40)		
Yes	5 (25)	11 (55)	12 (60)		
Psychiatric comorbidityⁱ, n (%)				—	.004 ^e
No	20 (100)	19 (95)	13 (65)		
Yes	0 (0)	1 (5)	7 (35)		
Clinical baseline data					
PANSS^l score, mean (SD)					

Characteristics and categories	HC ^a (n=20)	Schizophrenia (n=20)	Depression (n=20)	Statistical value, <i>F</i> test (<i>df</i>) or chi-square (<i>df</i>)	<i>P</i> value
Total	—	98.75 (31.67)	78.55 (19.33)	5.93 (1,31.43) ^{bj}	.02 ^j
Positive symptoms	—	18.85 (7.57)	10.4 (3.05)	21.44 (1,25.013) ^{bj}	<.001 ^j
Negative symptoms	—	18.7 (8.09)	17.5 (7.86)	0.17 (1) ^{c,f} , N=40	.68 ^f
Disorganization	—	22.1 (7.31)	14.35 (3.72)	17.85 (1,28.207) ^{bj}	<.001 ^j
Excitement	—	18.05 (6.99)	13.8 (3.9)	5.64 (1,29.785) ^{bj}	.02 ^j
Emotional distress	—	21.05 (7.48)	22.5 (5.33)	0.50 (1,38) ^b	.48
BDI ^k score, mean (SD)	—	18.45 (8.62)	27.3 (9.83)	9.16 (1,38) ^b	.004

^aHC: healthy control.

^b*F* test.

^cChi-square value.

^dNot applicable.

^eFisher exact test.

^fKruskal-Wallis test.

^gMultichoice option→percentage of n=20.

^hSomatic comorbidity refers to any diagnosis of a somatic disorder.

ⁱPANSS: Positive and Negative Syndrome Scale.

^jWelch-ANOVA.

^kBDI: Beck Depression Inventory.

Hypothesis Outcomes

Hypothesis 1: Patients With a Depressive or Psychotic Disorder and HCs Show Differential Effects of the iVR Intervention on Nature Connectedness

We found a strong increase in nature connectedness on comparing before and after experiences for all groups. Nature connectedness from before to after the iVR experience, INS delta, increased for participants with depression (mean 28.85, SD 26.55), psychotic disorder (mean 32.65, SD 25.86), and HCs (mean 30.70, SD 18.00). The repeated measures ANOVA showed a significant effect of time on changes in nature connectedness with a narrow CI, suggesting a high degree of precision in the effect size estimate ($F_{1,57}=100.12$, $\eta_p^2=0.637$, 95% CI 0.48-0.74; $P<.001$) but no significant interaction of group and time ($F_{2,57}=0.128$, $\eta_p^2=0.004$, 95% CI 0.00-0.11; $P=.88$). Hypothesis 1 has to be rejected.

Hypothesis 2: Patients With a Depressive or Psychotic Disorder and HCs Show Differential Effects of the iVR Intervention on Empathy

The repeated measures ANOVA showed neither significant effects of time on emotional empathy explicit ($F_{1,57}=0.03$, $\eta_p^2=0.001$, 95% CI 0.00-0.06; $P=.86$) or cognitive empathy ($F_{1,57}=0.77$, $\eta_p^2=0.013$, 95% CI 0.0-0.1; $P=.38$) for the 3 groups, nor significant interaction of group and time (emotional empathy explicit: $F_{2,57}=0.68$, $\eta_p^2=0.023$, 95% CI 0.00-0.22; $P=.51$; and cognitive empathy: $F_{2,57}=0.54$, $\eta_p^2=0.018$, 95% CI 0.0-0.2; $P=.59$). Hypothesis 2 has to be rejected.

Hypothesis 3: Patients With a Depressive or Psychotic Disorder and HCs Show Differential Effects of the iVR Intervention on Compassion

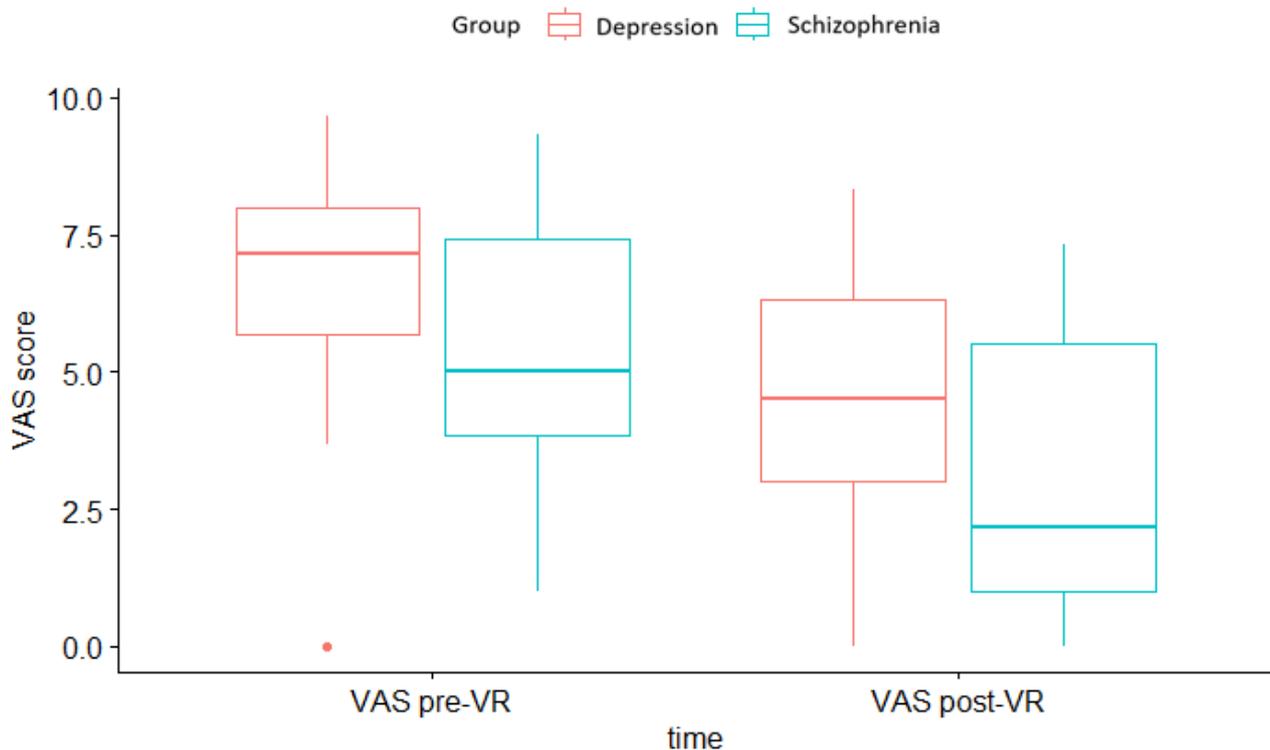
The repeated measures ANOVA revealed a significant effect of time ($F_{1,57}=12.86$, $\eta_p^2=0.18$, 95% CI 0.04-0.36; $P<.001$) for the change in compassion but no significant interaction of group×time ($F_{2,57}=1.48$, $\eta_p^2=0.05$, 95% CI 0.0-0.3; $P=.24$). Hence, compassion increased for all participants after iVR exposure but did not differ significantly between groups. Calculations were repeated after the removal of 3 outliers, showing no significant differences. Hypothesis 3 has to be rejected.

Hypothesis 4: Patients With Psychotic or Depressive Disorder Have a Lower Symptom Burden After the VR Experience Than Before the iVR Experience

Repeated measures ANOVA revealed a significant effect of time—both groups, patients with psychotic and depressive disorders, showed significantly fewer symptoms after iVR exposure ($F_{1,38}=40.93$, $\eta_p^2=0.52$, 95% CI 0.29-0.67; $P<.001$), indicating a large and precisely estimated effect. There was also a significant effect of group ($F_{1,38}=4.68$, $\eta_p^2=0.11$, 95% CI 0.000-0.314; $P=.04$), indicating significant differences between the groups, but no interaction effect of group and time ($F_{1,38}=0.021$, $\eta_p^2=0.001$, 95% CI 0.000-0.074; $P=.89$). Patients diagnosed with depression had a mean VAS score of 6.58 (SD 2.21) before and a mean VAS score of 4.42 (SD 2.50) after the iVR experience. Patients diagnosed with schizophrenia had a mean VAS score of 5.18 (SD 2.31) before and 2.92 (SD 2.52) after iVR. Hypothesis 4 was confirmed. Figure 3 illustrates a

boxplot of pre- and post-VR VAS measures as a measure of symptom burden in both patient groups.

Figure 3. Boxplot of pre- and post-VR VAS measures in both patient groups.



To further explore the influence of the iVR intervention on the change in symptom load of patients with psychosis or depression measured by VAS of the 3 individual main symptoms, we performed an ANCOVA controlling for the covariates, psychotic and depressive symptom severity as assessed with the PANSS, and the BDI before the iVR experience.

The overall model was not statistically significant ($F_{7,32}=1.44$, $R^2=0.239$, 95% CI 0.000-0.364; $P=.23$). Furthermore, 2 covariates, however, emerged as significant—PANSS factor excitement ($\beta=-0.198$, $\eta_p^2=0.127$, 95% CI 0.000-0.353, $\omega^2=0.086$; $P=.03$) and PANSS factor negative symptoms ($\beta=0.165$, $\eta_p^2=0.063$, 95% CI 0.00-0.27, $\omega^2=0.096$; $P=.03$). This suggests that higher levels of excitement symptoms are rather associated with a decrease in change of symptom load (delta VAS), whereas higher levels of negative symptoms are rather associated with an increase in change of symptom load (delta VAS).

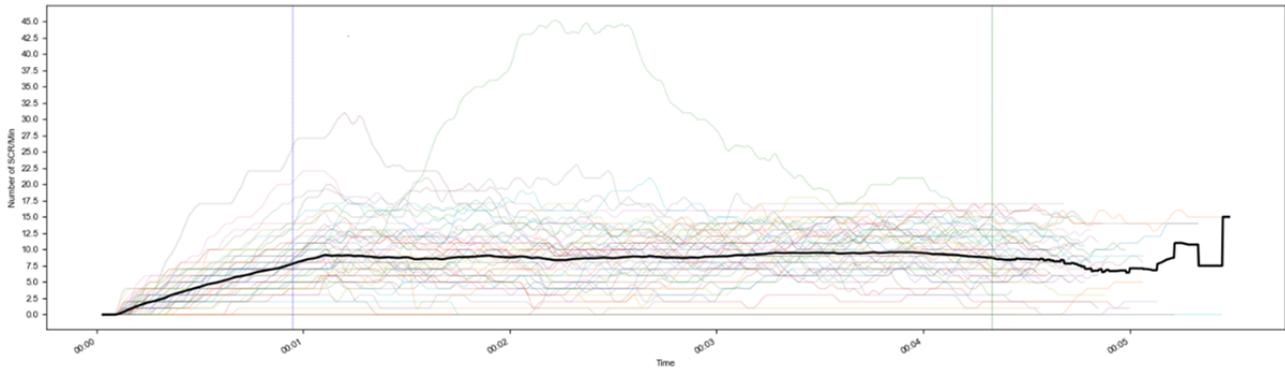
Hypothesis 5: The Measurement of EDA Before, During, and After the iVR Application Allows Emotional Arousal to Be Objectified

In total, 59 participants were tested for emotional arousal. In 1 case, the EDA device did not work properly because it lost

connectivity during the intervention after 3 minutes. In 4 cases, the EDA device recorded no change over time, hinting at a bad connection of the sensor. The average number of counted NS-SCRs at the end of the first full minute in iVR of the remaining 54 participants was interpreted as a baseline (mean 6.92, SD 3.65, 95% CI 5.93-7.90). A relatively low level of physical arousal (mean of the average number of NS-SCRs per minute) during the experience was measured over time (mean 9.30, SD 4.27, 95% CI 8.13-10.05). A paired-samples t test was conducted to examine the difference between the average number of counted NS-SCRs at baseline (point 1) and the mean of NS-SCRs during iVR (point 2). The analysis showed that the difference between point 1 (baseline) and point 2 (during) was statistically significant but with a wide CI (2-tailed $t_{53}=4.40$, $P<.001$, mean difference=2.38, 95% CI 1.30-3.46). EDA before the iVR session was not assessed. Therefore, hypothesis 5 was partly confirmed, as analysis of EDA during iVR showed significantly higher NS-SCR numbers than at the beginning of the iVR application.

In Figure 4, the black line presents the average number of NS-SCRs per minute of all participants, whereas the colored lines are the individual numbers of NS-SCRs per minute. One point in time is marked as the first event when breaking through the soil. The second point in time is marked as an event by means of the end of the scenario.

Figure 4. Number of NS-SCRs per minute.



Hypothesis 6: Spirituality of the Participants, Presence, and Virtual Body Ownership in iVR Have an Influence on Intervention-Induced Changes in Nature Connectedness, Empathy, and Compassion in all 3 Groups

The multiple linear regression analyses revealed that the combination of the 3 predictors (spirituality, presence, and virtual body ownership) did not explain changes in nature connectedness (state; $R^2=0.04$, $F_{3,56}=0.72$, 95% CI 0.00-0.13; $P=.55$).

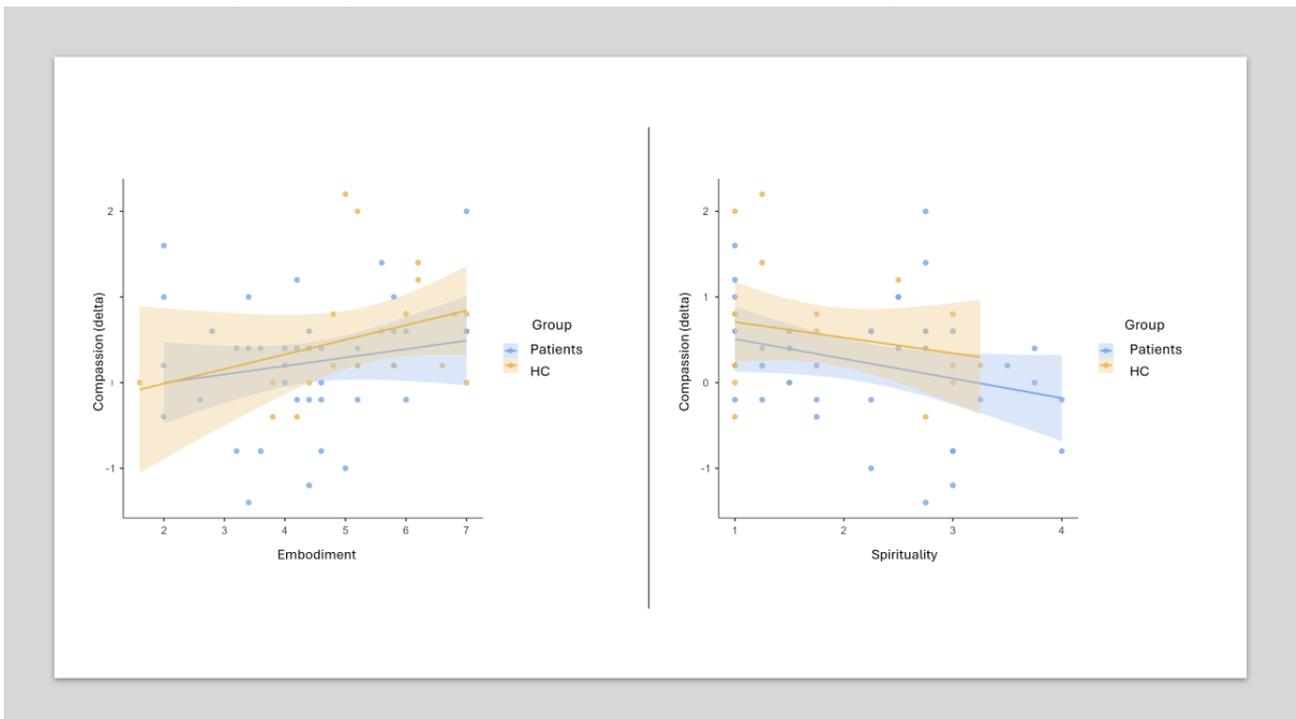
The multiple linear regression analysis also revealed that the combination of the 3 predictors (spirituality, presence, and virtual body ownership) did not explain changes in emotional empathy (state; $R^2=0.06$, $F_{3,56}=1.21$, 95% CI 0.00-0.18; $P=.31$).

Similar results were observed for changes in cognitive empathy. The test revealed that the combination of the 3 predictors (spirituality, presence, and virtual body ownership) did not

explain changes in cognitive empathy (state; $R^2=0.02$, $F_{3,56}=0.382$, 95% CI 0.0-0.09; $P=.77$).

For compassion, the test revealed that the combination of the 3 predictors (spirituality, presence, and virtual body ownership) explained 20.2% of the variance in participants' change in compassion (state; $R^2=0.20$, $F_{3,56}=4.73$, 95% CI 0.04-0.34; $P=.005$). Whereas the predictor variable virtual body ownership was a predictor of participants' change in compassion ($\beta=0.29$, 2-tailed $t_{59}=2.15$, $\eta_p^2=0.08$, 95% CI 0.00-0.23; $P=.04$) and the predictor variable spirituality was a predictor of participants' change in compassion ($\beta=-0.35$, 2-tailed $t_{59}=-2.89$, $\eta_p^2=0.13$, 95% CI 0.01-0.30; $P=.005$), the feeling of presence was not a significant predictor ($\beta=0.08$, 2-tailed $t_{59}=0.6$, $\eta_p^2=0.001$, 95% CI 0.00-0.10; $P=.55$). This means that the changes in compassion were predicted by participants' spirituality and perceived level of virtual body ownership with moderate to large effects, but not by the perceived feeling of presence (Figure 5).

Figure 5. Prediction of changes in the dependent variable compassion by the independent variables spirituality and embodiment.



Discussion

Principal Findings

The study results indicate that engaging with the iVR nature embodiment scenario led to an increase in participants' sense of nature connectedness, as well as a rise in compassion levels across all groups, despite no corresponding change in empathy, and no differential effects between groups (hypotheses 1, 2, and 3). Regarding the effect on symptom load (hypothesis 4), participants reported a reduction in subjective symptom load, as measured by VAS. Complementing these self-reported data, exploratory physiological measures via EDA revealed consistent arousal during the iVR application (hypothesis 5), highlighting that the immersive environment elicited a substantial affective response. Finally, changes in compassion were influenced by individual differences in spirituality and the degree of perceived virtual body ownership (hypothesis 6). Together, these results add further support to the multifaceted effects of iVR interventions on emotional and physiological states.

In general, the investigation of iVR nature exposure on symptom load was an important aim of this study. Here, we could show a significant decrease in individual symptom load in both patient groups, independent of previous symptom burden in the overall model, but significantly influenced by 2 covariates. The significant influence of the 2 PANSS factors, excitement and negative symptoms, calculated according to van der Gaag et al [34], is notable. In view of the known relationship of negative symptoms with poor treatment response [37], this positive association between negative symptoms and change in symptom load gives hope for a potentially effective intervention. Response of different symptom groups to comparable iVR interventions and individual characteristics influencing this response should be further studied, as it is still an insufficiently studied but promising field. Besides the above-mentioned study by Riches et al [16] with $n=42$ who found a decrease in VAS-measured stress, anxiety, and sadness [16], Ilioudi et al [38] conducted a quasi-randomized cross-sectional trial with 60 patients and compared the effects of iVR and physical calm rooms in acute psychiatric settings. Similar to our study, the authors not only use self-report questionnaires but also physiological parameters (blood pressure and heart rate). Of the 60 patients, 40 used the iVR environment, and 20 entered the physical calm room. Both groups showed an improvement in well-being, measured by VAS, after the intervention without showing significant differences between the 2 options [38]. Considering the potential benefits of iVR, as a cost-effective and space-saving tool, this study gives hope for an easily accessible nondrug therapeutic option. In a second study by the authors, this gratefulness of patients about nonpharmacological options was an important result, speaking for the good acceptance of such approaches [15]. In a longitudinal study over 3-4 weeks, Browning et al [39] examined the effect of daily iVR nature experiences on anxiety, depression, and rumination ($n=24$ in the iVR intervention group vs $n=16$ in the control group); however, not in a clinical cohort but in college students. Symptoms of anxious arousal and anxious apprehension decreased, whereas the VR nature experience did not influence anhedonia or rumination,

thereby underlining the importance of specifying different symptom groups in a longitudinal setting [39].

Additionally, we could replicate findings from our earlier investigations [13,17] showing that using the *Tree* application fosters nature connectedness assessed with the INS scale. With this study, more than 200 participants tested the application, and in all studies, a significant increase in INS was found [13,17]. We also replicated the finding of an increase in compassion after exposure found in Spangenberg et al [17], as well as the perceived virtual body ownership of participants being a predictor for compassion. As an embodiment of a tree was an integral component of the used iVR scenario (embodying a tree) and a predictor of the changes in compassion, its role should be further studied with a focus on its potential in body psychotherapy, an established approach in the treatment of severe mental disorders [40]. Psychotic negative symptoms in particular are sometimes understood as a form of disembodiment and alienation from the self and are therefore addressed in body-oriented therapies [41]. Even though the relationship between nature embodiment, especially the feeling of growing virtually in this study, and specific symptoms or symptom load in the long term has not yet been sufficiently researched, the promising results from previous studies and this study provide grounds for more detailed investigations.

In this study, we also assessed empathy and observed that it is rather compassion than emotional or cognitive empathy that is elicited by the perceived virtual body ownership experience [17]. Again, we were able to support earlier findings that while compassion has some overlaps with the concept of empathic concern, it is recognized as a distinct emotion, separate from empathy, and is associated with activation in different brain regions [42-44]. Although different impairments of empathy are described in the 2 examined disorders [18,19], we could not find any differences regarding the change of empathy through the exposure between the 3 groups in our study.

Limitations

Several limitations should be acknowledged when interpreting the findings of this study. First, our data were derived from an explorative study with a relatively small sample size, which restricts the generalizability of our results. Thus, further research with larger sample sizes is needed to assess the effects of iVR exposure on emotional or clinical aspects, for example, different categories of individual symptoms. The groups in this study were too small for the latter. Second, singular interventions do not allow for drawing conclusions about the long-term impact of this intervention. Longitudinal studies are necessary to clarify whether these effects change over time. Third, we did not ask participants' previous iVR experience, which may influence observed outcomes. Fourth, we did not include a control group (eg, a different VR scenario) to differentiate between the effects of nature embodiment in iVR and the general effects of iVR, which should be supplemented in future studies. Fifth, the absence of systematic psychological side effect monitoring is due to the exploratory design of this study, but should be addressed in the following studies. Finally, the quality of EDA data collected was limited, likely due to issues related to the lack of a pre-VR baseline, sensor placement, signal noise, or

participants' movement. Even though the decision to not integrate a pre-VR baseline measure of EDA was made in order to limit the time commitment of the patients in this explorative study, future studies should integrate a pre-exposure baseline and separate tonic and phasic activity as recommended [45,46]. Larger samples might be essential for obtaining more reliable physiological measures.

Conclusions

To conclude, the experience of embodying a tree in iVR had immediate benefits, as indicated by positive effects on individual symptoms, an increase in participants' sense of nature connectedness, and a rise in compassion levels. It can be highlighted that patients with psychiatric conditions during inpatient treatment, including patients diagnosed with psychosis, were able to receive an iVR-based intervention and profited from the experience. Despite a still high technological effort to

implement VR applications into therapeutic sessions, our study suggests that even a short exposure to becoming a virtual tree can elicit a strong affective response in patients. Given that some clinical facilities in urban areas offer little opportunity for recreation in natural surroundings, and that patients with acute psychiatric disorders are sometimes not able to seek out these surroundings on their own, the use of virtual nature experiences appears to be a promising approach. While these findings are a starting point, long-term effects in follow-up examinations remain to be investigated. Additional research examining sustained outcomes, different intervention dosages (longer or repeated sessions), sham or alternative VR control conditions, as well as the integration of such virtual embodiment practices into established body therapy interventions, would help determine whether this innovative form of iVR exposure can become an effective and easily accessible nonpharmacological tool for patients with severe mental disorders.

Acknowledgments

AL is a participant in the BIH Charité Junior Digital Clinician Scientist Program funded by the Charité – Universitätsmedizin Berlin, and the Berlin Institute of Health at Charité (BIH).

The authors disclose that generative artificial intelligence was not used in writing this manuscript.

Data Availability

The source data used in this research are currently unavailable for public sharing due to strict data safety and confidentiality protocols mandated by our university. These restrictions are in place to ensure compliance with ethical standards, privacy regulations, and institutional policies.

Access to the data may be granted under specific circumstances, subject to appropriate data use agreements and ethical approvals. For more information about the data, please contact the corresponding author.

Funding

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Authors' Contributions

AL and PS were responsible for concept and design of the study, acquisition of original data, statistical analysis, the interpretation of data as well as the first draft. AL took the lead in writing the manuscript. KL was responsible for data acquisition and statistical analysis. CM was responsible for the concept and design of the study, statistical analysis, interpretation of data, and revised the first draft. SG was responsible for the concept and design of the study, interpretation of data, and revised the first draft. GFR was responsible for processing, analysis, and interpretation of electrodermal activity data. FB revised the concept and design of the study, was responsible for interpretation of data, and revised the first draft. SN contributed to the first draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional information on statistical analysis.

[[DOCX File, 63 KB - games_v14i1e74337_app1.docx](#)]

References

1. Bell IH, Pot-Kolder R, Rizzo A, Rus-Calafell M, Cardi V, Cella M, et al. Advances in the use of virtual reality to treat mental health conditions. *Nat Rev Psychol* 2024;3(8):552-567. [doi: [10.1038/s44159-024-00334-9](https://doi.org/10.1038/s44159-024-00334-9)]
2. Grassini S. A systematic review and meta-analysis of nature walk as an intervention for anxiety and depression. *J Clin Med* 2022;11(6):1731 [FREE Full text] [doi: [10.3390/jcm11061731](https://doi.org/10.3390/jcm11061731)] [Medline: [35330055](https://pubmed.ncbi.nlm.nih.gov/35330055/)]
3. McMahan EA, Estes D. The effect of contact with natural environments on positive and negative affect: A meta-analysis. *The Journal of Positive Psychology* 2015;10(6):507-519. [doi: [10.1080/17439760.2014.994224](https://doi.org/10.1080/17439760.2014.994224)]

4. Paredes-Céspedes DM, Vélez N, Parada-López A, Toloza-Pérez YG, Téllez EM, Portilla C, et al. The effects of nature exposure therapies on stress, depression, and anxiety levels: A systematic review. *Eur J Investig Health Psychol Educ* 2024;14(3):609-622 [FREE Full text] [doi: [10.3390/ejihpe14030040](https://doi.org/10.3390/ejihpe14030040)] [Medline: [38534901](https://pubmed.ncbi.nlm.nih.gov/38534901/)]
5. United Nations. *The World's Cities in 2018*. New York City: United Nations; 2018.
6. Wiebe A, Kannen K, Selaskowski B, Mehren A, Thöne AK, Pramme L, et al. Virtual reality in the diagnostic and therapy for mental disorders: A systematic review. *Clin Psychol Rev* 2022;98:102213 [FREE Full text] [doi: [10.1016/j.cpr.2022.102213](https://doi.org/10.1016/j.cpr.2022.102213)] [Medline: [36356351](https://pubmed.ncbi.nlm.nih.gov/36356351/)]
7. Maheta B, Kraft A, Interrante N, Fereydooni S, Bailenson J, Beams B, et al. Using virtual reality to improve outcomes related to quality of life among older adults with serious illnesses: Systematic review of randomized controlled trials. *J Med Internet Res* 2025;27:e54452 [FREE Full text] [doi: [10.2196/54452](https://doi.org/10.2196/54452)] [Medline: [40009834](https://pubmed.ncbi.nlm.nih.gov/40009834/)]
8. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. *Front. Robot. AI* 2016;3. [doi: [10.3389/frobt.2016.00074](https://doi.org/10.3389/frobt.2016.00074)]
9. Kilteni K, Groten R, Slater M. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* 2012;21(4):373-387. [doi: [10.1162/pres_a.00124](https://doi.org/10.1162/pres_a.00124)]
10. Browning MHEM, Shipley N, McAnirlin O, Becker D, Yu C, Hartig T, et al. An actual natural setting improves mood better than its virtual counterpart: A meta-analysis of experimental data. *Front Psychol* 2020;11:2200 [FREE Full text] [doi: [10.3389/fpsyg.2020.02200](https://doi.org/10.3389/fpsyg.2020.02200)] [Medline: [33101104](https://pubmed.ncbi.nlm.nih.gov/33101104/)]
11. Sarcinella ED, Chirico A, Mancuso V, Pizzolante M, Murtas V, Lauro V, et al. How many degrees of emotional separation are there between real and virtual nature? A pilot study. *Annual Review of CyberTherapy and Telemedicine* 2023;21:47-52.
12. Newman M, Gatersleben B, Wyles K, Ratcliffe E. The use of virtual reality in environment experiences and the importance of realism. *Journal of Environmental Psychology* 2022;79:101733. [doi: [10.1016/j.jenvp.2021.101733](https://doi.org/10.1016/j.jenvp.2021.101733)]
13. Spangenberg P, Freytag SC, Geiger S. Embodying nature in immersive virtual reality: Are multisensory stimuli vital to affect nature connectedness and pro-environmental behaviour? *Computers & Education* 2024;212:104964 [FREE Full text] [doi: [10.1016/j.compedu.2023.104964](https://doi.org/10.1016/j.compedu.2023.104964)]
14. Spangenberg P, Geiger SM, Freytag S. Becoming nature: Effects of embodying a tree in immersive virtual reality on nature relatedness. *Sci Rep* 2022;12(1):1311 [FREE Full text] [doi: [10.1038/s41598-022-05184-0](https://doi.org/10.1038/s41598-022-05184-0)] [Medline: [35079051](https://pubmed.ncbi.nlm.nih.gov/35079051/)]
15. Ilioudi M, Wallström S, Steingrimsson S, Lindner P, Thunström AO, Ali L. Patient experience of a virtual reality calm room in a psychiatric inpatient care setting in Sweden: A qualitative study with inpatients. *BMJ Open* 2023;13(12):e076285 [FREE Full text] [doi: [10.1136/bmjopen-2023-076285](https://doi.org/10.1136/bmjopen-2023-076285)] [Medline: [38101823](https://pubmed.ncbi.nlm.nih.gov/38101823/)]
16. Riches S, Nicholson S, Fialho C, Little J, Ahmed L, McIntosh H, et al. Integrating a virtual reality relaxation clinic within acute psychiatric services: A pilot study. *Psychiatry Res* 2023;329:115477 [FREE Full text] [doi: [10.1016/j.psychres.2023.115477](https://doi.org/10.1016/j.psychres.2023.115477)] [Medline: [37802013](https://pubmed.ncbi.nlm.nih.gov/37802013/)]
17. Spangenberg P, Krüger JM, Geiger SM, Reuth GF, Baumann L, Nebel S. Compassion is key: How virtually embodying nature increases connectedness to nature. *Journal of Environmental Psychology* 2025;102:102521. [doi: [10.1016/j.jenvp.2025.102521](https://doi.org/10.1016/j.jenvp.2025.102521)]
18. Lehmann A, Bağçesular K, Brockmann EM, Biederbick SE, Dziobek I, Gallinat J, et al. Subjective experience of emotions and emotional empathy in paranoid schizophrenia. *Psychiatry Res* 2014;220(3):825-833. [doi: [10.1016/j.psychres.2014.09.009](https://doi.org/10.1016/j.psychres.2014.09.009)] [Medline: [25288043](https://pubmed.ncbi.nlm.nih.gov/25288043/)]
19. Guhn A, Merkel L, Hübner L, Dziobek I, Sterzer P, Köhler S. Understanding versus feeling the emotions of others: How persistent and recurrent depression affect empathy. *J Psychiatr Res* 2020;130:120-127. [doi: [10.1016/j.jpsychires.2020.06.023](https://doi.org/10.1016/j.jpsychires.2020.06.023)] [Medline: [32805521](https://pubmed.ncbi.nlm.nih.gov/32805521/)]
20. Winter D, Malighetti C, Cipolletta S, Ahmed S, Benson B, Röhrich F. Construing and body dissatisfaction in chronic depression: A study of body psychotherapy. *Psychiatry Res* 2018;270:845-851. [doi: [10.1016/j.psychres.2018.10.061](https://doi.org/10.1016/j.psychres.2018.10.061)] [Medline: [30551334](https://pubmed.ncbi.nlm.nih.gov/30551334/)]
21. Liu X. *Inward to Outward*. Cambridge, Massachusetts: Massachusetts Institute of Technology; 2017.
22. Tree. New Reality Co. URL: <https://www.treeofficial.com/> [accessed 2026-02-20]
23. Kleespies MW, Braun T, Dierkes PW, Wenzel V. Measuring connection to nature—A illustrated extension of the inclusion of nature in self scale. *Sustainability* 2021;13(4):1761. [doi: [10.3390/su13041761](https://doi.org/10.3390/su13041761)]
24. Dziobek I, Rogers K, Fleck S, Bahnemann M, Heekeren HR, Wolf OT, et al. Dissociation of cognitive and emotional empathy in adults with asperger syndrome using the multifaceted empathy test (MET). *J Autism Dev Disord* 2008;38(3):464-473. [doi: [10.1007/s10803-007-0486-x](https://doi.org/10.1007/s10803-007-0486-x)] [Medline: [17990089](https://pubmed.ncbi.nlm.nih.gov/17990089/)]
25. Pfattheicher S, Sassenrath C, Schindler S. Feelings for the suffering of others and the environment. *Environment and Behavior* 2015;48(7):929-945. [doi: [10.1177/0013916515574549](https://doi.org/10.1177/0013916515574549)]
26. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 1987;13(2):261-276. [doi: [10.1093/schbul/13.2.261](https://doi.org/10.1093/schbul/13.2.261)] [Medline: [3616518](https://pubmed.ncbi.nlm.nih.gov/3616518/)]
27. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:561-571. [doi: [10.1001/archpsyc.1961.01710120031004](https://doi.org/10.1001/archpsyc.1961.01710120031004)] [Medline: [13688369](https://pubmed.ncbi.nlm.nih.gov/13688369/)]

28. First MB, Gibbon M. The structured clinical interview for DSM-IV axis I disorders (SCID-I) and the structured clinical interview for DSM-IV axis II disorders (SCID-II). In: *Comprehensive Handbook of Psychological Assessment, Vol 2: Personality Assessment*. Hoboken, NJ, US: John Wiley & Sons, Inc; 2004:134-143.
29. Kim HK, Park J, Choi Y, Choe M. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Appl Ergon* 2018;69:66-73. [doi: [10.1016/j.apergo.2017.12.016](https://doi.org/10.1016/j.apergo.2017.12.016)] [Medline: [29477332](https://pubmed.ncbi.nlm.nih.gov/29477332/)]
30. Ahn SJG, Bostick J, Ogle E, Nowak KL, McGillicuddy KT, Bailenson JN. Experiencing nature: Embodying animals in immersive virtual environments increases inclusion of nature in self and involvement with nature. *J Comput-Mediat Comm* 2016;21(6):399-419. [doi: [10.1111/jcc4.12173](https://doi.org/10.1111/jcc4.12173)]
31. Slater M, Spanlang B, Sanchez-Vives MV, Blanke O. First person experience of body transfer in virtual reality. *PLoS One* 2010;5(5):e10564 [FREE Full text] [doi: [10.1371/journal.pone.0010564](https://doi.org/10.1371/journal.pone.0010564)] [Medline: [20485681](https://pubmed.ncbi.nlm.nih.gov/20485681/)]
32. Hampel ND, Brähler E, Zenger M, Albani C, Heuft G. Psychometrische Überprüfung der kurzversion der skala "Transpersonales Vertrauen" in einer deutschlandweiten repräsentativen bevölkerungsstichprobe. *Psychother Psychosom Med Psychol* 2019;70(06):237-245. [doi: [10.1055/a-1033-7627](https://doi.org/10.1055/a-1033-7627)]
33. Bailenson JN, Swinik K, Hoyt C, Persky S, Dimov A, Blascovich J. The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments* 2005;14(4):379-393. [doi: [10.1162/105474605774785235](https://doi.org/10.1162/105474605774785235)]
34. van der Gaag M, Hoffman T, Remijnsen M, Hijman R, de Haan L, van Meijel B, et al. The five-factor model of the positive and negative syndrome scale II: A ten-fold cross-validation of a revised model. *Schizophr Res* 2006;85(1-3):280-287. [doi: [10.1016/j.schres.2006.03.021](https://doi.org/10.1016/j.schres.2006.03.021)] [Medline: [16730429](https://pubmed.ncbi.nlm.nih.gov/16730429/)]
35. Field AP. *Discovering Statistics Using IBM SPSS Statistics: And Sex and Drugs and Rock 'n' Roll*. Thousand Oaks, CA: SAGE Publications, Limited; 2015.
36. 7.4.7.1. Tukey's method. NIST. URL: <https://www.itl.nist.gov/div898/handbook/prc/section4/prc471.htm> [accessed 2026-03-14]
37. Correll CU, Schooler NR. Negative symptoms in schizophrenia: A review and clinical guide for recognition, assessment, and treatment. *Neuropsychiatr Dis Treat* 2020;16:519-534 [FREE Full text] [doi: [10.2147/NDT.S225643](https://doi.org/10.2147/NDT.S225643)] [Medline: [32110026](https://pubmed.ncbi.nlm.nih.gov/32110026/)]
38. Ilioudi M, Lindner P, Ali L, Wallström S, Thunström AO, Ioannou M, et al. Physical versus virtual reality-based calm rooms for psychiatric inpatients: Quasi-randomized trial. *J Med Internet Res* 2023;25:e42365 [FREE Full text] [doi: [10.2196/42365](https://doi.org/10.2196/42365)] [Medline: [37204858](https://pubmed.ncbi.nlm.nih.gov/37204858/)]
39. Browning MHEM, Shin S, Drong G, McAnirlin O, Gagnon RJ, Ranganathan S, et al. Daily exposure to virtual nature reduces symptoms of anxiety in college students. *Sci Rep* 2023;13(1):1239 [FREE Full text] [doi: [10.1038/s41598-023-28070-9](https://doi.org/10.1038/s41598-023-28070-9)] [Medline: [36690698](https://pubmed.ncbi.nlm.nih.gov/36690698/)]
40. Röhrich F. Body psychotherapy for the treatment of severe mental disorders – an overview. *Body, Movement and Dance in Psychotherapy* 2014;10(1):51-67. [doi: [10.1080/17432979.2014.962093](https://doi.org/10.1080/17432979.2014.962093)]
41. Martin LAL, Koch SC, Hirjak D, Fuchs T. Overcoming disembodiment: The effect of movement therapy on negative symptoms in schizophrenia-A multicenter randomized controlled trial. *Front Psychol* 2016;7:483 [FREE Full text] [doi: [10.3389/fpsyg.2016.00483](https://doi.org/10.3389/fpsyg.2016.00483)] [Medline: [27064347](https://pubmed.ncbi.nlm.nih.gov/27064347/)]
42. Bloom P. *Against Empathy: The Case for Rational Compassion*. New York: HarperCollins; 2016.
43. Klimecki OM, Leiberg S, Ricard M, Singer T. Differential pattern of functional brain plasticity after compassion and empathy training. *Soc Cogn Affect Neurosci* 2014;9(6):873-879 [FREE Full text] [doi: [10.1093/scan/nst060](https://doi.org/10.1093/scan/nst060)] [Medline: [23576808](https://pubmed.ncbi.nlm.nih.gov/23576808/)]
44. Strauss C, Lever Taylor B, Gu J, Kuyken W, Baer R, Jones F, et al. What is compassion and how can we measure it? A review of definitions and measures. *Clin Psychol Rev* 2016;47:15-27 [FREE Full text] [doi: [10.1016/j.cpr.2016.05.004](https://doi.org/10.1016/j.cpr.2016.05.004)] [Medline: [27267346](https://pubmed.ncbi.nlm.nih.gov/27267346/)]
45. Horvers A, Tombeng N, Bosse T, Lazonder AW, Molenaar I. Detecting emotions through electrodermal activity in learning contexts: A systematic review. *Sensors (Basel)* 2021;21(23):7869 [FREE Full text] [doi: [10.3390/s21237869](https://doi.org/10.3390/s21237869)] [Medline: [34883870](https://pubmed.ncbi.nlm.nih.gov/34883870/)]
46. Wang D, Peng Y, Haddouk L, Vayatis N, Vidal P. Assessing virtual reality presence through physiological measures: A comprehensive review. *Front. Virtual Real* 2025;6. [doi: [10.3389/frvir.2025.1530770](https://doi.org/10.3389/frvir.2025.1530770)]

Abbreviations

BDI: Beck Depression Inventory

DSM-IV: Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)

EDA: electrodermal activity

HC: healthy controls

HMD: head-mounted display

INS: Inclusion of Nature in Self

iVR: immersive virtual reality

MET: Multifaceted Empathy Test

NS-SCR: nonspecific skin conductance response

PANAS: Positive and Negative Affect Schedule

PANSS: Positive and Negative Syndrome Scale

SCID I: Structured Clinical Interview for DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition)

VAS: visual analog scale

VR: virtual reality

Edited by S Brini; submitted 22.Mar.2025; peer-reviewed by J Swidrak, J Ferrer Costa; comments to author 30.Sep.2025; revised version received 07.Feb.2026; accepted 09.Feb.2026; published 24.Mar.2026.

Please cite as:

Lütt A, Lütgens K, Montag C, Gutwinski S, Reuth GF, Nebel S, BERPohl F, Spangenberg P

Effects of Virtual Nature Embodiment on Compassion, Empathy, Nature Connectedness, and Symptom Burden in Patients With Psychosis or Depression: An Explorative Clinical Study

JMIR Serious Games 2026;14:e74337

URL: <https://games.jmir.org/2026/1/e74337>

doi: [10.2196/74337](https://doi.org/10.2196/74337)

PMID: [41874554](https://pubmed.ncbi.nlm.nih.gov/41874554/)

©Alva Lütt, Kristina Lütgens, Christiane Montag, Stefan Gutwinski, Georg Felix Reuth, Steve Nebel, Felix BERPohl, Pia Spangenberg. Originally published in JMIR Serious Games (<https://games.jmir.org>), 24.Mar.2026. 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.

Original Paper

Mental Health Professionals' Views on Gaming to Inform Game-Based Interventions: Qualitative Cross-Sectional Study

Lauri Lukka¹, DSC; Veli-Matti Karhulahti², PhD; J Matias Palva^{1,3,4}, PhD

¹Department of Neuroscience and Biomedical Engineering, School of Science, Aalto University, Espoo, Finland

²Department of Music, Art and Culture Studies, University of Jyväskylä, Jyväskylä, Finland

³Neuroscience Center, Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland

⁴Centre for Cognitive Neuroimaging, School of Psychology and Neuroscience, University of Glasgow, Glasgow, United Kingdom

Corresponding Author:

Lauri Lukka, DSC

Department of Neuroscience and Biomedical Engineering

School of Science

Aalto University

Rakentajanaukio 2

Espoo, 02150

Finland

Phone: 358 440375666

Email: lauri.lukka@aalto.fi

Abstract

Background: Few game-based digital mental health interventions have been adopted in clinical practice, where mental health professionals (MHPs) play a critical role in the uptake of new technologies. Existing evidence suggests that MHPs' views on game-based interventions and entertainment video games are mixed, reflecting broader tensions surrounding video games, which are perceived as both harmful and beneficial. However, the underlying reasons for these perceptions have remained unclear, even though they may substantially influence MHPs' willingness to adopt or refuse new clinical solutions.

Objective: This qualitative cross-sectional study investigates how MHPs view entertainment video games and gaming in clinical contexts. By examining these perceptions, the study aims to inform the implementation of game-based digital mental health interventions in health care.

Methods: This study combined 3 qualitative interview datasets (n=19, n=16, and n=6) capturing Finnish MHPs' views on video games and gaming, resulting in a combined sample of 41 participants, of whom 56% (n=23) were women and 59% (n=24) reported playing games. The interview data were analyzed using reflexive thematic analysis. Additionally, 2 post hoc analyses were conducted with complementary qualitative questionnaire data (n=80) collected alongside the first dataset.

Results: In total, 3 themes were generated to reflect the MHPs' views. First, personal recreation, clinically harmful: MHPs demonstrated a self-client attitude asymmetry, describing their own gaming primarily as positive and recreational, while associating clients' gaming with problems. Second, adverse technology and meaningful culture: MHPs expressed attitudinal ambivalence, making sense of gaming through conflicting frames as both potentially harmful technology and meaningful culture. Third, holistic exploration of clients' gaming: MHPs evaluated their clients' gaming within the broader context of the clients' lives. The first post hoc analysis reinforced the observed self-client attitude asymmetry, showing that MHPs associated their own gaming experiences with more positive and fewer negative meanings compared to their clients' gaming. The second post hoc analysis suggested that MHPs expected game-based interventions to be approachable, motivating, and complementary to other treatments, and particularly suited for children and youth, people with strong digital competencies, and clients who were withdrawn.

Conclusions: Existing qualitative research on MHPs' perceptions of video games remains limited and has not examined clinicians' broader sense-making of gaming to inform the implementation of game-based interventions. This study identifies how clinician education can build on MHPs' existing perceptions while addressing potential misconceptions by clearly differentiating game-based interventions from entertainment games, clarifying their clinical aims and mechanisms of action, situating them within clients' broader care and recovery processes, and challenging narrow assumptions about their target audience. Together, these contributions address a critical gap in the literature and inform strategies to improve clinician education, communication, and the adoption of game-based interventions in mental health care.

KEYWORDS

clinical practice; digital interventions; mental health; mental health professionals; implementation science; qualitative study; serious games; thematic analysis; user-centered design; video games

Introduction

Game-Based Interventions as a New Treatment Modality

Game-based digital mental health interventions use elements from entertainment games to achieve health-related aims [1]. They include both fully fledged serious games and gamified interventions [2], and here we focus on the former category, referred to as game-based interventions. The rationale for designing such interventions includes increasing intervention appeal to facilitate treatment access [3], increasing user engagement with digital interventions to alleviate the well-known challenge of low user engagement [4], and establishing new modes of treatment to complement psychological interventions and pharmacotherapy, which do not help all patients [5,6]. Preliminary evidence supports the efficacy of game-based interventions for depression [7], anxiety [8], and other mental disorders [9,10]. Game-based interventions include, for instance, SPARX for adolescent depression [11]; StarStarter for social anxiety disorder [12]; Embers the Dragon, a child-parent intervention supporting children's emotional development [13]; and EndeavorRx, which has received Food and Drug Administration authorization for the treatment of pediatric attention-deficit/hyperactivity disorder [14]. However, most interventions are still in the research and development stage, few commercial game-based therapeutics are available, and few have been broadly or routinely integrated into clinical practice [15]. Best practices around their use are thus only emerging. The translation from the controlled research environments to real-world settings remains a central challenge for many digital interventions [16], including game-based interventions.

Beyond clinical efficacy, the success of game-based interventions depends on their acceptability [17-19]. The evaluation of acceptability has typically focused on end users [18], but in health care, clinical stakeholders also hold sway; they act as gatekeepers who recommend and prescribe digital interventions [20]. As health care providers' attitudes and consequent behavior can make or break the implementation of novel digital tools [21,22], it is vital to understand their perceptions and consider them in intervention implementation [23-27]. This is particularly significant considering that game-based interventions suggest a change of meaning [28]—what looks like an entertainment video game can be therapeutic. Such a category shift offers opportunities for reconceptualizing and renewing mental health interventions [29], but it can also provoke hesitancy and resistance. Our aim in this study is to examine mental health professionals' (MHPs') views on entertainment video games and gaming in clinical contexts to inform the implementation of game-based interventions.

MHPs' Views on Games Are Mixed

Game-based interventions are still uncommon in clinical practice, and there are few studies examining MHPs' views on them. Surveys indicate that only 12%-16% of psychotherapists know of the existence of such treatments, and only 1%-7% use them [30-32], while there are no data on the uptake of game-based interventions in the regional context of this study, Finland. Game-based interventions are perceived to offer opportunities for clients to engage with therapeutic content, foster self-efficacy and responsibility, and improve therapy access and patient motivation [31]. However, they are also viewed as potentially increasing avoidance and isolation, promoting addictive gambling habits, and devaluing therapy. Moreover, another study found that 5% of psychotherapists exhibit unfavorable and 10% highly favorable attitudes toward game-based interventions [32]. However, the previous studies do not capture the reasons for MHPs' sentiments, which are particularly important to understand when implementing novel solutions in health care contexts.

MHPs' mixed views on game-based interventions may arise from their perceptions of the underlying media, namely, video games. They may associate video games with addictive and violence-inciting potential, while simultaneously being open to using them in clinical contexts [33]. MHPs recognize both disadvantages and benefits related to gaming: "*Paradoxically, it seems that the way gaming benefits one individual may disrupt another and vice versa,*" summarize Alho and Mankinen [34]. Such conflicting attitudes mirror lay perspectives regarding video games [33,35] that they are associated with aggressive [36,37] and addictive behavior [38,39], while also being popular entertainment [40,41] and linked to positive effects on affective well-being [42], cognitive skills [43,44], and even intelligence [45,46]. Przybylski [47] encapsulates this attitudinal tension concisely: "*Whether electronic games are on balance a good or bad thing remains an open question, but is a topic about which people hold strong opinions.*" This contradiction may be a key distinction between MHPs' views on game-based and non-game-based interventions and underscores the need for further qualitative research to explore the tensions MHPs may perceive. Because most Finnish MHPs do not have first-hand experience with game-based interventions, this study focuses on their views of entertainment video games, the medium on which these interventions are based, and the activity of playing entertainment video games, that is, gaming.

Research Question

The research question guiding this study is as follows: How do MHPs view entertainment video games and gaming in clinical contexts? By examining these perceptions, the study aims to advance understanding of gaming in clinical contexts for researchers and designers involved in the creation, evaluation,

and implementation of game-based digital mental health interventions in health care.

Methods

Research Design Overview

This qualitative cross-sectional study investigated the views of Finnish MHPs on entertainment video games and gaming. This aimed to inform the implementation of game-based interventions, where designers and developers interact with health care professionals and need to be mindful of their perceptions. Aligned with this aim, the research paradigm was interpretive [48], seeking to understand the meanings MHPs assign to video games and gaming. The study data included 3 distinct interview datasets, which were gathered independently by Lukka et al (dataset 1, n=19) [20], Alho and Mankinen (dataset 2, n=16) [49], and Siuttila et al (dataset 3, n=6) [50], for a total of 41 MHP interviews. All 3 datasets concerned MHPs' views on video games but used slightly different interview guides as described further in this study in detail. The combined datasets were analyzed using reflexive thematic analysis (RTA) [51,52], which led to the generation of 3 themes. After the interview data analysis, 2 post hoc analyses were conducted using questionnaire data (n=80) that were gathered with dataset 1.

Ethical Considerations

The study and data gathering for dataset 1 were reviewed and approved by the Aalto University research ethics committee (D/508/03.04/2022), and the study was preregistered on the Open Science Framework [53]. The study involving the gathering of dataset 2 did not involve any of the 6 criteria that lead to ethics review in Finland, and therefore, ethics review was not required or applied for [54]. The study gathering dataset 3 was reviewed and approved by the Human Sciences and Ethics Committee of the University of Jyväskylä (80/13.00.04.00/2021).

Dataset 1 is stored securely on Aalto University servers with restricted access to only the researchers of this study, while datasets 2 and 3 are anonymized and stored securely in the Finnish Social Science Data Archive (FSD). The dataset 2 has been published in FSD with permission for secondary analysis [49], and the dataset reuse [55] was justified by the rationale that this study was guided by a different and more interpretive research question. The dataset 3 has also been published in FSD with permission for secondary analysis [50]. The dataset 3 reuse was justified because it was earlier analyzed via completely different research questions and analytical process, which focused primarily on players rather than MHPs.

In all 3 studies, the participants provided informed consent in Finnish. For studies generating datasets 2 and 3, the original informed consent allows for the secondary analysis without additional consent. No compensation was provided to the participants. All data have been reported in a manner that ensures no individual participant can be identified from the study results.

Study Participants

The study participants were Finnish MHPs or psychology students (Table 1). Dataset 1 was gathered to study the role digital tools and video games had on Finnish MHPs' clinical practices. Dataset 2 was gathered to study clinical psychologists' and psychology students' attitudes toward video games, problematic gaming, and gaming disorder [34]. The second author acted as the supervisor for this work. Dataset 3 was gathered to study the difference between 2 types of intensive play—one related to treatment-seeking and the other related to thriving well-being. The research effort was led by the second author (VMK) [56]. The primary participants in the study were people who played digital games actively and had sought help for their gaming-related problems, and those who actively played esports without health problems. These 2 samples of video game players were contextualized with interviews from Finnish MHPs (n=6) who had experience in treating at least 1 client with gaming-related health problems (3 MHPs were therapists, others were clinical psychologists).

Table 1. The demographic data of study participants within and across the three datasets.

Characteristic	Dataset 1	Dataset 2	Dataset 3	All datasets	
	Interviewees 1-19 (MH- Ps ^a , n=19), n (%)	Interviewees 20-25 (psy- chology students, n=6), n (%)	Interviewees 26-35 (psy- chologists, n=10), n (%)	Interviewees 36-41 (MHPs, n=6), n (%)	Interviewees 1-41 (n=41), n (%)
Gender					
Woman	13 (68)	3 (50)	6 (60)	1 (17)	23 (56)
Man	6 (32)	3 (50)	4 (40)	5 (83)	18 (44)
Age (y)					
18-29	1 (5)	0 (0)	4 (40)	0 (0)	5 (12)
30-39	3 (16)	0 (0)	4 (40)	0 (0)	7 (17)
40-49	6 (32)	0 (0)	1 (10)	0 (0)	7 (17)
50-59	7 (37)	0 (0)	1 (10)	0 (0)	8 (20)
60-66	2 (11)	0 (0)	0 (0)	0 (0)	2 (5)
Unknown	0 (0)	6 (100)	0 (0)	6 (100)	12 (29)
Years of working experience					
0	0 (0)	6 (100)	0 (0)	0 (0)	6 (15)
<5	4 (21)	0 (0)	7 (70)	0 (0)	11 (27)
5-10	3 (16)	0 (0)	1 (10)	0 (0)	4 (10)
10-15	2 (11)	0 (0)	1 (10)	0 (0)	3 (7)
>15	10 (53)	0 (0)	1 (10)	0 (0)	11 (27)
Unknown	0 (0)	0 (0)	0 (0)	6 (100)	6 (15)
Plays games					
Yes	9 (47)	3 (50)	6 (60)	6 (100)	24 (59)
No	10 (53)	3 (50)	4 (40)	0 (0)	17 (41)

^aMHP: mental health professional.

Participant Recruitment

For dataset 1, the study participants were recruited using a questionnaire distributed through professional associations, health care organizations, social media, and snowballing. A total of 80 MHPs completed an online questionnaire; 46 (58%) wanted to participate only in the questionnaire, and 34 (43%) also in the interview. The first author (LL) contacted 24 MHPs in the latter group via email, primarily in registration order, and received 21 (88%) responses, of whom 2 withdrew before the interview. For dataset 2, the study recruitment occurred in 2 parts. First, psychology students were recruited through the university email list, which led to 8 responses, of whom 6 were interviewed. Then, gaming and nongaming licensed psychologists were recruited using social media and emails. For dataset 3, the participants were recruited through health care institutions and snowballing.

Data Collection

For dataset 1, the first author (LL) conducted the semistructured interviews [57] with the initial aim of examining MHPs' views and needs regarding game-based digital interventions. The researcher's preinterview assumptions and perspectives are described in the study preregistration [53]. The original interview guide consisted of 4 sections—MHP background,

views on digital interventions, views on digital games, and views and needs concerning game-based digital interventions. During the early interviews, it became evident that participants had very limited first-hand experience with digital interventions or game-based digital interventions. Consequently, the interview guide was revised (available in the study by Lukka et al [20]) to focus on MHP background, digital tools more broadly, and video games and gaming. The interviews were conducted as a dialogue in which meaning was actively cocreated, and shared understanding was confirmed through the interviews. Member checking was not conducted [58]. The interviews were conducted in Finnish via Zoom (Zoom Video Communications Inc) between May and September 2022, and they had an average duration of 57 (SD 9) minutes. They were audio-recorded with the interviewee's explicit permission, and the first author (LL) transcribed the recordings verbatim for analysis.

For dataset 1, data sufficiency was evaluated using the concept of saturation, such that data collection was concluded when no substantively new information emerged. Saturation was evaluated through analytic memoing conducted during and after the interviews [59]. After 12 interviews, diminishing returns in new information were observed, and the first (LL) and second (VMK) authors agreed that 19 interviews were sufficient to address the original research questions. Data analysis then

commenced, focusing first on the digital tools used by the MHPs, which were reported elsewhere [20]. This analysis did not examine MHPs' views on video games and gaming, because this topic was found to be largely independent of how MHPs used digital tools in their clinical practice.

Subsequently, when preparing the analysis of MHPs' perceptions of video games and gaming, the second author (VMK), who had been involved in the studies gathering datasets 2 and 3, proposed the use of data triangulation. Incorporating datasets 2 and 3 allowed for the inclusion of material gathered by different researchers and using slightly different interview guides and recruitment networks. Consequently, dataset 1 was complemented with these additional datasets to enable data source and investigator triangulation [60], thereby enhancing data richness and credibility.

For dataset 2, the development of the interview guide was guided by the aim of investigating psychology students' and psychologists' views on video games, problematic gaming, and gaming disorder. To study these topics, the interview guide (Alho et al [34]) included 4 sections, that is, background questions to explore participants' clinical background and gaming experience; an adapted version of the Attitudes Toward Gambling Scale 8 [61], in which the word "gambling" was replaced by the word "gaming," to study attitudes toward gaming; follow-up questions on attitudes toward video games; and questions about the new diagnostic category of gaming disorder. The interviews had an average duration of 35 minutes, and they were conducted in Finnish via Zoom between February and March 2021. The interviews were audio-recorded with permission from the interviewee and transcribed verbatim.

For dataset 3, the development of the interview guide was guided by the aim of contextualizing problematic and esports play through interviews with MHPs who had worked with at least 1 individual experiencing gaming-related problems. VMK and Miia Siutila conducted interviews with MHPs using a semistructured 12-item interview guide (Siutila et al [50]), which examined their work context, clinical experiences, practices related to treating clients who lived with gaming-related problems, their personal gaming experiences, and their perceptions of gaming disorders. The interviews had an average duration of 48 (SD 8) minutes, and they were conducted remotely via Zoom between June and November 2021. The interviews were recorded with the interviewees' permission and transcribed verbatim.

Interview Data Analysis

The interview data analysis approach was inductive, interpretive, and pragmatic. This approach was selected due to the characteristics of the research domain [62]; there is little previous research on MHPs' attitudes toward video games and game-based digital interventions, and the existing findings are mixed [63]. Therefore, it was logical to proceed with a ground-up data analysis approach [64] while acknowledging that no research can be fully inductive [65]. The study paradigm was interpretive [48] and aimed to understand the meanings MHPs attributed to video games and gaming in their clinical contexts. The analysis was also influenced by the pragmatic aim of furthering game-based digital intervention development,

which is derived from the pragmatist stance in health care [66], which aims to alleviate suffering and improve well-being.

The data was analyzed using RTA [51,52]. The methodological choice was based on our overall evaluation of the research aims, the quality and nature of the data, and our familiarity with the method. We chose to use RTA over coding reliability, thematic analysis, or codebook thematic analysis [65] due to our inductive and interpretive approach. RTA emphasizes the researchers' active and reflexive role in the interpretation of data that occurs at the intersection of theoretical assumptions, analytic resources, and skills, and the data [65,67]. The researcher does not discover, but develops, constructs, and generates themes that are defined as "*patterns of shared meaning, united by a central concept or idea*" [65]. This should lead to a "*compelling, insightful, thoughtful, rich, complex, deep, and nuanced*" analysis, which acted as a quality criterion [68].

The first author (LL) conducted the analysis. It began with combining the datasets 1-3 and familiarization with them while making notes. Then, the combined data were coded using Atlas.Ti 23 software (ATLAS.ti GmbH). RTA emphasizes deep engagement with the data [65], and considerable time was given to generating the themes over a period of over half a year. During this time, the researcher engaged with the data, coded it, took a distance from it, and returned to review the themes several times. The second author (VMK) reviewed and helped revise the generated themes in 4 sense-making sessions with the first author (LL). Through this process, 3 themes were eventually considered to be meaningful and stable. They are illustrated with translated interview quotes and numbers (eg, #1) that connect them to a particular interviewee and dataset (Table 1).

We consider the combined sample and the data, diverse and adequate, to answer the research question. In RTA, it is not recommended to use the concept of saturation [69]. Given the exploratory nature of this study and the complexity of clinicians' perceptions of the topic of gaming, this study does not capture all possible views on the topic. For instance, this study did not gather data on MHPs' views on specific game platforms or genres, or on specific game-based interventions. However, we consider that the sample is sufficient to answer the research question of how MHPs view video games and gaming in clinical contexts. For reference, comparable qualitative research typically requires only 9-17 interviews [70,71], whereas our combined sample included 41 interviews. The inclusion of 3 complementary datasets enhanced the diversity and comprehensiveness of the qualitative material and enabled data source and investigator triangulation [60]. The manuscript was reviewed and revised in consultation with the Standards for Reporting Qualitative Research [72] (Multimedia Appendix 1) and Journal Article Reporting Standards for Qualitative Research [73].

The first author (LL) rooted the analysis in the data while being mindful that their own experiences in the domain inevitably contribute to their interpretations—in RTA, this awareness acts as a resource [65]. Considering positionality, the first author (LL) is a Finnish psychologist who has worked in Finnish mental health care, which allowed them to relate to the participant's

clinical working context. On the other hand, they are experienced in games; they play analog and digital games, have a degree in game design and production, and they have worked with entertainment games [74], gaming disorder [75], and game-based intervention for adult major depressive disorder [76]. This allowed them to connect with the participant's personal and professional experiences regarding video games. The second author (VMK) is an interdisciplinary researcher of play and games, with a specific focus on clinical and psychological approaches to gaming. A personal history of active gaming has also contributed to the second author's researcher position.

Post Hoc Analysis

After the interview data analysis, 2 qualitative post hoc analyses were designed to provide complementary perspectives to the primary interview data analysis. The post hoc data were gathered with dataset 1 using an online questionnaire (Multimedia Appendix 2). The respondent characteristics were 80% (64/80) identified as women, 81% (65/80) were 30-59 years old, 74% (59/80) were psychologists, 78% (62/80) worked full time, 40% (32/80) worked in specialized health care, 39% (31/80) played video games daily or weekly, 35% (28/80) did not play video games at all, and the respondents had an average of 16 (SD 12) years of mental health working experience (Multimedia Appendix 3).

The first theme generated via RTA revealed how MHPs' personal and professional gaming experiences differed, suggesting self-client attitude asymmetry. The gathered questionnaire data allowed for the methodological triangulation of this theme [60]. Therefore, the first post hoc analysis compared the questionnaire respondents' attitudes toward video games in their personal lives (question 22: "What does playing

digital games mean to you?") and in their professional context (question 25: "How does video game play exhibit itself in your client work?"). All participants answered both questions. The analysis was conducted by deductively coding MHPs' open-ended responses by their emotional valence from positive to neutral to negative [59]. Responses containing multiple positions were coded as including multiple categories. For example, the statement "[Gaming is] a way to spend time, a way to fill time, fun activity, numbing, something that you can use to avoid things and that you can get addicted to" was coded as both positive and negative. The results were reported as domain summaries [67]. Because these post hoc analyses were qualitative and nonstatistical and used to illustrate patterns of meaning, CIs for coding frequencies are not reported.

The 3 RTA themes did not provide insights into what kind of benefits game-based interventions could have and for whom. Therefore, we decided to complement the interview data analysis with the second post hoc analysis, which analyzed MHPs' responses to the questions regarding the expected benefits of game-based interventions (question 27: "What benefits in particular would you expect from game-based digital therapies?") and their target audience (question 30: "For whom do you think game-based digital therapies would be most useful and why?"). In total, 79 (99%) participants answered question 27, and all participants answered question 30. The open-ended responses were categorized with descriptive coding [59] and reported as domain summaries.

Results

Overview

The RTA of interview data generated 3 themes (Table 2).

Table 2. Reflexive thematic analysis of interview data generated 3 themes, which are illustrated with quotes.

Theme	Description	Illustrative quotes
Personal recreation, clinically harmful	MHPs ^a views on video games were influenced by the context of their gaming experiences. MHPs' self-selected personal gaming experiences were typically positive and associated with recreation and pastime. In contrast, MHPs' professional experiences with gaming occurred in a problem-focused clinical environment, which associated their clients' gaming with mental health challenges, harm, and potential addiction. However, the client cases could also expand MHPs' views on gaming.	<ul style="list-style-type: none"> • "Games are close to my heart and I have played since I was a kid. My earliest memories are associated with gaming: I played Super Mario Bros 2 on the Nintendo Entertainment System when I was 4-5 years old." [#14] • "I have two children and they play video games. I see what they are doing, and I have learned about that world through them, and I know what it is about." [#17] • "In my work, I am mostly faced with the problems gaming causes." [#1] • "[In the clinical context] I have maybe seen the breadth of problematic gaming more than what I would have seen in a hobby. This has broadened my views on how different people benefit from and have problems with games." [#26]
Adverse technology and meaningful culture	MHPs framed video games as both adverse technology and meaningful culture. Here, framing refers to a process of meaning-making through comparison and association. When MHPs framed video games as technology, gaming was associated with smartphones and digital media, which were perceived to have the potential for excessive use and harms. When MHPs framed video games as culture, gaming was compared to storytelling, hobbies, and arts, which were characterized as meaningful activities and recreation. Many MHPs acknowledged the simultaneous presence of both frames, which reflected the complexity of gaming as a societal phenomenon.	<ul style="list-style-type: none"> • "We encourage people to read, watch movies, listen to music, and engage in other cultural experiences. I think games could very well be in the same category." [#27] • "I am occasionally frustrated that gaming is presented alongside snus, cannabis, and alcohol in a preventative substance abuse event. Gaming does not belong to that category, it is a hobby." [#40] • "Schools have pretty negative views on gaming. The teachers are concerned that children play a lot and spend a lot of time with digital devices." [#32] • "Games are not a black-and-white thing, but a new phenomenon to cope with. The same goes for smartphones, some people spend way too much time on them." [#16]
Holistic exploration of clients' gaming	MHPs explored their clients' lives holistically: their responsibilities, social interactions, rest, exercise, and recreation. Particular attention was paid to whether any of these activities, including but not limited to gaming, had a detrimental effect on other life domains. Thus, the positive and adverse effects of any activity could only be identified by examining life as a whole. The balance between activities could change over time, and MHPs identified two courses where gaming could evolve from a recreational activity to a problem: from connection to loneliness and from comfort to avoidance.	<ul style="list-style-type: none"> • "I try to think about life as a whole, about all the things that are important to the person. If you miss out on important because of playing, you play too much. This is true for other things as well: sleeping, working, seeing friends, and spending time with the spouse; so I don't think it merely concerns playing." [#33] • "I think more important than the hours played is whether gaming has a negative influence on other fields of life." [#24] • "I don't believe that [excessive] gaming is the root cause, but a symptom." [#14] • "Most people do not have problems [with gaming]. Those children who have, have other mental health problems or risk factors to them, such as anxiety disorders, social anxiety, maybe difficulties in forming friendships, domestic challenges." [#29]

^aMHP: mental health professional.

Personal Recreation, Clinically Harmful

MHPs' views on video games depended on whether they considered their own experiences (as players or spectators) or those of their clients. Gaming represented a long-term interest and hobby for several MHPs. One interviewee explained:

I have played quite a lot, for as long as I can remember. My first memory, in fact, is of a video game. [#23]

Such a close relationship with gaming was understandably associated with perceiving the activity positively, even enthusiastically. Other MHPs played digital games more casually or had a more distant, spectator relationship with gaming. Their second-hand gaming experiences involved observing the gaming of their friends, siblings, spouse, and children, and such connections fostered familiarity with gaming and the perception of it as a meaningful recreational activity.

Few MHPs associated video games with themes of threat and control, and they discussed the necessity of and challenges with limiting their children's playtime to ensure moderation and daily rhythm.

The positive connotations associated with gaming in personal contexts were less present in problem-focused professional clinical contexts. Working with people who had mental health challenges often focused attention on problems, implicitly associating gaming with risks, harms, procrastination, interpersonal problems, and addictive behaviors. One participant reflected:

A psychotherapist always faces things that have become a problem. Then, the first association is that someone has a problem with games, that they are addicted or something. [#17]

Such a perspective was reinforced by the clients having little need to explore gaming when it was recreational, positive, and meaningful. Therefore, MHPs' professional perspective on gaming could be limited, and some felt their views were not representative of the general population. However, this did not necessarily mean their views were negative. Many MHPs acknowledged that gaming could offer their clients a meaningful way to pass the time, relax, socialize, and manage emotional states, but they commonly considered gaming a conflicted or divided topic that was shadowed by its potential harms.

MHPs' personal gaming experiences were often self-affirmingly positive, whereas their professional context was associated with a greater possibility for attitudinal change and enrichment. Some MHPs, enthusiastic about video games, found that their professional experiences had made them more aware of potential problems associated with gaming. One participant described:

Before [running a video-game-related group intervention] I was more open-minded and liberal regarding problematic gaming. Now, I recognize and understand its divisiveness better: even if gaming is meaningful, it is potentially harmful. [#27]

Conversely, a clinician with initial reservations about video games described how they came to appreciate the value gaming offered their clients:

If I was not working here at the neuropsychiatric clinic, I would probably think differently. I would think that games do more harm than good. [#9]

These attitudinal changes stemmed from client cases that broadened MHPs' perspectives beyond their usual social circumstances. Indeed, some MHPs reflected on case stories, which included cautionary examples, such as a compulsively playing mother and a boy walking into a ditch while playing on their phone, but also inspiring anecdotes, such as a physically disabled youth who found rich social contact through games and highly talented players who earned their living from competitive gaming. These exceptional case stories expanded MHPs' views on gaming through idiosyncratic examples.

Adverse Technology and Meaningful Culture

MHPs framed games as both adverse technology and meaningful culture. By framing, we refer to a process in which the interviewee associated and compared gaming with other activities, thus establishing categorical similarity and connection. This allowed MHPs to make sense of gaming by positioning it within broader societal developments and phenomena, which in turn also rationalized their attitudes and cyclically maintained them. Some MHPs emphasized one frame over another, while others acknowledged that gaming exhibited characteristics of both frames. This introduced tension and made it difficult to make sense of gaming as it was not merely harmful or beneficial, but rather had the simultaneous potential for both. Thus, gaming was a complex, divisive, and even polarizing topic that could not be described in black-and-white terms.

When MHPs framed video games as technology, they grouped and compared them with smartphones, television, streaming services, and social media. Video games were viewed as an inseparable part of societal technologization and digitalization,

as well as the growing use of digital media and smartphones, in a negative sense. One interviewee reflected:

I think that the problems with video game playing are part of a broader discussion on the use of technology. – – The same types of addictive features are also present in other digital software, not only in games. Therefore, examining video games independently does not capture the whole they are a part of. [#23]

Some MHPs considered this change to be linked to the technology and gaming industries, which aimed to design highly engaging products to increase their revenues while presenting themselves favorably as providers of employment and tax revenue. The frame of technology was associated with concerns about excessive digital device use, which could disrupt sleep, detach patients from their bodily experiences, and encourage social withdrawal. One participant explained:

I am afraid of what kind of a country or world this will become if there is too much digital interaction. It feels as if there is a technological religiosity, and digital playing is only a part of it. [#21]

When gaming was framed as technology, it was associated with the disproportionate, pervasive, and adverse influence of digital devices on society.

In contrast, some MHPs framed video games as culture. They described gaming as a hobby, pastime, and recreational activity, which associated it with widely accepted activities such as storytelling, reading, doing crossword puzzles, listening to music, going to the movies, and doing handcrafts. This normalized and depathologized gaming and linked it with positive, life-enriching, and meaningful experiences, relaxation, and ways to pass the time. Video games were seen as the latest manifestation of the continuous cultural evolution of recreational activities. One interviewee described:

My generation played with dolls, cars, and in sandboxes—these activities have shifted [to video games]. [#19]

Concerns regarding the ill influence of gaming were perceived as an instance of media panic, where the latest form of media was unnecessarily feared and blamed. Within the frame of culture, the distinct feature of video games was their capability to invite participation in an immersive and sensory-rich story. One clinician characterized this as follows:

Video games are a form of art that enables processing different themes, telling stories, and stepping into another person's shoes. They do it in a way that is different or unique compared to literature, movies, theater, or other forms of art. [#31]

These features set video games apart from other forms of culture to which they were compared.

Holistic Exploration of Clients' Gaming

MHPs explored their clients' lives holistically. They found that the universal pillars of a good life consisted of sufficient sleep, nutrition, exercise, social relationships, pursuing livelihood through studies or work, and recreation. MHPs consistently positioned video games within recreation where gaming offered

challenges, feelings of competence, opportunities for social interaction, and a way to unwind and relax. However, they found that evaluating the overall impact of gaming required considering it against the other life domains, because well-being was a matter of balance. One participant described this as “a golden middle way” (#32). Another elaborated:

If you have a lot of social relationships, do sports, play an instrument, and then sometimes play video games, even a lot, things are in balance. [The influence of gaming] depends on the whole. [#21]

Vice versa, MHPs agreed that disturbances to the balance between activities lead to problems. This meant that one activity, such as gaming, could cause “harm to other fields of life” (#9). This included negative influence on responsibilities, such as work or studies, or the fulfillment of basic human needs, such as nutrition, rest, daily rhythm, exercise, and social contact. Therefore, the activity itself was not perceived as a problem, but its influence on other necessary and intrinsically valuable domains of life. Moreover, the participants did not find that this principle was specific to video games, because anything in excess could be detrimental.

Therapeutic interaction included exploring the position and role video games played in the client’s life. The interviewees found that challenges in regulating gaming were rarely the primary reason for help-seeking, but other mental health issues, such as depression and anxiety, were. These symptoms, in turn, had predisposing factors, such as developmental and neuropsychological problems, learning disabilities, and external stressors (such as challenging domestic environments, bullying, bereavement, breakup, and loneliness). From this perspective, an existing behavioral problem, such as excessive gaming, could be a coping attempt to an emotional, social, or cognitive issue. Over time, the role of playing may have changed from an initial solution to a rigid, unsuccessful coping mechanism. One clinician characterized:

The motivation to playing may not be the same as it once was. It may be more compulsive and routinized, and you may not enjoy it as much. [#41]

Thus, using games as self-help could turn from the alleviation of symptoms to the maintenance of the client’s issues.

MHPs described 2 problematic pathways in which initial recreational gaming could evolve into a problem—from connection to loneliness and from comfort to avoidance. Regarding the former, MHPs found that players may seek social connection from digital games, for instance, due to shyness or social anxiety. However, finding such online connections may lead to further withdrawal from social contacts outside gaming

and exacerbate social anxiety, because the patients do not get exposure to real-life social situations. One clinician described:

Multiplayer games are an opportunity to find friends and community from the other side of the world – – but it can also offer an easy way out so that you don’t have to think about your feelings of loneliness and find other solutions as you can withdraw to the game world. [#31]

Considering the latter pathway, the clients may also struggle with stress and depression, and fail to experience meaning in their lives, which can lead to seeking comfort, feelings of competence, and motivation from digital games. However, because gaming requires less effort than studying or working, this may lead to detrimental avoidance behavior, which prevents solving the underlying issues. Using gaming to manage stress, anxiety, and depression may end up aggravating the problems through self-defeating behavior. One clinician portrayed:

At least initially, gaming may have been therapeutic. You may have anxiety, depression, or a life crisis, and then you seek comfort from playing. But over time the amount of gaming may increase and become a problem of its own. [#37]

Differentiating between gaming as an avoidance behavior and a coping strategy could be challenging, but it was considered important in understanding the influence of gaming.

Post Hoc Analysis of Questionnaire Data

Do MHPs’ Views on Their Own and Their Clients’ Gaming Differ?

This first post hoc analysis triangulated the first theme generated in the RTA, investigating whether MHPs’ views on their own and their clients’ gaming differed. The post hoc analysis found that MHPs associated their own video game experiences with more positive meanings compared with their clients’ gaming (Table 3). When thinking about their own life, 73% (58/80) of MHPs associated gaming with positive meanings, such as recreation, relaxation, benefits, and a way to spend time, whereas considering their clients, only 38% (30/80) made such positive associations. Examining their own life, 16% (13/80) viewed video games in a negative light as uninteresting, uninspiring, addictive, and a waste of time, whereas when thinking about their clients, 64% (51/80) were concerned about excessive gaming, gaming-related interpersonal problems, avoidance behavior, and problems with daily rhythm. Overall, this analysis provided further evidence of the self-client attitude asymmetry concerning video games.

Table 3. The post hoc analysis of open-ended questionnaire responses (n=80) with quotes annotated with the respondent number (eg, #q1). Responses could be coded into multiple categories, and therefore percentages do not sum to 100.

Question and category	Quote	n (%)
“What does playing digital games mean to you?” (question 22)		
Positive or beneficial	<ul style="list-style-type: none"> “A way to spend time with my spouse and friends. Pleasant recreation.” [#q54] “An opportunity to momentarily disrupt my thoughts by intensively focusing on something else.” [#q46] 	58 (73)
Neutral	<ul style="list-style-type: none"> “I don’t have a relationship with playing.” [#q29] “It is not currently a part of my life at all. I last played in the 90s.” [#q17] 	19 (24)
Negative or harmful	<ul style="list-style-type: none"> “They are something I don’t want to give any time from my life.” [#q6] “Senseless waste of time and a residue of childhood.” [#q64] 	13 (16)
“How does video game play exhibit itself in your client work?” (question 25)		
Positive or beneficial	<ul style="list-style-type: none"> “Some clients consider playing a meaningful hobby.” [#q9] “It offers many people a way to relax like watching TV did earlier.” [#q39] 	30 (38)
Neutral	<ul style="list-style-type: none"> “It only shows itself in conversations with clients.” [#q44] “Almost all clients play.” [#q21] 	25 (31)
Negative or harmful	<ul style="list-style-type: none"> “People increasingly use these games to avoid real social contacts.” [#q6] “It maintains insomnia.” [#q20] 	51 (64)

What Benefits Could Game-Based Interventions Have and for Whom?

This second post hoc analysis complemented the RTA of interview data. It used the open-ended questionnaire responses to investigate what kind of benefits game-based interventions could have and for whom. Many MHPs (n=37, 46%) found that game-based interventions could be approachable and lower the threshold to participate in treatment. Some MHPs (n=20, 25%) expected that game-based interventions could create therapeutic impact by motivating the client to take care of themselves, be more active, build new skills, and self-reflect. Some MHPs (n=18, 23%) found that game-based interventions could augment and complement traditional therapies by providing structured content between sessions. Few MHPs (n=4, 5%) considered that game-based approaches were not feasible because they do not incorporate the crucial salutatory component of therapeutic, interpersonal interaction.

Many MHPs (n=34, 44%) considered that game-based interventions were particularly suitable for children and youth. This view was occasionally rationalized with younger generations being digitally native, and many MHPs (n=33, 41%) considered that competence with computers or playing video games could facilitate the uptake of game-based interventions. While the mentions of gender were infrequent, MHPs exclusively mentioned boys and men (n=6, 8%) when making such remarks. The participants also considered that game-based interventions could be feasible for clients who were hard to reach, withdraw, or have challenges with interpersonal interaction (n=25, 31%), or have neuropsychological challenges (n=5, 6%). In contrast to these cases that exhibited rather severe psychiatric problems, some participants suggested that the client’s symptoms should be mild and that the client should be proactive and capable of using such interventions (n=9, 11%). Refer to [Multimedia Appendix 4](#) for quotes.

Discussion

Principal Findings

This qualitative study investigated Finnish MHPs’ views on gaming to inform the implementation of game-based interventions in health care. Although previous research highlights the importance of clinicians’ views on the adoption of new health care technologies [23-27], detailed accounts of how clinicians make sense of gaming within clinical contexts have remained limited. To address this gap, we analyzed 3 distinct qualitative interview datasets (n=41) using RTA, which converged into three themes: (1) MHPs exhibited a self-client attitude asymmetry where they associated their own gaming positively with recreation while linking their clients’ gaming to problems; (2) MHPs’ views expressed attitudinal ambivalence: they made sense of gaming through conflicting frames, viewing it simultaneously as harmful technology and meaningful culture; and (3) MHPs evaluated clients’ gaming holistically within the broader context of clients’ lives (Table 2). Furthermore, 2 complementary post hoc analyses of open-ended questionnaire responses (n=80) further contextualized these themes. The first corroborated the self-client attitude asymmetry identified in the interviews (Table 3). The second expanded the core findings by showing that MHPs expected game-based interventions to be approachable, motivating, and complementary to other treatments, and particularly suitable for children and youth, digitally native generations, and clients who were withdrawn or difficult to reach. Overall, these findings provide nuanced insights into the perceptions shaping MHPs’ decisions to adopt and recommend game-based interventions in clinical practice.

Self-Client Attitude Asymmetry

In the first theme, we found that MHPs' personal gaming experiences were positive and associated primarily with recreation. This aligns with previous work showing that limited gaming experience is associated with negative assumptions [47] and that first-hand exposure to gameplay can shift these views. Ferguson et al [77] demonstrated that older adults developed more positive attitudes toward video games after playing one themselves, concluding that "*negative attitudes toward video games exists mainly in the abstract and do not survive direct exposure to individual games.*" Because game-based mental health interventions are still rarely used in Finland, MHPs had no first-hand experience of such tools. Offering clinicians opportunities to try game-based interventions as they are introduced in health care may therefore help challenge preconceptions, improve acceptability, and support intervention uptake. These efforts may be further supported by lead users or internal champions [78] who have personal and professional experience of gaming and game-based interventions and can facilitate informed discussions within their organizations.

Previous research has shown that MHPs hold mixed views about gaming and game-based interventions [32,33], but the reasons underlying these perceptions have remained unclear. A key contribution of this study is the identification of a self-client attitude asymmetry; although MHPs viewed their own gaming positively, they tended to perceive their clients' gaming as potentially problematic. This discrepancy may reflect the problem-oriented nature of psychiatric practice [79], where clinical encounters focus on difficulties and challenges, leaving the potential benefits of gaming underrecognized, even when they contribute to clients' well-being. From an implementation perspective, these findings highlight the value of providing clinicians with concrete case examples that illustrate the therapeutic potential of game-based interventions.

Negotiating Attitudinal Ambivalence

The second theme illustrates how MHPs use 2 nonexclusive and at times conflicting frames to make sense of gaming—adverse technology and meaningful culture. Entman [80] defines framing as selecting: "*some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation.*" Thus, these frames allowed MHPs to interpret the complex phenomenon of video games within a broader societal context [81]. Previous research on serious games has identified 4 complementary frames [82]. This study expands the framing approach by focusing specifically on clinical contexts, revealing the 2 frames most salient to MHPs. Furthermore, these 2 frames exhibit attitudinal ambivalence [83], in which both positive and negative evaluations coexist, and suggest that MHPs' perspectives on gaming cannot be accurately represented on a simple positive-negative continuum or measured with a single-dimensional scale.

The conflicting frames may surface in the implementation of game-based interventions, particularly when MHPs draw from broader societal discourses about the risks or cultural significance of video games and apply these discourses to

game-based interventions. It may therefore be beneficial to acknowledge these frames and the complexity surrounding the issue of gaming, and to exhibit that the frames applied to entertainment gaming are not necessarily sufficient when considering game-based interventions; entertainment games are used for a range of recreational reasons [84,85], whereas game-based interventions are designed to achieve health benefits [1]. Clients' expectations for game-based interventions are driven by a health-related need that the interventions aim to address through a therapeutic mechanism of action, which are not present in entertainment games. Thus, although game-based interventions may resemble commercial entertainment video games, they function as therapeutic tools or medical devices intended to meet specific health needs. Highlighting this distinction in clinician education and communications can help MHPs focus on the unique features of the technology and its intended purpose.

Contextualizing Gaming

The third theme illustrates how MHPs adopt a holistic view when assessing clients' gaming. They examined the role gaming played in clients' lives and sought to differentiate problematic gaming from gaming that functions as a personal resource [86]. MHPs did not single out video games as a unique threat to well-being [63] but viewed gaming as one influence among many. When gaming presented problems, clinicians often explored how the meaning and motivations behind gaming had changed. Here, we identified 2 pathways through which recreational gaming could evolve into problematic gaming—from connection to loneliness and from comfort to avoidance. These pathways highlight that the functions of gaming are not static, but dynamic, shifting by the players' life context.

The implementation of game-based interventions could benefit from adopting a similar holistic and temporal perspective. The educational and communication materials could emphasize the potential client-experienced benefits and position them within the broader context of their lives. To illustrate the role of game-based interventions in clients' health care and recovery processes, implementation efforts could apply a patient journey approach [87]. Case examples may be particularly effective in demonstrating how game-based interventions match diverse life trajectories and contribute to the lived experience of recovery [88], thereby adding detail to health-related claims.

Broadening Views on Users

The second post hoc analysis revealed that MHPs' attitudes reflected common stereotypes about video game players as male, young, and socially withdrawn—perceptions that have been upheld by media representations [89]. References to player gender were exclusively about boys and men (Multimedia Appendix 4), consistent with the marginalization of female video game players. While men may more commonly identify as gamers [90], approximately half of video game players are female [41], and 59% of Finnish women play video games at least monthly [40]. Given that depressive and anxiety disorders are more common in women [91], it is likely that potential users of game-based interventions for mood disorders may be disproportionately female. This is consistent with our previous

evaluation of a game-based intervention for adult depression, in which 57% (255/445) of participants identified as female, 34% (152/445) as male, and 7% (31/445) as nonbinary [76].

In this study, we also found common age-related stereotypes; many MHPs considered gaming to be an activity for children and youth. However, while the frequency of playing video games declines with age [40], an average player is 33 years old [41], and nearly half of Finns in their 30s play video games weekly [40]. Earlier, we found that a game-based intervention for adult depression attracted two-thirds of participants in their 20s and 30s, with a fourth aged 40-59 years [76], suggesting that there is an adult audience for these interventions.

Finally, some MHPs in this study associated video game playing with social anxiety, withdrawal, and awkwardness. However, previous work found no significant differences in social networks or exercise between online players, offline players, and nonplayers [89]. Overall, there appears to be a clear contrast between MHPs' expectations and the actual characteristics of many game-based intervention users.

This finding has significant implications for the implementation of game-based interventions; their potential target audience may be broader than clinicians assume, and inaccurate stereotypes may inadvertently limit who MHPs recommend such interventions to. To expand understanding of actual intervention users, researchers may need to identify who engages with and benefits from game-based interventions. Target audience profiling [76] and user-journey analyses [92] can help broaden clinicians' perceptions of the intervention user population and highlight cases in which interventions reach underrecognized patient groups, such as women older than 40 years.

Limitations

This study was conducted in Finland, a digitally advanced country [93] where digital solutions are widely used, including in health care [94]. Finland also has an active entertainment game industry, employing 4100 people in 2022 [95], which may contribute to higher digital and gaming familiarity. Regarding representativeness, interview participation may reflect self-selection bias, as individuals with an interest in or favorable attitudes toward the topic may be more likely to participate [96]. The interviews did not capture MHPs' perceptions of the regulatory status of digital interventions, which is a critical factor for their broader implementation. Regulatory frameworks differ across countries, are constantly evolving [97,98], and debate remains regarding which MHPs have the authority to prescribe such interventions [99], which warrant further investigation. Future research should also examine whether the

identified frames and self-client attitude asymmetry are expressed in other cultural contexts. There are also opportunities for studying how MHPs' perceptions may differ across certain gaming platforms and genres. Additionally, subsequent studies could broaden the analytical focus from individual MHPs to organizational factors, exploring how clinicians are introduced to and supported in using game-based interventions.

Conclusions

Existing qualitative research on MHPs' perceptions of video games is scarce and has not examined clinicians' broader sense-making of gaming, limiting its usefulness for informing the implementation of game-based interventions. To fill this gap, this qualitative study explored the views of Finnish MHPs and suggests ways to improve clinician education by building on MHPs' existing perceptions while addressing potential misconceptions. Because many MHPs have limited experience with game-based interventions, they tended to interpret them either through their own, often positive, recreational gaming experiences or through clients' gaming-related problems, revealing a previously underexplored self-client attitude asymmetry. Their views also exhibited attitudinal ambivalence, such that they made sense of gaming through conflicting frames as both potentially harmful technology and meaningful culture. To alleviate these tensions, clinician education could more clearly differentiate game-based interventions from recreational entertainment games and highlight their clinical purpose, mechanisms of action, and expected benefits. This communication could be further enhanced by providing clinicians with first-hand exposure to these interventions and by illustrating their therapeutic potential with case examples.

MHPs also tended to evaluate the impact that video games had on their clients' lives holistically and temporally, and they identified two pathways through which gaming could evolve into a problem: (1) from connection to loneliness and (2) from comfort to avoidance. Clinician education could use a similar approach and situate game-based interventions within clients' broader health care and recovery processes. Finally, MHPs often assumed that game-based interventions are best suited for stereotypically male, young, and socially withdrawn individuals. Clinician communication could help challenge these preconceptions and demonstrate how game-based interventions are appealing and relevant to a wider range of users than MHPs may initially assume. Together, these findings address an important gap in the literature and offer practical guidance for improving clinician education, communication strategies, and the adoption of game-based interventions in mental health care.

Acknowledgments

The authors would like to thank the researchers who gathered dataset 2 (Joel Alho and Katja Mankinen), the researchers who gathered dataset 3 (Miia Siuttila, Antti Päivinen, Soli Perttu, Tanja Laitinen, and Venla Korhonen), Otso Saariluoma for their comments on the article draft, and all the study participants for sharing their views. Artificial intelligence was not used in the preparation of this manuscript.

Data Availability

The dataset 1 used and/or analyzed during this study is available from the corresponding author on reasonable request.

The dataset 2 supporting the conclusions of this article is available in the Finnish Social Science Data Archive repository (FSD3671).

The dataset 3 supporting the conclusions of this article is available in the Finnish Social Science Data Archive repository (FSD3685).

Funding

This project was funded by Jane and Aatos Erkkö Foundation and Technology Industries of Finland Centennial funding for “The Future Makers Program” and Business Finland Research2Business funding (42173/31/2020) to JMP. VMK was supported by the European Research Council under the European Union’s Horizon Europe research and innovation program grant (101042052).

Authors' Contributions

The contributions are listed using the CRediT statement [100].

LL: Conceptualization, methodology, investigation, writing – original draft, writing – review and editing, visualization, project administration.

VMK: Writing – review and editing, supervision.

JMP: Funding acquisition, supervision, writing – review and editing.

Conflicts of Interest

LL and JMP are cofounders in Soihtu DTx Ltd, which develops game-based digital mental health interventions.

Multimedia Appendix 1

Standards for Reporting Qualitative Research.

[[DOCX File , 19 KB - games_v14i1e69236_app1.docx](#)]

Multimedia Appendix 2

Questionnaire content.

[[DOCX File , 38 KB - games_v14i1e69236_app2.docx](#)]

Multimedia Appendix 3

Questionnaire respondents.

[[DOCX File , 18 KB - games_v14i1e69236_app3.docx](#)]

Multimedia Appendix 4

The second post hoc analysis summary - Game-based intervention benefits and target audience.

[[DOCX File , 20 KB - games_v14i1e69236_app4.docx](#)]

References

1. Lukka L, Palva JM. The development of game-based digital mental health interventions: Bridging the paradigms of health care and entertainment. *JMIR Serious Games* 2023;11:e42173 [[FREE Full text](#)] [doi: [10.2196/42173](https://doi.org/10.2196/42173)] [Medline: [37665624](https://pubmed.ncbi.nlm.nih.gov/37665624/)]
2. Warsinsky S, Schmidt-Kraepelin M, Rank S, Thiebes S, Sunyaev A. Conceptual ambiguity surrounding gamification and serious games in health care: Literature review and development of game-based intervention reporting guidelines (GAMING). *J Med Internet Res* 2021;23(9):e30390 [[FREE Full text](#)] [doi: [10.2196/30390](https://doi.org/10.2196/30390)] [Medline: [34505840](https://pubmed.ncbi.nlm.nih.gov/34505840/)]
3. Fleming TM, Bavin L, Stasiak K, Hermansson-Webb E, Merry SN, Cheek C, et al. Serious games and gamification for mental health: Current status and promising directions. *Front Psychiatry* 2016;7:215 [[FREE Full text](#)] [doi: [10.3389/fpsy.2016.00215](https://doi.org/10.3389/fpsy.2016.00215)] [Medline: [28119636](https://pubmed.ncbi.nlm.nih.gov/28119636/)]
4. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: Evidence, theory and improvements. *Evid Based Ment Health* 2018;21(3):116-119 [[FREE Full text](#)] [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
5. Cuijpers P, Karyotaki E, Ciharova M, Miguel C, Noma H, Furukawa TA. The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: A meta-analysis. *Acta Psychiatr Scand* 2021;144(3):288-299 [[FREE Full text](#)] [doi: [10.1111/acps.13335](https://doi.org/10.1111/acps.13335)] [Medline: [34107050](https://pubmed.ncbi.nlm.nih.gov/34107050/)]
6. Cuijpers P, Miguel C, Harrer M, Ciharova M, Karyotaki E. The outcomes of mental health care for depression over time: A meta-regression analysis of response rates in usual care. *J Affect Disord* 2024;358:89-96 [[FREE Full text](#)] [doi: [10.1016/j.jad.2024.05.019](https://doi.org/10.1016/j.jad.2024.05.019)] [Medline: [38710332](https://pubmed.ncbi.nlm.nih.gov/38710332/)]
7. Abd-Alrazaq A, Al-Jafar E, Alajlani M, Toro C, Alhuwail D, Ahmed A, et al. The effectiveness of serious games for alleviating depression: Systematic review and meta-analysis. *JMIR Serious Games* 2022;10(1):e32331 [[FREE Full text](#)] [doi: [10.2196/32331](https://doi.org/10.2196/32331)] [Medline: [35029530](https://pubmed.ncbi.nlm.nih.gov/35029530/)]

8. Abd-Alrazaq A, Alajlani M, Alhuwail D, Schneider J, Akhu-Zaheya L, Ahmed A, et al. The effectiveness of serious games in alleviating anxiety: Systematic review and meta-analysis. *JMIR Serious Games* 2022;10(1):e29137 [FREE Full text] [doi: [10.2196/29137](https://doi.org/10.2196/29137)] [Medline: [35156932](https://pubmed.ncbi.nlm.nih.gov/35156932/)]
9. Dewhirst A, Laugharne R, Shankar R. Therapeutic use of serious games in mental health: Scoping review. *BJPsych Open* 2022;8(2):e37 [FREE Full text] [doi: [10.1192/bjo.2022.4](https://doi.org/10.1192/bjo.2022.4)] [Medline: [35105418](https://pubmed.ncbi.nlm.nih.gov/35105418/)]
10. Fitzgerald M, Ratcliffe G. Serious games, gamification, and serious mental illness: A scoping review. *Psychiatr Serv* 2020;71(2):170-183. [doi: [10.1176/appi.ps.201800567](https://doi.org/10.1176/appi.ps.201800567)] [Medline: [31640521](https://pubmed.ncbi.nlm.nih.gov/31640521/)]
11. 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;344:e2598 [FREE Full text] [doi: [10.1136/bmj.e2598](https://doi.org/10.1136/bmj.e2598)] [Medline: [22517917](https://pubmed.ncbi.nlm.nih.gov/22517917/)]
12. de Rutte J, Myruski S, Davis E, Findley A, Dennis-Tiwary TA. A randomized clinical trial investigating the clinical impact of a game-based digital therapeutic for social anxiety disorder. *J Anxiety Disord* 2025;111:103000. [doi: [10.1016/j.janxdis.2025.103000](https://doi.org/10.1016/j.janxdis.2025.103000)] [Medline: [40068434](https://pubmed.ncbi.nlm.nih.gov/40068434/)]
13. Selby E, Allabyrne C, Keenan JR. Delivering clinical evidence-based child-parent interventions for emotional development through a digital platform: A feasibility trial. *Clin Child Psychol Psychiatry* 2021;26(4):1271-1283 [FREE Full text] [doi: [10.1177/13591045211041580](https://doi.org/10.1177/13591045211041580)] [Medline: [34479455](https://pubmed.ncbi.nlm.nih.gov/34479455/)]
14. Kollins SH, DeLoss DJ, Cañadas E, Lutz J, Findling RL, Keefe RSE, et al. A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): A randomised controlled trial. *Lancet Digit Health* 2020;2(4):e168-e178 [FREE Full text] [doi: [10.1016/S2589-7500\(20\)30017-0](https://doi.org/10.1016/S2589-7500(20)30017-0)] [Medline: [33334505](https://pubmed.ncbi.nlm.nih.gov/33334505/)]
15. Gómez-Cambronero Á, Mann A, Mira A, Doherty G, Casteleyn S. Smartphone-based serious games for mental health: A scoping review. *Multimed Tools Appl* 2024;83(36):84047-84094. [doi: [10.1007/s11042-024-18971-w](https://doi.org/10.1007/s11042-024-18971-w)] [Medline: [39553422](https://pubmed.ncbi.nlm.nih.gov/39553422/)]
16. Lipschitz J, Hogan TP, Bauer MS, Mohr DC. Closing the research-to-practice gap in digital psychiatry: The need to integrate implementation science. *J Clin Psychiatry* 2019;80(3):18com12659 [FREE Full text] [doi: [10.4088/JCP.18com12659](https://doi.org/10.4088/JCP.18com12659)] [Medline: [31091029](https://pubmed.ncbi.nlm.nih.gov/31091029/)]
17. Treanor CJ, Kouvonen A, Lallukka T, Donnelly M. Acceptability of computerized cognitive behavioral therapy for adults: Umbrella review. *JMIR Ment Health* 2021;8(7):e23091 [FREE Full text] [doi: [10.2196/23091](https://doi.org/10.2196/23091)] [Medline: [34255714](https://pubmed.ncbi.nlm.nih.gov/34255714/)]
18. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework. *BMC Health Serv Res* 2017;17(1):88 [FREE Full text] [doi: [10.1186/s12913-017-2031-8](https://doi.org/10.1186/s12913-017-2031-8)] [Medline: [28126032](https://pubmed.ncbi.nlm.nih.gov/28126032/)]
19. Klaic M, Kapp S, Hudson P, Chapman W, Denehy L, Story D, et al. Implementability of healthcare interventions: An overview of reviews and development of a conceptual framework. *Implement Sci* 2022;17(1):10 [FREE Full text] [doi: [10.1186/s13012-021-01171-7](https://doi.org/10.1186/s13012-021-01171-7)] [Medline: [35086538](https://pubmed.ncbi.nlm.nih.gov/35086538/)]
20. Lukka L, Karhulahti V, Palva JM. Factors affecting digital tool use in client interaction according to mental health professionals: Interview study. *JMIR Hum Factors* 2023;10:e44681 [FREE Full text] [doi: [10.2196/44681](https://doi.org/10.2196/44681)] [Medline: [37428520](https://pubmed.ncbi.nlm.nih.gov/37428520/)]
21. Whitten PS, Mackert MS. Addressing telehealth's foremost barrier: Provider as initial gatekeeper. *Int J Technol Assess Health Care* 2005;21(4):517-521. [doi: [10.1017/S0266462305050725](https://doi.org/10.1017/S0266462305050725)] [Medline: [16262977](https://pubmed.ncbi.nlm.nih.gov/16262977/)]
22. Connolly SL, Hogan TP, Shimada SL, Miller CJ. Leveraging implementation science to understand factors influencing sustained use of mental health apps: A narrative review. *J Technol Behav Sci* 2021;6(2):184-196 [FREE Full text] [doi: [10.1007/s41347-020-00165-4](https://doi.org/10.1007/s41347-020-00165-4)] [Medline: [32923580](https://pubmed.ncbi.nlm.nih.gov/32923580/)]
23. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: Application to digital health-related behavior change interventions. *J Med Internet Res* 2015;17(1):e30 [FREE Full text] [doi: [10.2196/jmir.4055](https://doi.org/10.2196/jmir.4055)] [Medline: [25639757](https://pubmed.ncbi.nlm.nih.gov/25639757/)]
24. Lyon AR, Koerner K. User-centered design for psychosocial intervention development and implementation. *Clin Psychol (New York)* 2016;23(2):180-200 [FREE Full text] [doi: [10.1111/cpsp.12154](https://doi.org/10.1111/cpsp.12154)] [Medline: [29456295](https://pubmed.ncbi.nlm.nih.gov/29456295/)]
25. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. *J Med Internet Res* 2017;19(5):e153 [FREE Full text] [doi: [10.2196/jmir.7725](https://doi.org/10.2196/jmir.7725)] [Medline: [28490417](https://pubmed.ncbi.nlm.nih.gov/28490417/)]
26. van Gemert-Pijnen JEW, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. *J Med Internet Res* 2011;13(4):e111 [FREE Full text] [doi: [10.2196/jmir.1672](https://doi.org/10.2196/jmir.1672)] [Medline: [22155738](https://pubmed.ncbi.nlm.nih.gov/22155738/)]
27. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: Conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health* 2011;38(2):65-76 [FREE Full text] [doi: [10.1007/s10488-010-0319-7](https://doi.org/10.1007/s10488-010-0319-7)] [Medline: [20957426](https://pubmed.ncbi.nlm.nih.gov/20957426/)]
28. Verganti R. Design, meaning, and radical innovation: A metamodel and a research agency. *Journal of Product Innovation Management* 2008;25(5):436-456.
29. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research . . . and three things we can do about them. *Psychiatr Serv* 2017;68(5):427-429 [FREE Full text] [doi: [10.1176/appi.ps.201600541](https://doi.org/10.1176/appi.ps.201600541)] [Medline: [28412890](https://pubmed.ncbi.nlm.nih.gov/28412890/)]

30. Spies G, Huss J, Motswage O, Seedat S, Eichenberg C. Attitudes of patients and mental health professionals towards hypothetical use of serious games in psychotherapy. *Health Informatics J* 2024;30(2):14604582241259343 [FREE Full text] [doi: [10.1177/14604582241259343](https://doi.org/10.1177/14604582241259343)] [Medline: [38900576](https://pubmed.ncbi.nlm.nih.gov/38900576/)]
31. Eichenberg C, Grabmayer G, Green N. Acceptance of serious games in psychotherapy: An inquiry into the stance of therapists and patients. *Telemed J E Health* 2016;22(11):945-951. [doi: [10.1089/tmj.2016.0001](https://doi.org/10.1089/tmj.2016.0001)] [Medline: [27045452](https://pubmed.ncbi.nlm.nih.gov/27045452/)]
32. Tiburcio M, Martínez-Vélez NA, Bernal-Pérez RP, Huss J, Eichenberg C. Exploring Mexican psychotherapist's attitudes towards knowledge and use of serious games in clinical practice. *Front Digit Health* 2025;7:1584171 [FREE Full text] [doi: [10.3389/fdgh.2025.1584171](https://doi.org/10.3389/fdgh.2025.1584171)] [Medline: [40718500](https://pubmed.ncbi.nlm.nih.gov/40718500/)]
33. Ferguson CJ. Clinicians' attitudes toward video games vary as a function of age, gender and negative beliefs about youth: A sociology of media research approach. *Computers in Human Behavior* 2015;52:379-386. [doi: [10.1016/j.chb.2015.06.016](https://doi.org/10.1016/j.chb.2015.06.016)]
34. Alho J, Mankinen K. PPsychologists' and psychology students' attitudes about gaming, problematic gaming, and the new diagnosis for gaming disorder. Master's thesis. University of Jyväskylä. URL: <http://urn.fi/URN:NBN:fi:jyu-202203151822> [accessed 2023-09-19]
35. Segev A, Rovner M, Appel DI, Abrams AW, Rotem M, Bloch Y. Possible biases of researchers' attitudes toward video games: Publication trends analysis of the medical literature (1980-2013). *J Med Internet Res* 2016;18(7):e196 [FREE Full text] [doi: [10.2196/jmir.5935](https://doi.org/10.2196/jmir.5935)] [Medline: [27430187](https://pubmed.ncbi.nlm.nih.gov/27430187/)]
36. Ferguson CJ. Do angry birds make for angry children? A meta-analysis of video game influences on children's and adolescents' aggression, mental health, prosocial behavior, and academic performance. *Perspect Psychol Sci* 2015;10(5):646-666. [doi: [10.1177/1745691615592234](https://doi.org/10.1177/1745691615592234)] [Medline: [26386002](https://pubmed.ncbi.nlm.nih.gov/26386002/)]
37. Kühn S, Kugler DT, Schmalen K, Weichenberger M, Witt C, Gallinat J. Does playing violent video games cause aggression? A longitudinal intervention study. *Mol Psychiatry* 2019;24(8):1220-1234 [FREE Full text] [doi: [10.1038/s41380-018-0031-7](https://doi.org/10.1038/s41380-018-0031-7)] [Medline: [29535447](https://pubmed.ncbi.nlm.nih.gov/29535447/)]
38. Stevens MW, Dorstyn D, Delfabbro PH, King DL. Global prevalence of gaming disorder: A systematic review and meta-analysis. *Aust N Z J Psychiatry* 2021;55(6):553-568. [doi: [10.1177/0004867420962851](https://doi.org/10.1177/0004867420962851)] [Medline: [33028074](https://pubmed.ncbi.nlm.nih.gov/33028074/)]
39. Kim HS, Son G, Roh E, Ahn W, Kim J, Shin S, et al. Prevalence of gaming disorder: A meta-analysis. *Addict Behav* 2022;126:107183. [doi: [10.1016/j.addbeh.2021.107183](https://doi.org/10.1016/j.addbeh.2021.107183)] [Medline: [34864436](https://pubmed.ncbi.nlm.nih.gov/34864436/)]
40. Kinnunen J, Tuomela M, Mäyrä F. *Pelaajabarometri 2022: Kohti Uutta Normaalia*. 2022. URL: <https://urn.fi/URN:ISBN:978-952-03-2732-3> [accessed 2023-01-09]
41. Essential facts about the video game industry. Entertainment Software Association. 2022. URL: <https://www.theesa.com/resource/2022-essential-facts-about-the-video-game-industry/> [accessed 2022-12-29]
42. Johannes N, Vuorre M, Przybylski AK. Video game play is positively correlated with well-being. *R Soc Open Sci* 2021;8(2):202049 [FREE Full text] [doi: [10.1098/rsos.202049](https://doi.org/10.1098/rsos.202049)] [Medline: [33972879](https://pubmed.ncbi.nlm.nih.gov/33972879/)]
43. Bediou B, Adams DM, Mayer RE, Tipton E, Green CS, Bavelier D. Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychol Bull* 2018;144(1):77-110. [doi: [10.1037/bul0000130](https://doi.org/10.1037/bul0000130)] [Medline: [29172564](https://pubmed.ncbi.nlm.nih.gov/29172564/)]
44. Powers KL, Brooks PJ, Aldrich NJ, Palladino MA, Alfieri L. Effects of video-game play on information processing: A meta-analytic investigation. *Psychon Bull Rev* 2013;20(6):1055-1079. [doi: [10.3758/s13423-013-0418-z](https://doi.org/10.3758/s13423-013-0418-z)] [Medline: [23519430](https://pubmed.ncbi.nlm.nih.gov/23519430/)]
45. Sauce B, Liebherr M, Judd N, Klingberg T. The impact of digital media on children's intelligence while controlling for genetic differences in cognition and socioeconomic background. *Sci Rep* 2022;12(1):7720 [FREE Full text] [doi: [10.1038/s41598-022-11341-2](https://doi.org/10.1038/s41598-022-11341-2)] [Medline: [35545630](https://pubmed.ncbi.nlm.nih.gov/35545630/)]
46. Kovess-Masfety V, Keyes K, Hamilton A, Hanson G, Bitfoi A, Golitz D, et al. Is time spent playing video games associated with mental health, cognitive and social skills in young children? *Soc Psychiatry Psychiatr Epidemiol* 2016;51(3):349-357 [FREE Full text] [doi: [10.1007/s00127-016-1179-6](https://doi.org/10.1007/s00127-016-1179-6)] [Medline: [26846228](https://pubmed.ncbi.nlm.nih.gov/26846228/)]
47. Przybylski AK. Who believes electronic games cause real world aggression? *Cyberpsychol Behav Soc Netw* 2014;17(4):228-234. [doi: [10.1089/cyber.2013.0245](https://doi.org/10.1089/cyber.2013.0245)] [Medline: [24256132](https://pubmed.ncbi.nlm.nih.gov/24256132/)]
48. Fossey E, Harvey C, McDermott F, Davidson L. Understanding and evaluating qualitative research. *Aust N Z J Psychiatry* 2002;36(6):717-732. [doi: [10.1046/j.1440-1614.2002.01100.x](https://doi.org/10.1046/j.1440-1614.2002.01100.x)] [Medline: [12406114](https://pubmed.ncbi.nlm.nih.gov/12406114/)]
49. Alho J, Mankinen K. FSD3671 psykologien ja psykologian opiskelijoiden näkemyksiä pelaamisesta ja ongelmapelaamisesta 2021. Aila. Tietoarkisto. URL: <https://urn.fi/urn:nbn:fi:fsd:T-FSD3671> [accessed 2024-05-22]
50. Siuttila M, Karhulahti V, Päivinen A, Soli P, Laitinen T, Korhonen V. FSD3685 terveys ja videopelaaminen: Asiantuntijoiden haastattelut 2021. Tietoarkisto. 2023. URL: <https://urn.fi/urn:nbn:fi:fsd:T-FSD3685> [accessed 2023-11-01]
51. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2008;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
52. Braun V, Clarke V. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern - based qualitative analytic approaches. *Couns and Psychother Res* 2020;21(1):37-47. [doi: [10.1002/capr.12360](https://doi.org/10.1002/capr.12360)]
53. Lukka L, Karhulahti VM, Palva M. Mental health professionals' views and needs regarding game-based digital therapeutics. *Osf.io* 2022. [doi: [10.17605/OSF.IO/B8EG6](https://doi.org/10.17605/OSF.IO/B8EG6)]

54. The ethical principles of research with human participants and ethical review in the human sciences in Finland. Finnish National Board on Research Integrity TENK Guidelines 2019. 2019. URL: <https://tenk.fi/en/ethical-review> [accessed 2024-01-24]
55. Bishop L, Kuula-Luumi A. Revisiting qualitative data reuse. *Sage Open* 2017;7(1). [doi: [10.1177/2158244016685136](https://doi.org/10.1177/2158244016685136)]
56. Karhulahti VM, Siuttila M, Vahlo J, Koskimaa R. Phenomenological strands for gaming disorder and esports play: A qualitative registered report. *Collabra Psychol* 2022;8(1). [doi: [10.1525/collabra.38819](https://doi.org/10.1525/collabra.38819)]
57. McIntosh MJ, Morse JM. Situating and constructing diversity in semi-structured interviews. *Glob Qual Nurs Res* 2015;2:2333393615597674 [FREE Full text] [doi: [10.1177/2333393615597674](https://doi.org/10.1177/2333393615597674)] [Medline: [28462313](https://pubmed.ncbi.nlm.nih.gov/28462313/)]
58. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member checking: A tool to enhance trustworthiness or merely a nod to validation? *Qual Health Res* 2016;26(13):1802-1811. [doi: [10.1177/1049732316654870](https://doi.org/10.1177/1049732316654870)] [Medline: [27340178](https://pubmed.ncbi.nlm.nih.gov/27340178/)]
59. Saldaña J. *The Coding Manual for Qualitative Researchers*. Thousand Oak, California: SAGE Publications, Inc; 2017.
60. Thurmond VA. The point of triangulation. *J Nurs Scholarsh* 2001;33(3):253-258. [doi: [10.1111/j.1547-5069.2001.00253.x](https://doi.org/10.1111/j.1547-5069.2001.00253.x)] [Medline: [11552552](https://pubmed.ncbi.nlm.nih.gov/11552552/)]
61. Canale N, Vieno A, Pastore M, Ghisi M, Griffiths MD. Validation of the 8-item attitudes towards gambling scale (ATGS-8) in a British population survey. *Addict Behav* 2016;54:70-74 [FREE Full text] [doi: [10.1016/j.addbeh.2015.12.009](https://doi.org/10.1016/j.addbeh.2015.12.009)] [Medline: [26722993](https://pubmed.ncbi.nlm.nih.gov/26722993/)]
62. Barker C, Pistrang N, Elliott R. *Research methods in clinical psychology*. In: *An Introduction for Students and Practitioners*. New Jersey: Wiley; 2002.
63. Ferguson CJ, Colwell J. Lack of consensus among scholars on the issue of video game “addiction”. *Psychology of Popular Media* 2020;9(3):359-366. [doi: [10.1037/ppm0000243](https://doi.org/10.1037/ppm0000243)]
64. Elliott R, Timulak L. Descriptive and interpretive approaches to qualitative research. In: *A Handbook of Research Methods for Clinical and Health Psychology*. Oxford: Oxford University Press; 2005:147-159.
65. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology* 2020;18(3):328-352. [doi: [10.1080/14780887.2020.1769238](https://doi.org/10.1080/14780887.2020.1769238)]
66. Scott PJ, Briggs JS. A pragmatist argument for mixed methodology in medical informatics. *Journal of Mixed Methods Research* 2009;3(3):223-241. [doi: [10.1177/1558689809334209](https://doi.org/10.1177/1558689809334209)]
67. Braun V, Clarke V. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 2019;11(4):589-597. [doi: [10.1080/2159676x.2019.1628806](https://doi.org/10.1080/2159676x.2019.1628806)]
68. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. *Qualitative Psychology* 2022;9(1):3-26. [doi: [10.1037/qup0000196](https://doi.org/10.1037/qup0000196)]
69. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health* 2019;13(2):201-216. [doi: [10.1080/2159676x.2019.1704846](https://doi.org/10.1080/2159676x.2019.1704846)]
70. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Soc Sci Med* 2022;292:114523 [FREE Full text] [doi: [10.1016/j.socscimed.2021.114523](https://doi.org/10.1016/j.socscimed.2021.114523)] [Medline: [34785096](https://pubmed.ncbi.nlm.nih.gov/34785096/)]
71. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. *Field Methods* 2006;18(1):59-82. [doi: [10.1177/1525822X05279903](https://doi.org/10.1177/1525822X05279903)]
72. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: A synthesis of recommendations. *Acad Med* 2014;89(9):1245-1251. [doi: [10.1097/ACM.0000000000000388](https://doi.org/10.1097/ACM.0000000000000388)] [Medline: [24979285](https://pubmed.ncbi.nlm.nih.gov/24979285/)]
73. Levitt HM, Bamberg M, Creswell JW, Frost DM, Josselson R, Suárez-Orozco C. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA publications and communications board task force report. *Am Psychol* 2018;73(1):26-46 [FREE Full text] [doi: [10.1037/amp0000151](https://doi.org/10.1037/amp0000151)] [Medline: [29345485](https://pubmed.ncbi.nlm.nih.gov/29345485/)]
74. Lukka L. A study on the contemporary practices in mobile game development through nine cases and the introduction of the DPS model. *Conquering the mobile*. 2017. URL: <https://aaltodoc.aalto.fi/handle/123456789/28607> [accessed 2026-04-02]
75. Karhulahti VM, Behm S, Lukka L. Why do adults seek treatment for gaming (disorder)? A qualitative study. *Humanit Soc Sci Commun* 2023;10(1):299. [doi: [10.1057/s41599-023-01775-y](https://doi.org/10.1057/s41599-023-01775-y)]
76. Lukka L, Salonen A, Vesterinen M, Karhulahti V, Palva S, Palva JM. The qualities of patients interested in using a game-based digital mental health intervention for depression: A sequential mixed methods study. *BMC Digit Health* 2023;1(1):37. [doi: [10.1186/s44247-023-00037-w](https://doi.org/10.1186/s44247-023-00037-w)]
77. Ferguson CJ, Nielsen RKL, Maguire R. Do older adults hate video games until they play them? A proof-of-concept study. *Curr Psychol* 2016;36(4):919-926. [doi: [10.1007/s12144-016-9480-9](https://doi.org/10.1007/s12144-016-9480-9)]
78. Kirchner JE, Smith JL, Powell BJ, Waltz TJ, Proctor EK. Getting a clinical innovation into practice: An introduction to implementation strategies. *Psychiatry Res* 2020;283:112467 [FREE Full text] [doi: [10.1016/j.psychres.2019.06.042](https://doi.org/10.1016/j.psychres.2019.06.042)] [Medline: [31488332](https://pubmed.ncbi.nlm.nih.gov/31488332/)]
79. Borsboom D. A network theory of mental disorders. *World Psychiatry* 2017;16(1):5-13 [FREE Full text] [doi: [10.1002/wps.20375](https://doi.org/10.1002/wps.20375)] [Medline: [28127906](https://pubmed.ncbi.nlm.nih.gov/28127906/)]
80. Entman RM. Framing: Towards clarification of a fractured paradigm. *Journal of Communication* 1993;43(4):51-58. [doi: [10.1111/j.1460-2466.1993.tb01304.x](https://doi.org/10.1111/j.1460-2466.1993.tb01304.x)]

81. Scheff TJ. The structure of context: Deciphering frame analysis. *Sociological Theory* 2005;23(4):368-385. [doi: [10.1111/j.0735-2751.2005.00259.x](https://doi.org/10.1111/j.0735-2751.2005.00259.x)]
82. Mayer I, Warmelink H, Zhou Q. A frame - reflective discourse analysis of serious games. *Brit J Educational Tech* 2015;47(2):342-357. [doi: [10.1111/bjet.12245](https://doi.org/10.1111/bjet.12245)]
83. Schneider IK, Schwarz N. Mixed feelings: The case of ambivalence. *Current Opinion in Behavioral Sciences* 2017;15:39-45. [doi: [10.1016/j.cobeha.2017.05.012](https://doi.org/10.1016/j.cobeha.2017.05.012)]
84. Meriläinen M, Ruotsalainen M. The light, the dark, and everything else: Making sense of young people's digital gaming. *Front Psychol* 2023;14:1164992 [FREE Full text] [doi: [10.3389/fpsyg.2023.1164992](https://doi.org/10.3389/fpsyg.2023.1164992)] [Medline: [37388650](https://pubmed.ncbi.nlm.nih.gov/37388650/)]
85. Bäcklund C, Elbe P, Gavelin HM, Sörman DE, Ljungberg JK. Gaming motivations and gaming disorder symptoms: A systematic review and meta-analysis. *J Behav Addict* 2022;11(3):667-688 [FREE Full text] [doi: [10.1556/2006.2022.00053](https://doi.org/10.1556/2006.2022.00053)] [Medline: [36094861](https://pubmed.ncbi.nlm.nih.gov/36094861/)]
86. Giardina A, Starcevic V, King DL, Schimmenti A, Di Blasi M, Billieux J. Research directions in the study of gaming-related escapism: A commentary to melodia, canale, and griffiths (2020). *Int J Ment Health Addiction* 2021;21(2):1075-1081. [doi: [10.1007/s11469-021-00642-8](https://doi.org/10.1007/s11469-021-00642-8)]
87. Borycki EM, Kushniruk AW, Wagner E, Kletke R. Patient journey mapping: Integrating digital technologies into the journey. *Knowledge Management & E-Learning* 2020;12:521-535. [doi: [10.34105/j.kmel.2020.12.029](https://doi.org/10.34105/j.kmel.2020.12.029)]
88. Fusar-Poli P, Estradé A, Mathi K, Mabilia C, Yanayirah N, Floris V, et al. The lived experience of postpartum depression and psychosis in women: A bottom-up review co-written by experts by experience and academics. *World Psychiatry* 2025;24(1):32-45 [FREE Full text] [doi: [10.1002/wps.21264](https://doi.org/10.1002/wps.21264)] [Medline: [39810658](https://pubmed.ncbi.nlm.nih.gov/39810658/)]
89. Kowert R, Festl R, Quandt T. Unpopular, overweight, and socially inept: Reconsidering the stereotype of online gamers. *Cyberpsychol Behav Soc Netw* 2014;17(3):141-146. [doi: [10.1089/cyber.2013.0118](https://doi.org/10.1089/cyber.2013.0118)] [Medline: [24053382](https://pubmed.ncbi.nlm.nih.gov/24053382/)]
90. Paaßen B, Morgenroth T, Stratemeyer M. What is a true gamer? The male gamer stereotype and the marginalization of women in video game culture. *Sex Roles* 2016;76(7-8):421-435. [doi: [10.1007/s11199-016-0678-y](https://doi.org/10.1007/s11199-016-0678-y)]
91. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: A systematic analysis for the global burden of disease study 2019. *Lancet Psychiatry* 2022;9(2):137-150 [FREE Full text] [doi: [10.1016/S2215-0366\(21\)00395-3](https://doi.org/10.1016/S2215-0366(21)00395-3)] [Medline: [35026139](https://pubmed.ncbi.nlm.nih.gov/35026139/)]
92. Lukka L, Vesterinen M, Salonen A, Bergman V, Torkki P, Palva S, et al. User journey method: A case study for improving digital intervention use measurement. *BMC Health Serv Res* 2025;25(1):479 [FREE Full text] [doi: [10.1186/s12913-025-12641-9](https://doi.org/10.1186/s12913-025-12641-9)] [Medline: [40165237](https://pubmed.ncbi.nlm.nih.gov/40165237/)]
93. The digital economy and society index (DES) 2022 Finland. European Commission. 2022. URL: <https://digital-strategy.ec.europa.eu/en/policies/countries-digitisation-performance> [accessed 2026-04-02]
94. Kyytsönen M, Aalto A, Vehko T. Social and healthcare electronic services 2020-2021: experiences of the population. Finnish Institute of Health And Welfare. 2021. URL: <https://www.julkari.fi/server/api/core/bitstreams/2cf172ac-eaae-492c-abd5-acd0cafd028c/content> [accessed 2026-04-08]
95. The game industry of finland 2022. Neogames. 2022. URL: <https://neogames.fi/finnish-game-industry-report-2022/> [accessed 2023-12-08]
96. Lukka L, Karhulahti VM, Bergman VR, Palva JM. Measuring digital intervention user experience with a novel ecological momentary assessment (EMA) method, CORTO. *Internet Interv* 2024;35:100706 [FREE Full text] [doi: [10.1016/j.invent.2023.100706](https://doi.org/10.1016/j.invent.2023.100706)] [Medline: [38274123](https://pubmed.ncbi.nlm.nih.gov/38274123/)]
97. Essén A, Stern AD, Haase CB, Car J, Greaves F, Papparova D, et al. Health app policy: International comparison of nine countries' approaches. *NPJ Digit Med* 2022;5(1):31. [doi: [10.1038/s41746-022-00573-1](https://doi.org/10.1038/s41746-022-00573-1)] [Medline: [35304561](https://pubmed.ncbi.nlm.nih.gov/35304561/)]
98. Wang C, Lee C, Shin H. Digital therapeutics from bench to bedside. *NPJ Digit Med* 2023;6(1):38 [FREE Full text] [doi: [10.1038/s41746-023-00777-z](https://doi.org/10.1038/s41746-023-00777-z)] [Medline: [36899073](https://pubmed.ncbi.nlm.nih.gov/36899073/)]
99. Carl JR, Jones DJ, Lindhiem OJ, Doss BD, Weingardt KR, Timmons AC, et al. Regulating digital therapeutics for mental health: Opportunities, challenges, and the essential role of psychologists. *Br J Clin Psychol* 2022;61 Suppl 1:130-135. [doi: [10.1111/bjc.12286](https://doi.org/10.1111/bjc.12286)] [Medline: [33650131](https://pubmed.ncbi.nlm.nih.gov/33650131/)]
100. Allen L, O'Connell A, Kiermer V. How can we ensure visibility and diversity in research contributions? How the contributor role taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learned Publishing* 2019;32(1):71-74. [doi: [10.1002/leap.1210](https://doi.org/10.1002/leap.1210)]

Abbreviations

- FSD:** Finnish Social Science Data Archive
 - MHP:** mental health professional
 - RTA:** reflexive thematic analysis
-

Edited by S Brini; submitted 18.Aug.2025; peer-reviewed by T Baranowski, S Göbel, A Eisingerich; comments to author 01.Dec.2025; accepted 14.Feb.2026; published 20.Apr.2026.

Please cite as:

Lukka L, Karhulahti VM, Palva JM

Mental Health Professionals' Views on Gaming to Inform Game-Based Interventions: Qualitative Cross-Sectional Study

JMIR Serious Games 2026;14:e69236

URL: <https://games.jmir.org/2026/1/e69236>

doi: [10.2196/69236](https://doi.org/10.2196/69236)

PMID: [34779769](https://pubmed.ncbi.nlm.nih.gov/34779769/)

©Lauri Lukka, Veli-Matti Karhulahti, J Matias Palva. Originally published in JMIR Serious Games (<https://games.jmir.org>), 20.Apr.2026. 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.

Analysis of Machine Learning–Based Investigation Into Multivariate Factors of Team Performance in Serious Games: Cross-Sectional Retrospective Study

Gruyff Germain Abdul-Rahman¹, MSc, PhD; Freark de Lange¹, BSc; Andrej Zwitter², PhD; Noman Haleem¹, PhD

¹Department of Responsible Governance and Technology, Campus Fryslân, University of Groningen, Wirdumerdijk 34, Leeuwarden, The Netherlands

²Centre for Human and Data Sciences, University of Klagenfurt, Lakeside Park B07b.01.121, Klagenfurt, Carinthia, Austria

Corresponding Author:

Gruyff Germain Abdul-Rahman, MSc, PhD

Department of Responsible Governance and Technology, Campus Fryslân, University of Groningen, Wirdumerdijk 34, Leeuwarden, The Netherlands

Abstract

Background: Serious games (SGs) are increasingly used to study and enhance team performance in organizational and educational settings. While prior research has explored leadership and communication as isolated factors, the multivariate interactions between behavioral indicators remain poorly understood. A deeper understanding of these relationships can reveal which behavioral and demographic factors most strongly predict successful outcomes, offering insights relevant to both scientific research and practical training design.

Objective: This study aimed to develop machine learning (ML) models to predict team success in SGs. Specifically, it sought to identify the behavioral and demographic predictors that most strongly influence team performance outcomes.

Methods: This study used a cross-sectional retrospective design. Behavioral and demographic data were analyzed from 233 teams participating in escape room–based SGs delivered by JGM Serious eXperiences in The Netherlands. Teams of 2 - 8 players (mean age 25.8 y; 53 all-male, 55 all-female, and 125 mixed-gender) were scored by trained observers across collaboration, communication, and leadership constructs using Likert-scale indicators. Exploratory data analysis compared winning (n=141) and losing teams (n=92) using descriptive statistics, Pearson correlations, and significance testing (independent-samples *t* tests and Mann-Whitney *U* tests). Mean differences were interpreted with 95% CIs. A total of 4 ML models: logistic regression, random forest, multilayer perceptron, and support vector classifier, were trained using 5-fold cross-validation (F_1 -score). The best model was interpreted using SHAP (Shapley Additive Explanations).

Results: Winning teams scored higher on several behavioral constructs, but only 4: knowledge sharing, leadership, guidance, and extraversion, showed statistically significant differences between winners and losing teams. These effects were supported by 95% CIs, Shapiro-Wilk tests for normality, and Mann-Whitney *U* tests where assumptions were violated, indicating that only a subset of behavioral indicators meaningfully distinguishes successful teams. Among the ML models, logistic regression achieved the highest accuracy (88%), followed by multilayer perceptron (87%), random forest (87%), and support vector classifier (85%). SHAP analysis showed that gender composition and prior escape-room experience were the strongest demographic predictors of success, while “celebrating progress” (extern5) and “taking initiative when the team is stuck” (sturing5) were the most influential behavioral indicators.

Conclusions: This work demonstrates the usefulness of multivariate analysis in studying and understanding complex human behavior in SG environments as opposed to studying isolated behavioral indicators, often described in previous studies. The ML models developed using behavioral and demographic features of participating teams showed promising accuracies, and their interpretation led to unveiling a set of demographic and behavioral components as the most decisive factors leading to team success. This improved understanding of what makes a team win can be potentially translated into terms of improved productivity in business and organizational settings.

(*JMIR Serious Games* 2026;14:e83478) doi:[10.2196/83478](https://doi.org/10.2196/83478)

KEYWORDS

serious games; team performance; escape rooms; machine learning; explainable AI; explainable artificial intelligence

Introduction

Recent advances in data analytics have facilitated machine learning (ML) applications across several domains of science and business. A total of 1 area that has also revolutionized significantly from these developments is sports and gaming [1], where ML techniques can reveal hidden trends [2] that are important for the development of modern sports and provide a glimpse into the intricacy of player performance [1]. Predicting outcomes based on historical evidence has long been a central goal for researchers. This motivation has driven the development of various approaches, including sport-specific statistical simulations and sport-independent machine-learning techniques [2]. By using these tools, trends in sports data can be identified and used for personal, competitive, or economic advantage [2]. By using cutting-edge algorithms, sports analysts can uncover insights that were previously unattainable, ranging from physical attributes to strategic patterns, ultimately enabling a more comprehensive understanding that surpasses conventional analytical techniques [1]. The early work in this regard dates back more than half a century, when Samuel [3] used ML for game strategy optimization, marking the intersection of artificial intelligence (AI) and gaming. More recently, ML models have been applied to predict individual and team performance using historical data, which has enabled refined decision-making in competition [4,5]. For example, in football, variables such as shots on target, ball possession, and passing strategies were studied using ML techniques and subsequently linked to team success [6]. In cricket and basketball, ML models, such as Naive Bayes (NB) and support vector machines (SVMs), have demonstrated high predictive accuracy by accounting for complex interactions between team composition and match conditions [7,8]. ML techniques have also been effectively applied to understand team performance in sports, with methods such as random forest (RF) and Partial Least Squares Discriminant Analysis being applied to identify key performance metrics [9]. On an aggregate level, the application of ML has shown a promising potential in game analysis by capturing intricate performance patterns, which were otherwise difficult to analyze using traditional statistical techniques.

Recent work in the field of serious games (SGs) reflects a similar shift toward data-driven analysis and adaptive systems. Although most SGs do not currently incorporate ML, there is a growing trend toward integrating ML to assess participants and automatically adapt gameplay to individual needs [10]. These adaptive systems can personalize difficulty, improve engagement, and reduce learner frustration, but they also introduce practical challenges, such as selecting appropriate data sources, determining which game elements to modify, resolving cold-start problems, and validating whether adaptive changes genuinely support learning [10]. Evidence shows that ML techniques are increasingly used in SGs to monitor learner behavior, generate real-time feedback, and model player states such as motivation, skill level, or emotional reactions [11,12]. This has led to the emergence of AI-powered SGs that adjust game content based on individual progress, preferences, and training goals, thereby creating more personalized and effective learning experiences [13,14].

A recent systematic review highlights that advanced technologies, such as AI, immersive environments, and biosignal monitoring, can enhance meta-skills training by supporting adaptive learning, targeted practice, and improved monitoring of behavioral and cognitive states [15]. Studies demonstrate that AI-based SGs can develop key meta-skills, such as time management, self-regulation, and motivation, through difficulty adjustment and personalized support. For example, a study [16] used ML models to analyze players' biosignals in real time and modify game difficulty based on their emotional state, resulting in improved time-management skills. Despite these benefits, the integration of AI into SGs also raises several concerns. The development of AI-powered games requires significant technological infrastructure and expertise [11]. Concerns have been raised regarding data privacy, potential algorithmic biases, and inequities related to access and fairness [15,17]. To address these challenges, recent literature emphasizes the need for ethical design principles, secure data practices, and transparent model development to ensure trustworthy and inclusive applications of AI in SGs.

AI-based SGs offer various potential benefits for meta-skills training, but they also introduce important challenges that require mitigation. One key advantage is improved access to personalized training, as AI can adapt learning content and game difficulty to individual player capabilities and provide instant feedback with continuous performance tracking. AI systems can also support a deeper understanding of teamwork, communication, leadership, and behavioral patterns and enable training to be delivered at scale, at low cost, and remotely. However, concerns remain about how AI systems make decisions, the risk of confusing or misleading players, and an increasing dependence on automated feedback. Additional challenges include reduced human interaction, risks related to storing and handling sensitive data, and the possibility of biased or unfair outcomes if AI models are trained on unrepresentative datasets. To address these issues, the literature highlights the importance of transparent and responsible AI design, the use of simple and explainable adaptation rules, balancing automated feedback with human guidance, and maintaining opportunities for reflection and social interaction. Robust data protection measures, including encryption and restricted access, as well as regular bias checks using diverse datasets, are also essential to ensure ethical and effective deployment of AI-based SGs [11,14,15,17-21].

SGs are interactive experiences created with objectives that extend beyond mere entertainment, serving functions such as education, skill development, or problem-solving [22]. They typically represent real-world systems or processes in a simplified manner, enabling participants to engage with and better understand the complexities of real-life structures through a distilled, manageable format [23]. One common implementation of SGs is through escape rooms, where groups of individuals collaborate to solve complex puzzles under given time constraints. During such SGs, an expert technician monitors the teams and assesses their behavioral and psychological attributes and constructs, for example, the abilities of collaboration, decision-making, and leadership. These attributes are often quantitative as scores are assigned on a Likert scale.

The primary game outcome, that is, if the team completed the task within the allotted time, is also recorded. This data is then analyzed through a variety of methods to explore how different team attributes and compositions influence their performance and game outcome. With a promising potential to provide insights into the factors that contribute to team success, SGs are now increasingly recognized as a tool for enhancing teamwork, collaboration, and problem-solving [24].

Despite significant progress in general sports analytics [7,25], the application of ML in analyzing team performance within serious gaming environments and escape rooms is rather limited. The data acquired through escape rooms is usually analyzed using traditional statistical methods, for example, 1-way ANOVA, which cannot capture the multivariate nature of several variables and their interactions, limiting our ability to achieve deeper and accurate insights into complex team behaviors [9]. Some previous studies on serious gaming have focused on aspects such as team cohesion and leadership dynamics [26,27], but the quantification of specific behaviors contributing to team success is not yet fully understood. The growing body of literature on ML in sports analytics already highlights the utility of multivariate models in understanding team performance [4], but its application in serious gaming is yet to be realized. This gap highlights the need to apply data-driven, multivariate analytical approaches to understand the factors that drive team success. This knowledge will eventually contribute to forming better teams, which could lead to improved efficiency and productivity in organizational settings.

In this study, we aim to address the above-described research gap by applying exploratory and ML techniques to identify the multivariate feature combinations which influence team performance in serious gaming, specifically within escape room settings. Various ML-based models were developed using a serious gaming dataset to study key indicators in team behavior and dynamics that can distinguish successful teams from unsuccessful ones. In addition, explainable AI methods were used to tease out the most significant feature combinations that played a decisive role in determining game outcomes. This knowledge can likely be translated into terms of understanding to develop more efficient and productive teams in organizational and business settings.

Methods

Data Collection

This study used a cross-sectional retrospective design, as the outcome data had already been collected, and the analysis focused on examining existing data to explore potential relationships. This study was conducted in collaboration with JGM Serious eXperiences, a provider of escape rooms and serious gaming environments designed for corporate and educational training based in Leeuwarden, The Netherlands. The research partnership began in November 2022 and included regular in-person meetings between the research team and JGM staff through July 2023. The concept of SGs, as implemented by JGM, involves using traditional entertainment games for educational and instructional objectives [24]. Using SGs for educational purposes makes it possible to measure how

game-based activities influence participant behavior, supporting goals in education and organizational development.

The data collection process was carried out in multiple phases. First, an expert technician documents key demographic information before participants engage in the escape room experience. This includes variables such as age, gender distribution, team size, and prior escape room experience. These demographic factors provide context for the interpretation of behavioral patterns. Any personal data from the consent form of participants is stored in compliance with the relevant privacy regulations. Additionally, participation in JGM's escape room is fully voluntary. Participants are asked to read and sign an informed consent form, and a verbal agreement is also explained to all participants during the introduction session. Participants are also free to withdraw at any time and are properly debriefed afterward.

Second, during the escape room activity, the actions of the participants were continuously recorded using closed-circuit television cameras. Alongside, the expert technicians rated predefined behavioral indicators on a Likert scale, focusing on three constructs: (1) collaboration, (2) communication, and (3) leadership. The behavioral model guiding these observations was adapted from the Anesthetists' Non-Technical Skills (ANTS) Framework [28] and tailored for escape room settings in cooperation with psychologists and academics specialized in personal leadership. Although the ANTS Framework was designed for clinical settings, it draws on general teamwork and organizational behavior theory by focusing on observable behaviors that contribute to effective team performance [28]. The framework entails both interpersonal skills, for example, communication, teamwork, leadership, and cognitive skills, such as situation awareness and decision-making [28], which are critical in high-pressure collaborative environments such as escape rooms. In order to train nontechnical skills effectively, it is essential to first identify which skills are required in a specific environment and then assess them in a structured way. This makes it possible to give clear feedback and evaluate whether training leads to measurable improvement [28]. In collaboration with psychologists and experts in personal leadership, JGM adapted this framework to the escape-room environment, ensuring that each behavioral indicator reflected established constructs in teamwork research while remaining observable. This theoretical foundation helps connect our results to well-established models (such as the ANTS framework) of how teams operate.

Third, upon completion of the escape room session, the participants engaged in a structured reflection process. They reviewed the video recordings of their performance, received feedback from team members, and discussed insights from the analyzer report on aspects related to collaboration, communication, and leadership. The expert technicians also provided personalized feedback, highlighting strengths and areas for improvement on an individual basis. The disagreement of participants on a certain item in the feedback served as a catalyst for discussion on discrepancies between perceived and observed behaviors. Conversely, when there was an agreement, the focus shifted toward the strategies for translating these behavioral insights into real-world contexts, such as professional

settings. This reflective exercise not only served as a learning intervention but also reinforced the reliability of the observations recorded by expert technicians.

Various behavioral components for which the data were recorded on a Likert scale are shown in Figures 1-3. Figure 1 depicts the visual representation of the collaboration construct, Figure 2 shows the communication construct, and Figure 3 shows the leadership construct. For example, Figure 1 illustrates the construct of collaboration, which includes behavioral

components such as coordinating activities, knowledge sharing, and environmental awareness. Each of these components is assessed using 2 opposite indicators (central vs decentralized). These indicators are measured through specific variables, and the responses are used to calculate a score for each indicator. The scores reflect how well a team performs in each behavioral area. For instance, a score above 3.5 suggests strong team unity and coordination (central), while a score below 2.5 indicates a more individual, less coordinated approach (decentralized).

Figure 1. Graphical representation of the collaboration construct.

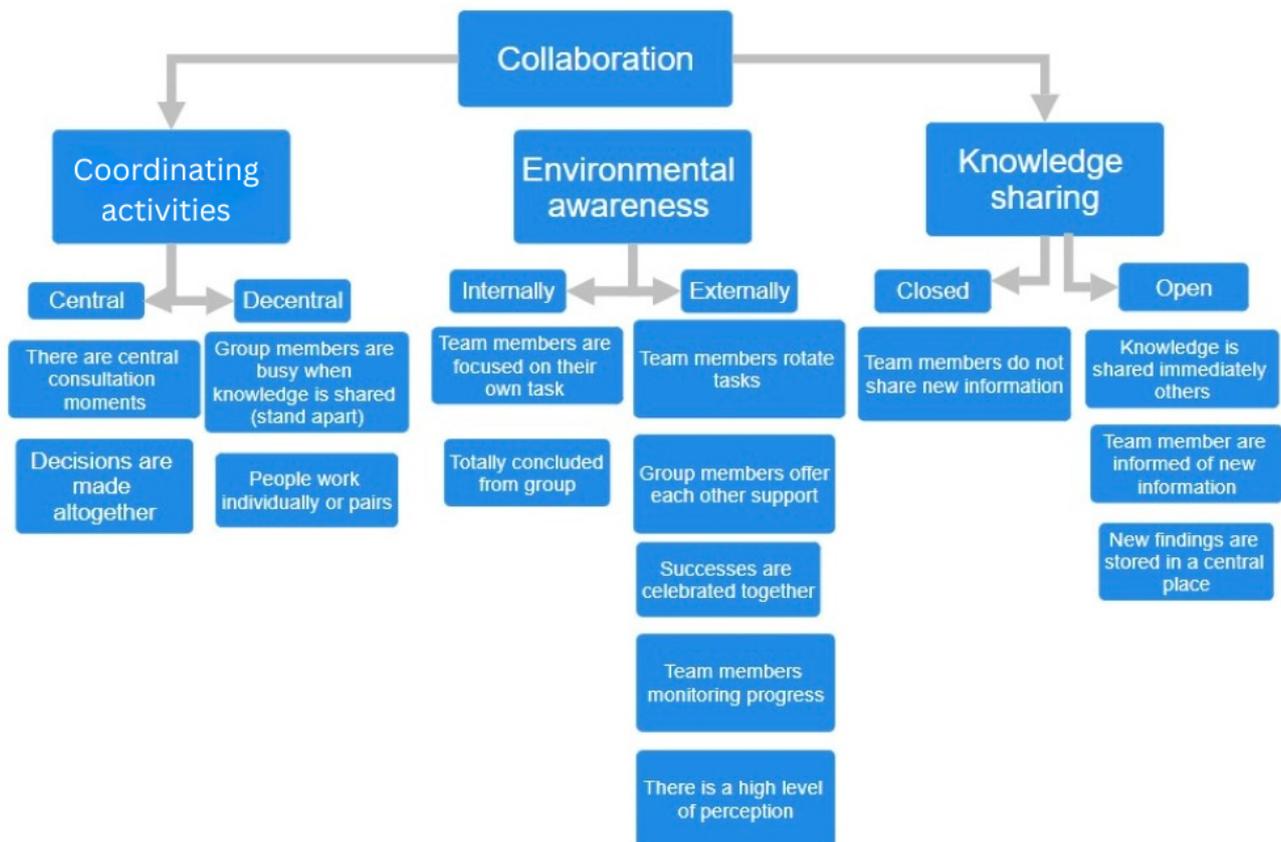


Figure 2. Graphical representation of the communication construct.

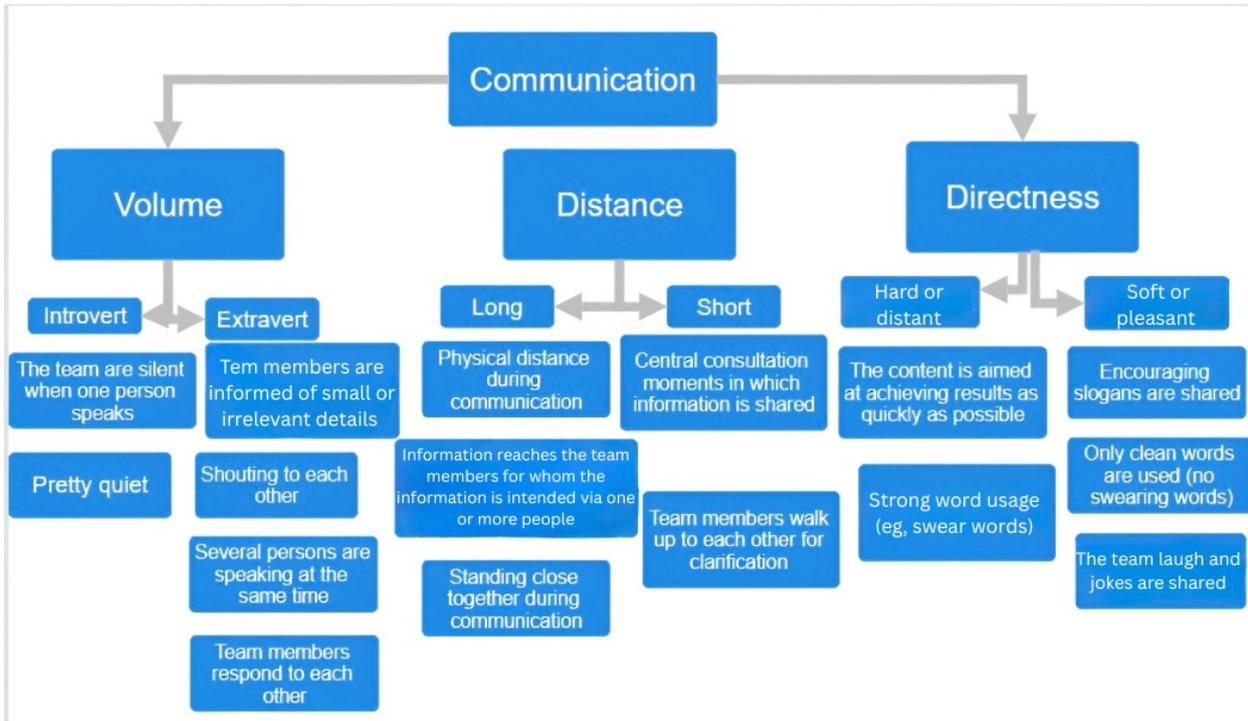
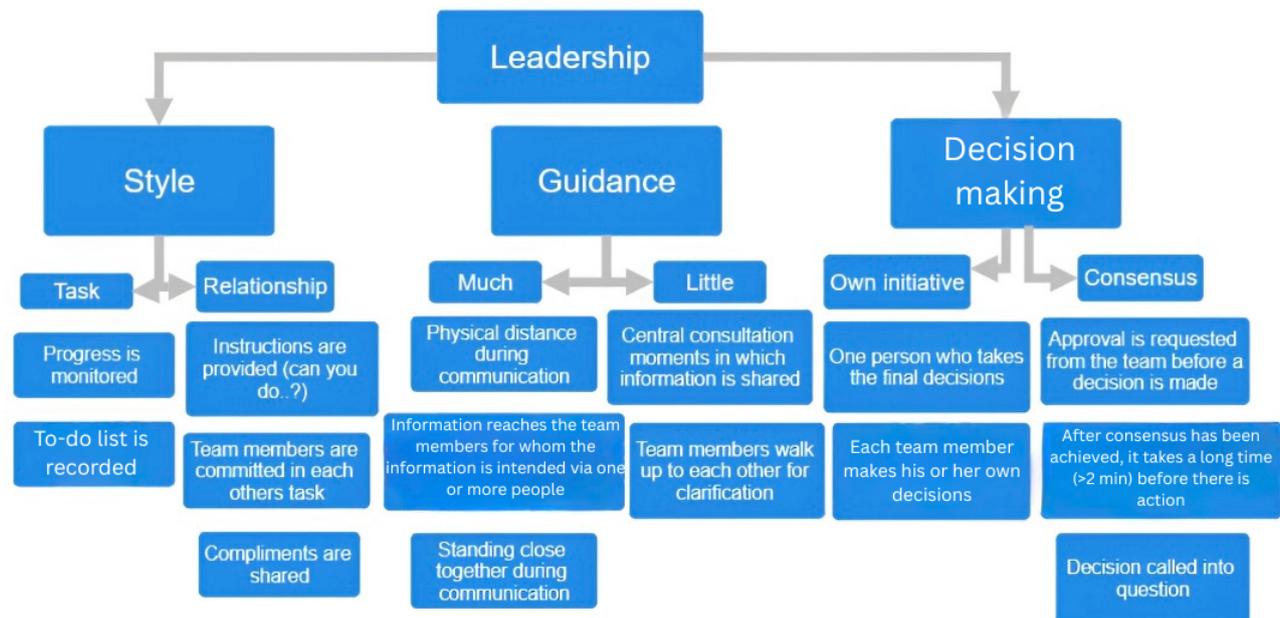


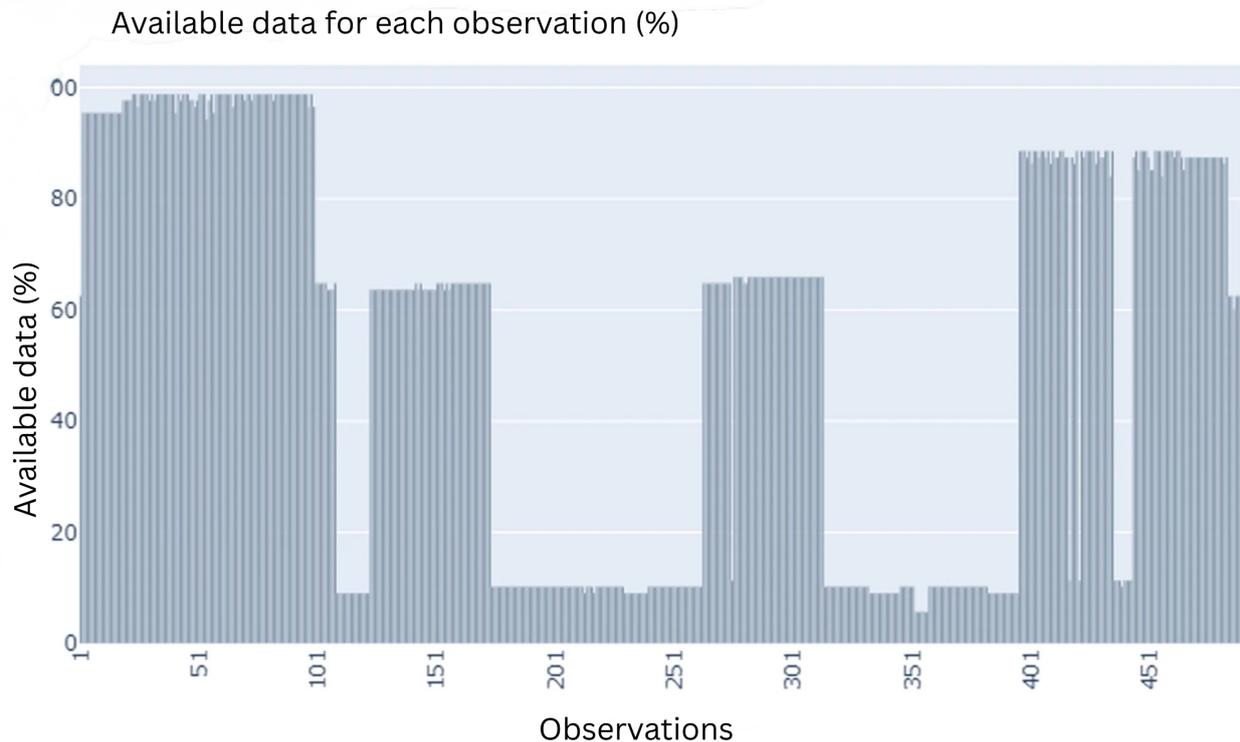
Figure 3. Graphical representation of the leadership construct.



Data Preparation and Exploratory Analysis

Once the collection of data had been completed, JGM’s expert technicians stored the dataset in Excel (Microsoft Corp) spreadsheet in compliance with privacy regulations. This resulted in a raw dataset containing 88 features, 490 rows, and 20,246 missing values. Figure 4 illustrates the percentage of available data per observation, grouped in bins of 50 observations. It indicates that missingness is not uniformly distributed across the dataset. The first 100 observations show very high data completeness, with approximately 95% - 100%

available data. Several subsequent observation blocks (eg, around observations 100 - 250 and 300 - 380) show substantially lower completeness, often around 5% - 15% available data, indicating a high level of missingness in these cases. Other observation ranges (eg, around observations 250 - 300 and 400 - 470) show moderate to high completeness, typically between 60% - 90% available data. Overall, Figure 4 shows that missing values are unevenly distributed, with some observations containing substantially more missing data than others.

Figure 4. Visualization of the percentage of available data for each observation.

The data preparation was performed using the Pandas library (version 2.2.3) in Python (Python Software Foundation) programming language (version 3.13.1) in Jupyter notebooks. All analyses were performed on an AMD Ryzen (9 processor, 40 GB memory; Advanced Micro Devices, Inc) laptop running Microsoft Windows 11 Pro (version 24H2; Microsoft Corp), provided by University of Groningen (Campus Fryslân). All ML models were developed and tested using the SciKit-Learn (version 1.6.1) library. The visualizations were performed using Plotly and Seaborn packages in a Python programming environment.

To handle missing data, all features and rows that contained more than 50% empty values were dropped. A threshold of 50% provided a balance between retaining as much data as possible. Additionally, rows that contained empty values in columns that were deemed not to be imputable were dropped. Lastly, rows that contained >50% empty values were also dropped, due to there not being enough data to impute them accurately. Any last empty values were imputed using a k-nearest neighbors (KNN) algorithm. The resulting dataset contained 53 features and 233 rows, consisting of 141 winning teams and 92 losing teams.

After the data preparation, an exploratory data analysis (EDA) was conducted to gain an initial understanding of the behavioral constructs and their association with team performance. The dataset was grouped by game outcome, that is, winning and losing teams, enabling comparative assessments across both groups. All 3 constructs, that is, collaboration, communication, and leadership, were analyzed by aggregating individual variables into relevant composite indices. For example, the scores of all behavioral indicators in the leadership construct

were added to calculate the value of the leadership construct, and so on.

For each composite variable, statistical descriptors such as mean and SE around the mean were calculated and visually analyzed. In addition, statistical tests for significance were applied to compare the differences between winning and losing teams. The Student *t* test was applied using independent samples, where the assumptions of normality were satisfied, and Mann-Whitney U tests based on Shapiro-Wilk tests [29] were used where the assumption of normality was violated.

Subsequently, to examine the correlations between behavioral indicators, Pearson correlation analysis was applied and visualized using heatmaps. The correlation analysis quantified the linear associations between behavioral indicators, which were further quantified by fitting a regression model into the scatter plot of respective variable groups. The quantification of linear associations visualized the extent of the difference between behavioral indicators exhibited by the winning and losing teams. The EDA resulted in some interesting insights about underlying group differences between winning and losing groups in the escape room activities.

ML-Based Modeling

In order to develop ML models for classifying between winning and losing teams, the preprocessed dataset was split into train and test sets in a proportion of 80% and 20%, respectively. Due to the unbalanced nature of the dataset, the split was stratified with respect to the target variable, which ensured an equal proportion of the target variable between the train and test sets.

All models were developed in the Scikit-Learn library (version 1.6.1) using the Python programming language (version 3.13.1)

with necessary preprocessing steps, including feature scaling and feature selection. The purpose of feature selection was to enhance the performance of models and reduce potential overfitting. It was implemented using the SelectKBest method [30] with the chi-squared test as its scoring function to evaluate statistical relationships between each feature and the target variables (win or lose) and selecting only the most relevant features based on their chi-squared scores.

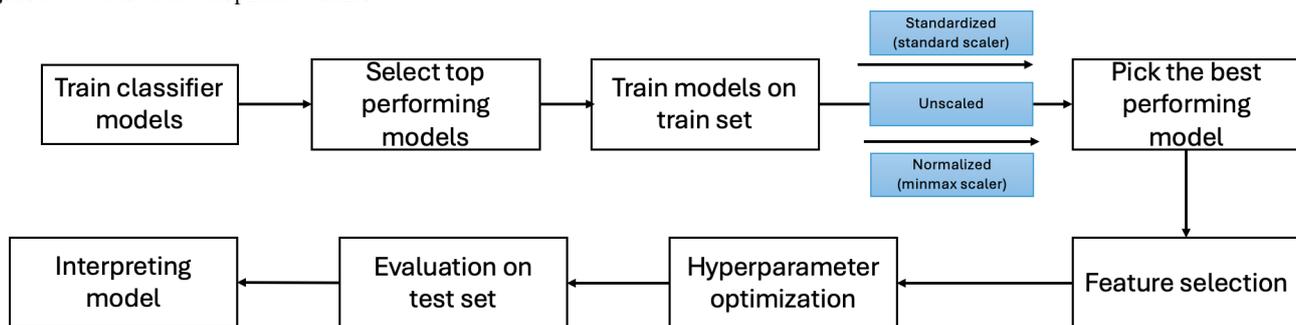
A selection of supervised classification models, namely: support vector classifier (SVC), KNN, Gaussian NB, RF, multilayer perceptron (MLP), gradient boosting classifier, logistic regression (LR), decision tree, adaptive boosting classifier, ridge classifier, and linear discriminant analysis was trained twice, once on unscaled dataset and once on a standardized dataset, and subsequently evaluated using a 5-fold cross-validation with F_1 -scoring. The hyperparameters were optimized using the GridSearch method, based on the highest 5-fold cross-validation with F_1 -scoring (≤ 3240 fits per model). The optimized hyperparameters were mostly related to the regularization complexity of the models. For SVC, these are C (regularization parameter) and kernel and gamma (kernel coefficient). For RF,

these are number of estimators, maximum depth, and minimum samples needed for a split and a leaf. For MLP, the size of the hidden layers, the activation function, the solver, and the maximum allowed iterations. Lastly, LR was optimized on the C (regularization parameter), the penalty function, and the maximum iterations.

Lastly, the best-performing model was selected for interpretation and to identify the importance of each feature in shaping the final prediction. This was achieved by (1) extracting the model coefficients to analyze their weights associated with each feature and (2) the SHAP (or Shapley Additive Explanations) [31] values, which were determined using the SHAP library (version 0.48).

This study followed the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) reporting guideline to ensure transparency and completeness in reporting the development and evaluation of the ML models [32]. A completed TRIPOD checklist has been included as a supplementary file (Checklist 1). The overall workflow used for developing, optimizing, and interpreting the ML models is illustrated in Figure 5.

Figure 5. N - Model Development Workflow.



Ethical Considerations

This study involved secondary analysis of anonymized behavioral data collected independently by JGM Serious eXperiences as part of their standard team-training activities. The Campus Fryslân Ethics Committee confirmed that formal ethical approval was not required for this study. This determination is consistent with the committee's review procedures, under which studies involving the secondary analysis of fully anonymized data with no identifiable information or participant interaction are classified as Tier 1 (no formal review required). The study was conducted in accordance with established ethical standards and institutional guidelines, including the World Medical Association Declaration of Helsinki and relevant data protection regulations such as the General Data Protection Regulation (GDPR). All data were handled in compliance with applicable privacy and confidentiality standards.

All participants who engage in JGM's escape-room training activities are required to read and sign a written informed consent form before participation. Participants are informed about (1) the purpose of the activity, (2) the presence of trained observers, (3) the nature of the behavioral data collected, and (4) their right to withdraw at any time. The provided JGM

consent form explicitly states that behavioral and observational data may be used anonymously for scientific research and that all identifying fields (name, age, and email) are removed before data storage and sharing. The consent form confirms that participants grant permission for the anonymized secondary use of their data for research purposes.

To protect participant confidentiality, JGM removes all personally identifying information before data are uploaded to their internal database, otherwise known as the team analyzer. The research team only received anonymized, team-level variables (eg, communication, collaboration, and leadership scores) with no access to names or any other identifying fields. All data was shared securely and stored on encrypted university servers. No reidentification of individuals was possible, as the research team had no access to raw video, audio, or personal data.

The research team did not compensate participants. Participation in JGM sessions occurred voluntarily as part of organizational training programs or recreational use, and JGM's standard activity fees (if applicable) were unrelated to the research component. No images, videos, or audio recordings of participants are included in this paper or supplementary materials. Therefore, no consent for identifiable images was required. All analyses were conducted solely on anonymized

numerical and categorical variables extracted from JGM's internal behavioral coding system.

Results

Characteristics of Data

The final dataset consisted of 233 team-level observations, with an imbalanced distribution of outcomes: 141 teams completed the escape room game (ie, winning teams), whereas 92 teams did not achieve success (ie, losing teams). The team sizes varied between 2 and 8 members, where most of the teams consisted of 4 or 5 members. The smaller teams (2 - 3 members) and the larger teams (6 - 8 members) were relatively uncommon, which resulted in most of the teams in the dataset being midsized. The experience in escape rooms was also recorded on an individual level and then aggregated at the team level. A significant number of teams ($n=126$) reported no experience. In contrast, a smaller group ($n=25$) indicated moderate familiarity, and an even smaller number ($n=9$) reported a higher level of prior engagement.

The mean age of participants across all teams was 25.8 years. The majority of teams ($n=107$) had a mean age of 24.9 years, whereas only 6 teams had a mean age above 50. In terms of gender composition, 53 teams were composed entirely of male members, whereas 55 teams were entirely composed of female members. The remaining teams were composed of both genders, and the gender proportion varied significantly across groups.

In summary, the dataset captures a diverse sample of teams, mainly composed of younger participants, with limited prior escape room experience and mixed gender composition. These demographic and background characteristics provide context for interpreting teams' performance in the escape room environment.

About EDA

Overall, winning teams showed higher scores in environmental awareness, leadership, guidance, and extraversion, with CIs that were narrow and showed little overlap, indicating more reliable group differences. In contrast, constructs such as coordination, knowledge sharing, consensus, and method of communication showed overlapping CIs, suggesting that any mean differences between winners and losers are small and less precise.

For environmental awareness, the mean difference between winners and losers was 2.33 (95% CI 1.27 to 3.38). Although this suggests a noticeable gap, the Shapiro-Wilk test indicated nonnormality, and the Mann-Whitney U test showed no statistically significant difference ($U=9595.00$, $P=.68$). This means the construct does not reliably distinguish the 2 groups.

Knowledge sharing showed a smaller mean difference of 0.43 (95% CI 0.01 to 0.84). Despite marginal overlap in CIs, the Mann-Whitney U test revealed a statistically significant difference ($U=13,687.50$, $P<.001$), indicating that winners scored meaningfully higher than losers.

Leadership also showed a clear difference, with a mean gap of 1.21 (95% CI 0.46 to 1.95). Normality was violated, so a Mann-Whitney U test was used, and it confirmed a statistically significant difference ($U=12,502.00$, $P<.001$). Guidance showed a similar pattern, with a mean difference of 1.27 (95% CI 1.46 to 2.47) and strong statistical significance ($U=14,577.00$, $P<.001$).

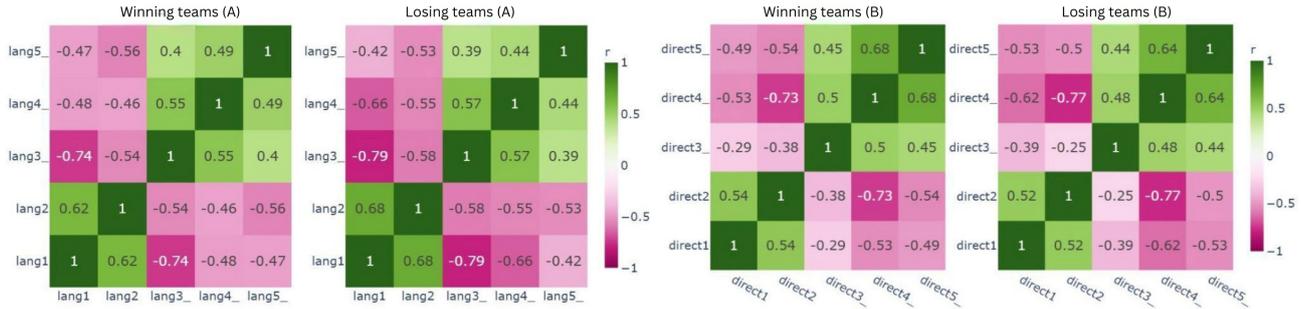
Extraversion of communication had a mean difference of 0.77 (95% CI 0.28 to 1.25), also supported by a significant Mann-Whitney U result ($U=12,347.00$, $P<.001$). Other communication-related constructs showed weaker differences. Directness of communication had a mean difference of 0.55 (95% CI 0.10 to 0.99), but the Mann-Whitney U test showed no significant difference ($U=10,842.50$, $P=.17$). Consensus within the team slightly favored losers with a mean difference of -0.34 (95% CI -0.86 to 0.19), but again, results were not significant ($U=8917.00$, $P=.16$). Method of communication showed a small and nonsignificant mean difference of 0.35 (95% CI -0.03 to 0.74; $U=10,842.50$, $P=.17$).

Coordination showed virtually no difference between groups with a mean difference of -0.09 (95% CI -0.49 to 0.32). This was the only construct meeting normality assumptions for both groups, yet the Mann-Whitney U test still indicated no significant difference ($U=9595.00$, $P=.68$).

In summary, only 4 constructs: knowledge sharing, leadership, guidance, and extraversion, showed statistically significant differences between winners and losers. These findings, supported through CIs, normality testing, and appropriate nonparametric analyses, suggest that only a subset of behavioral indicators meaningfully distinguish successful teams.

In addition, Pearson correlation analysis was used to analyze variable level relationships between the winning and losing groups. Figure 6 shows correlation maps for the communication construct. The winners displayed positive correlations between some individual variables (eg, lang1 and lang2, $r=0.62$) and negative correlations involving the lang3_ variable (eg, lang1 and lang3_, $r=-0.74$). In contrast, the losing teams mostly exhibited negative associations, suggesting a less differentiated structure.

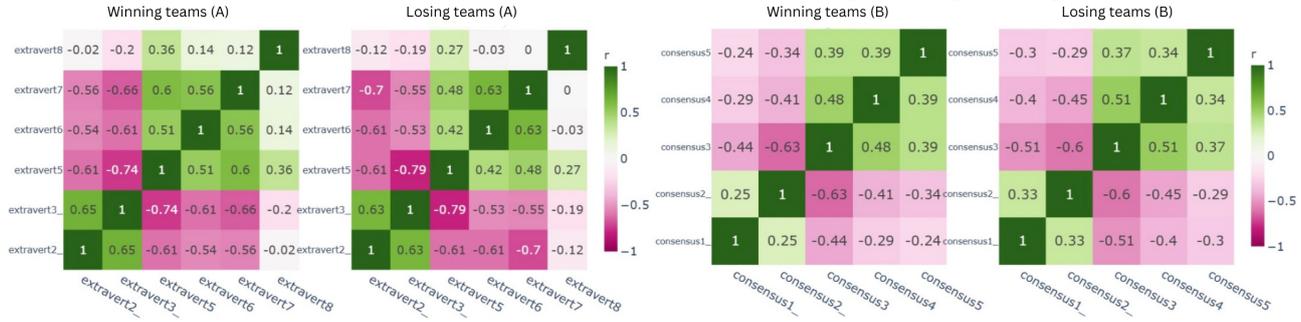
Figure 6. Pearson correlation for (A) communication and (B) directness of communication in winning and losing teams.



Within the “directness of communication” construct, a negative correlation was found between direct4_ (use of only clean words) and direct2 (use of strong language, including swearing), with coefficients of $r=-0.73$ for winners and $r=-0.77$ for losing teams. As both groups show a similar pattern, this construct does not differentiate between winners and losers as shown in Figure 6B.

This mirrored structure was also evident in other constructs. Figure 7A shows the extraversion of communication construct, extravert3 (lack of active communication) and extravert5 (continuous communication) were negatively correlated ($r=-0.74$ for winners; $r=-0.79$ for losing teams). Similarly, within the consensus construct in Figure 7B, consensus3 (waiting for team approval before acting) and consensus2 (seeking advice but acting independently) demonstrated a similarly negative correlation ($r=-0.63$ for winners; $r=-0.60$ for losing teams).

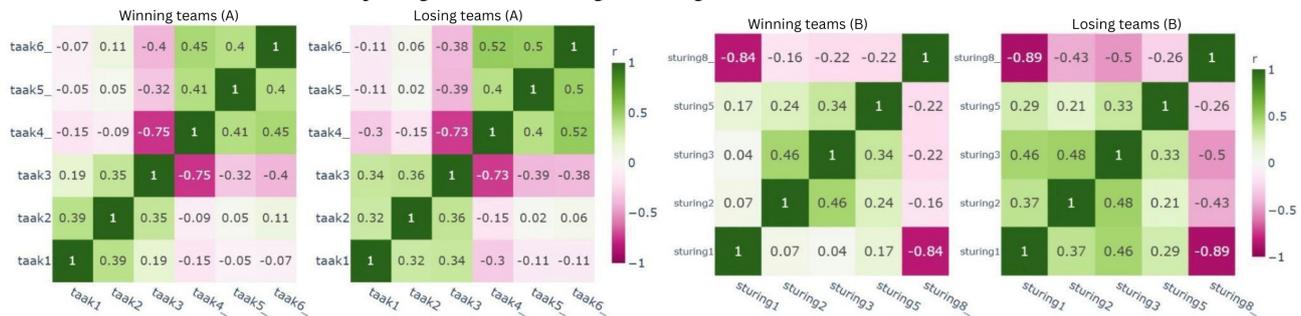
Figure 7. Pearson correlation for (A) extraversion of communication and (B) consensus-building in winning and losing teams.



Furthermore, Figure 8A displays variables in the leadership construct, showing taak4 (instructions are provided) and taak3 (group members give each other instructions) as inversely related, with correlation coefficients of $r=-0.75$ for winners and $r=-0.73$ for losing teams. Similarly, the guidance construct in

Figure 8B also displayed redundant variables between sturing1 (there is group activity) and sturing8_ (the group has a wait-and-see attitude), with correlation coefficients of $r=-0.84$ for winners and $r=-0.89$ for losing teams.

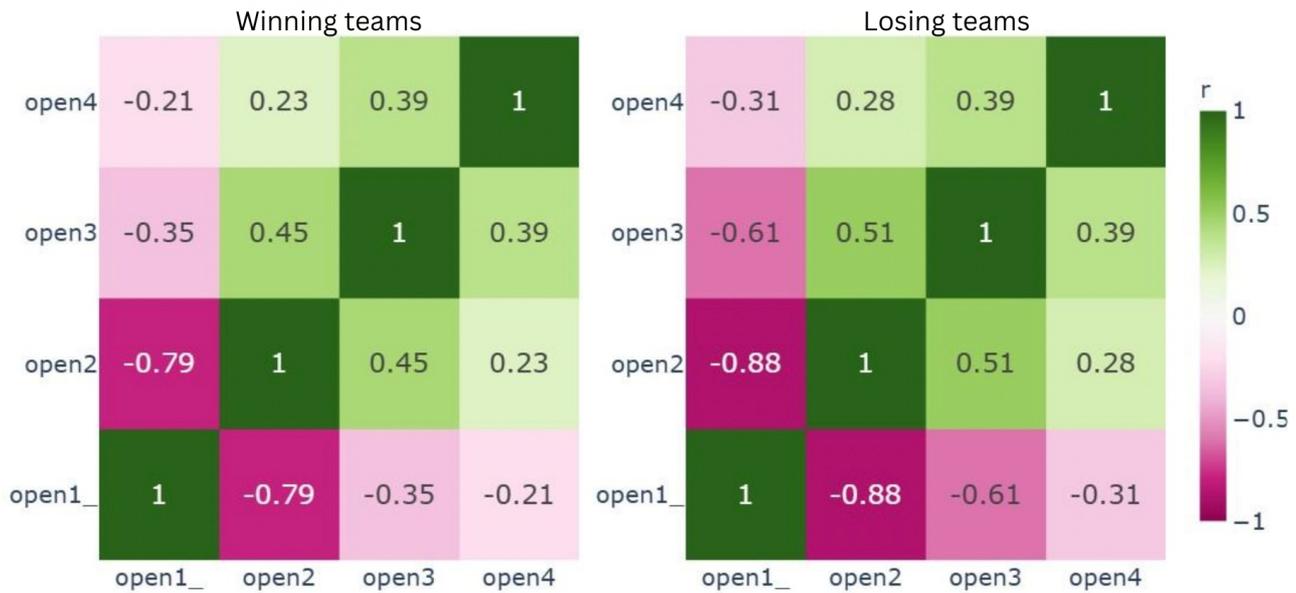
Figure 8. Pearson correlation for leadership and guidance in winning and losing teams.



The final construct showing overlap between variables was the sharing knowledge construct. For instance, open2 (knowledge is shared immediately with the group after learning about it) and open1_ (team members keep the knowledge they have gained to themselves) mirrored each other in winning teams,

with correlation coefficients of $r=-0.79$ for winners and $r=-0.88$ for losing teams. These relationships are illustrated in Figure 9, which presents the correlation patterns within the sharing knowledge construct for winning and losing teams.

Figure 9. Pearson correlation for sharing knowledge in winning and losing teams.



However, not all constructs exhibited opposing relationships. Constructs such as coordination within the team and environmental awareness showed generally inconsistent correlations among their variables, suggesting that these behavioral dimensions may capture more independent or multifaceted aspects of team dynamics. Unlike the previously discussed constructs, no clear inverse patterns emerged, indicating that the variables within these domains may reflect complementary rather than redundant behaviors.

Taken together, the exploratory analysis using descriptive statistics and Pearson correlation provided insights into group-level behavioral differences and variable relationships. Descriptive comparisons revealed that winners tended to score higher in constructs related to leadership, guidance, extraversion, and environmental awareness, whereas losing teams displayed less variation across several domains. Furthermore, the correlation analyses highlighted numerous instances of inverse

relationships within constructs, such as communication, directness, extraversion, and leadership, suggesting the presence of opposing behavioral features.

Performance of ML Models

Of the 12 models evaluated, 4 classifiers including LR, MLP, RF, and SVC, demonstrated consistently strong performance on F_1 -score across cross-validation. These models were subsequently selected for hyperparameter optimization and interpretation. The remaining models, including Gaussian NB, KNN, and others, showed substantially weaker performance and were therefore excluded from further analysis. Table 1 presents various evaluation metrics associated with 4 ML models developed to classify between winning and losing teams. The results were obtained by evaluating all 4 models on an unseen test set. In addition, 5-fold stratified cross-validation was performed using the F_1 -score as the accuracy metric (Table 1).

Table 1. ML^a model scores.

Model	Test-set scores				Cross-validation scores (F_1 -score)
	Accuracy (%)	Precision (%)	Recall (%)	F_1 -score (%)	Mean (SD, %)
Logistic regression	85	84	93	88	79 (0.037)
Support vector classifier	81	79	93	85	81 (0.044)
Multilayer perceptron	83	81	93	87	80 (0.039)
Random forest	83	81	93	87	80 (0.072)

^aML: machine learning.

Across the board, all 4 models yielded reasonable accuracies of over 80%. Interestingly, all models had an equal recall score of 93%, with slight differences in terms of precision score, where the LR model slightly outperformed the other models. The LR model also outperformed the other models in terms of accuracy and F_1 -score, which is a better representation of the

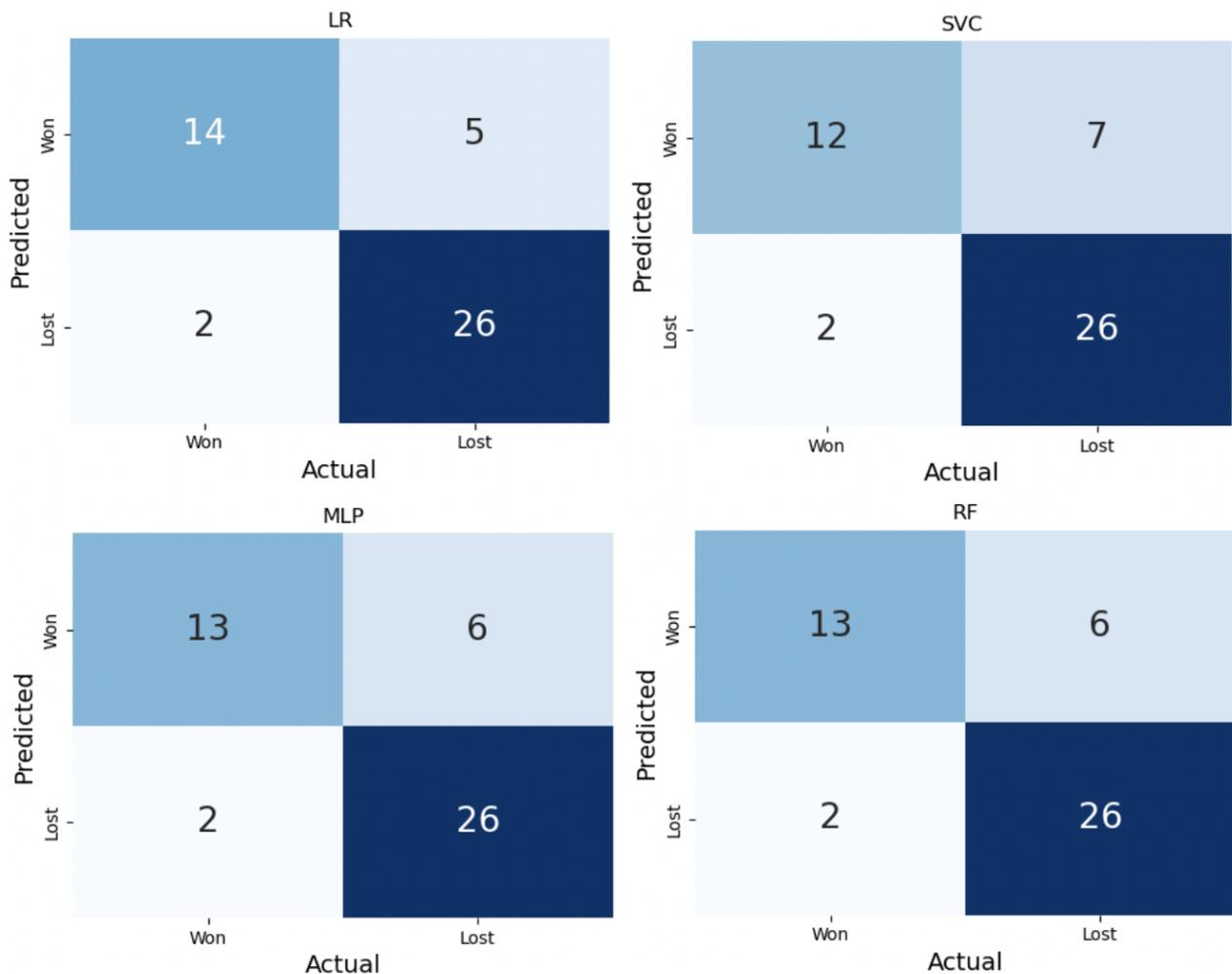
model's performance given a significant imbalance in the target variable. The SVC model accuracy was the lowest on the test set, but it outperformed all other models in terms of cross-validation score. On the other hand, the LR model scored the lowest score on the cross-validation metric (2 percentage points lower than SVC). However, the difference between the

accuracies and cross-validation scores for all 4 models was not substantially different from each other (within 5% range).

Figure 10 presents the confusion matrix for the 4 ML models used to classify team outcomes as “won” or “lost” based on team behavioral indicators. The MLP and RF models both correctly identified 13 winning teams and 26 losing teams, with 6 false positives and 2 false negatives. The SVC model performed similarly, but with slightly more false positives (ie,

7) among winning classifications. In contrast, the LR model demonstrated the optimal balance, that is, correctly classifying 14 winning teams with only 5 false positives and 2 false negatives. Across all models, the number of false negatives stayed consistently low, but the number of false positives varied substantially. The LR model, in particular, achieved the optimal trade-off between sensitivity and specificity in predicting team success.

Figure 10. Confusion matrix for 4 models: LR, SVC, MLP, and RF. LR: logistic regression; MLP: multilayer perceptron; RF: random forest; SVC: support vector classifier.



Explainability of Model

All 4 ML models were trained on 10 features that were selected during the feature selection process. To explain the contribution of each feature to the final classification outcome, the feature importances were extracted using various techniques based on model type. For the LR model, these are the coefficients, which

are the internal weights of the model (Figure 11). For RF, its Gini impurities quantify how much a feature helped in decreasing uncertainty (Figure 12). The feature importances of MLP (Figure 13) and SVC (Figure 14) were quantified using permutation importances, a method that finds the most important features by shuffling feature inputs.

Figure 11. Feature importance of logistic regression

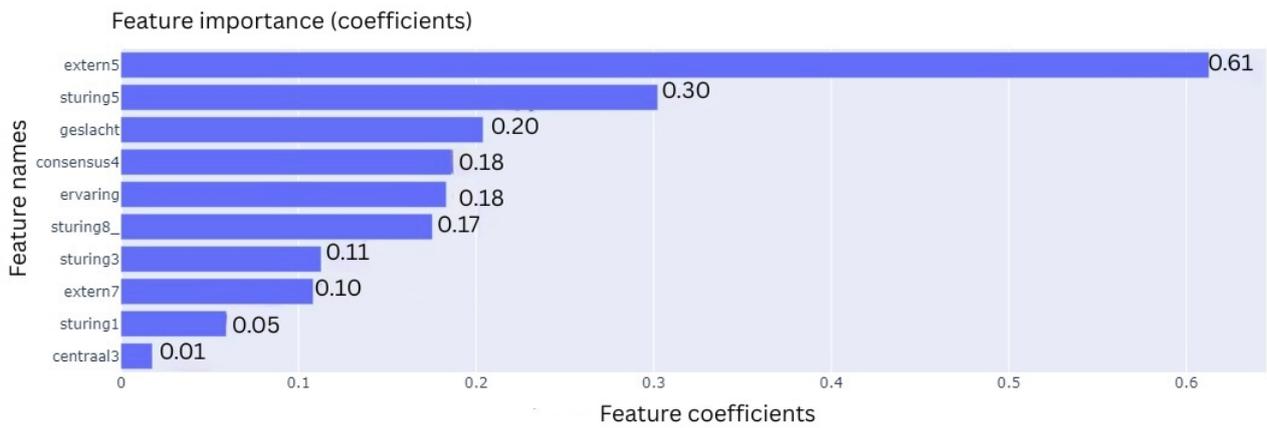


Figure 12. Feature importance of random forest.

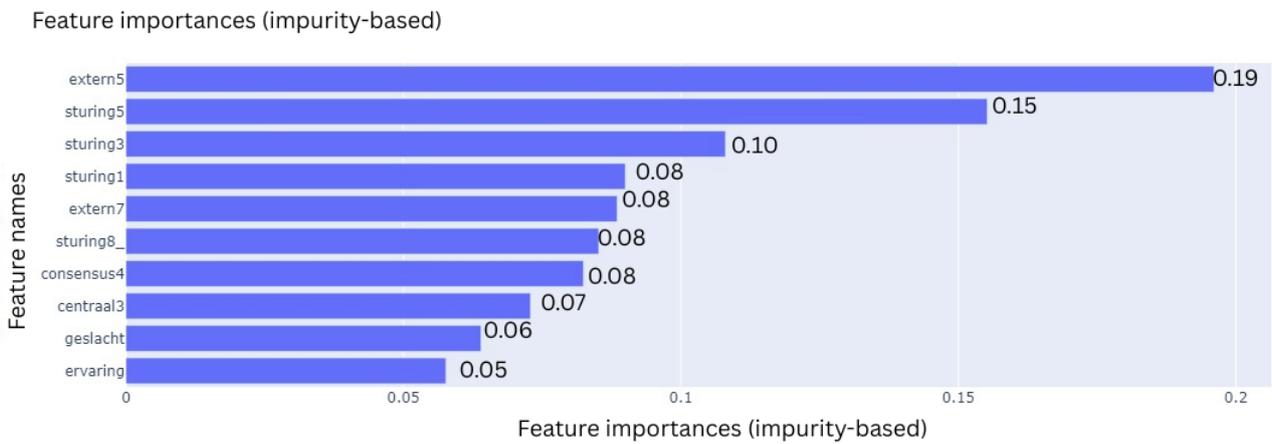


Figure 13. Permutation importance of the support vector classifier

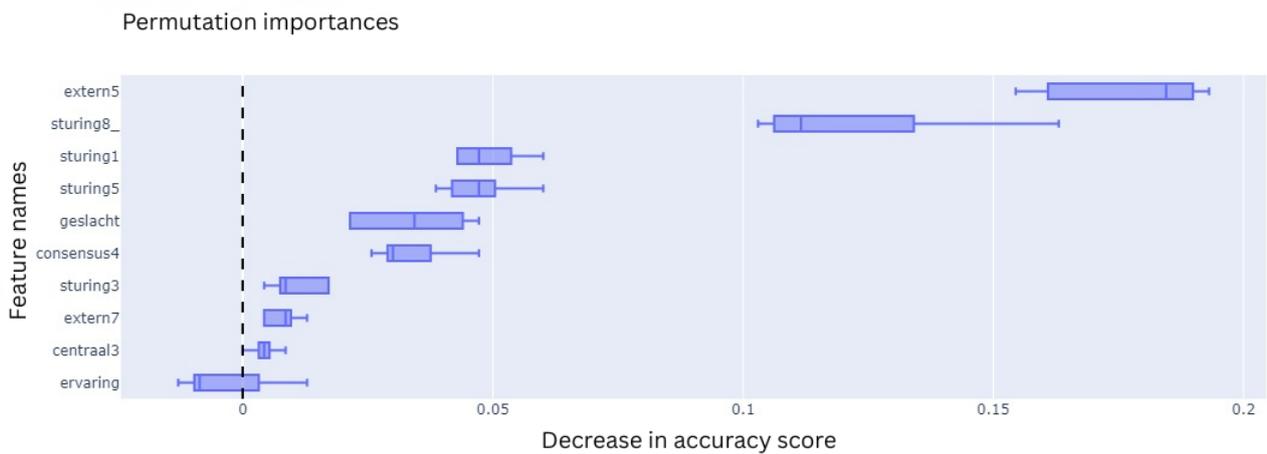
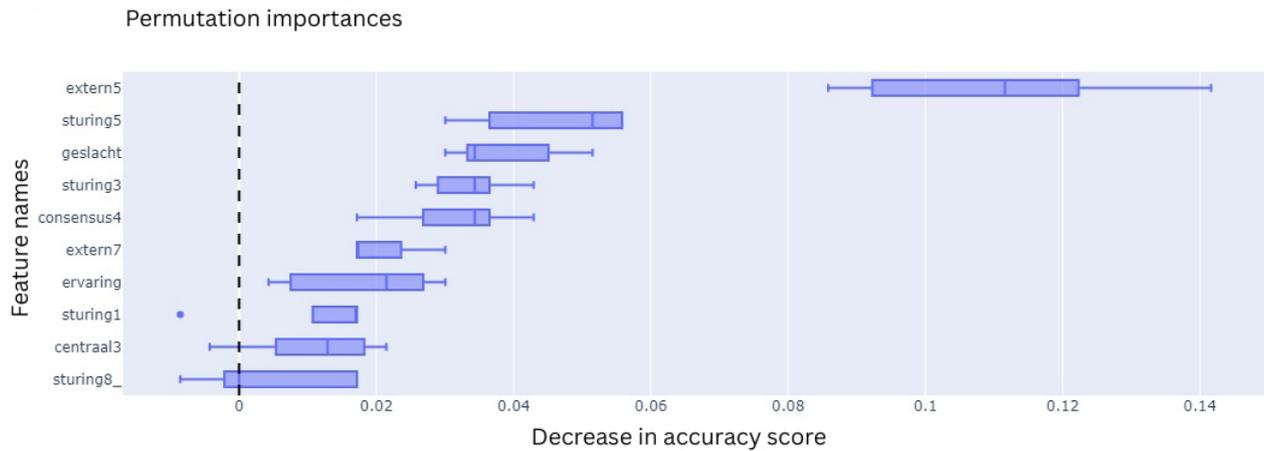
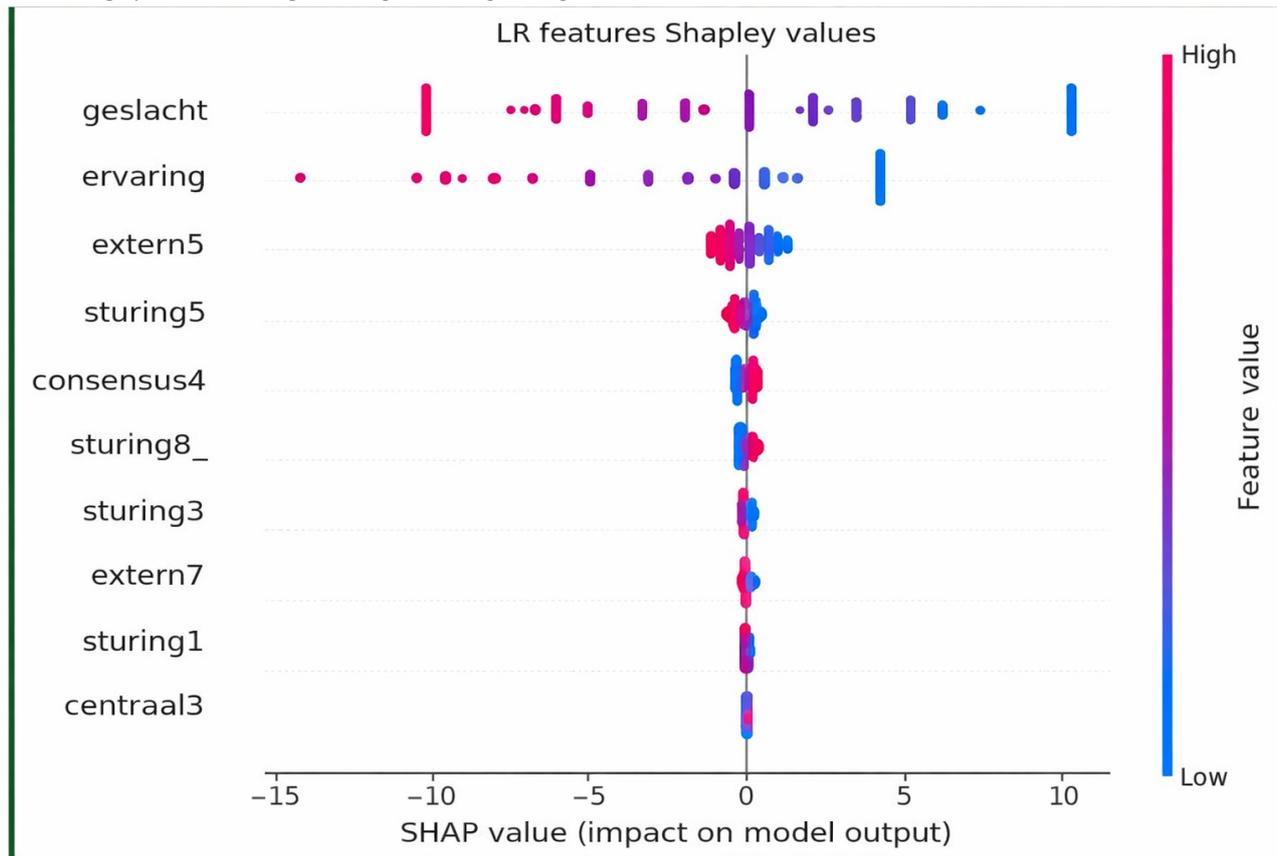


Figure 14. Permutation importance of the support vector classifier.

Across the models, *extern5* (team members celebrating progression with each other) and *sturing5* (someone taking the lead when the team is stuck) consistently ranked among the top features contributing to the prediction of team success. In the LR model, *extern5* had the highest absolute coefficient, followed by *sturing5*, *geslacht* (gender distribution), and *consensus4* (time to action after consensus). In the RF model, impurity-based feature importance ranked *extern5*, *sturing5*, *sturing3* (guidance is provided through questioning and progress monitoring), and *sturing1* (all team members are actively involved in tasks to find solutions as a group) as the top 4 features. In contrast, the MLP model showed higher permutation importance for *extern5*, *sturing8_* (the team has a wait-and-see attitude), and *sturing1*, with *sturing5* ranking fourth. For the SVC model, permutation importance highlighted *extern5*, *sturing5*, *geslacht*, and *sturing3* as the most influential features. While the top-ranked features varied slightly across models, *extern5* emerged as the most consistently important predictor.

In addition, SHAP values were computed for the best-performing model, that is, the LR model, to identify the features with the greatest impact on the model's output (Figure 15). The most influential features included *geslacht* (gender composition) and *ervaring* (prior escape room experience), both of which showed the highest SHAP value ranges. These were followed by *extern5* (celebrating progress) and *sturing5* (taking the lead when the team is stuck), which had consistently positive SHAP contributions. Other features, such as *consensus4* (waiting after consensus) and *sturing8_* (passive behavior), exhibited SHAP values that were negatively associated with the model output, whereas *sturing3*, *extern7*, and *sturing1* had comparatively lower but consistent influence on predictions. The SHAP summary plot thus reflects a combination of demographic and behavioral predictors, with varying degrees of contribution to the classification of team success.

Figure 15. Shapley values of (best performing model) logistic regression.



Discussion

Principal Findings

This study set out to examine whether team success in SGs can be better understood through a multivariate combination of behavioral and demographic factors, rather than through isolated indicators. The findings show that team performance in escape room-based SGs is shaped by multivariate patterns of behavior. Successful teams were not distinguished by single actions alone, but by coordinated behavioral dynamics related to leadership, guidance, and expressive communication, alongside demographic characteristics such as team composition and experience. By applying exploratory analysis together with ML and explainable AI techniques, this study demonstrates that these factors jointly contribute to predicting successful team outcomes. Overall, the results confirm this study's objective by showing that team success in SGs emerges from complex, multivariate behavioral patterns rather than from individual behavioral indicators considered in isolation.

Interpretation of Behavioral Patterns in Team Performance

The exploratory analyses highlight clear behavioral differences between winning and losing teams in serious gaming environments. Winning teams tended to display stronger behavioral patterns related to leadership, guidance, environmental awareness, and expressive communication. These behaviors appeared more consistently and coherently among successful teams, whereas losing teams showed greater variability. In contrast, behaviors such as coordination and

knowledge sharing were present in both groups but did not clearly differentiate successful from unsuccessful teams, suggesting that these behaviors alone are not decisive for performance in escape-room contexts.

Differences also emerged in how behaviors were combined within teams. Winning and losing teams often relied on similar behavioral elements, but they appeared to integrate them in distinct ways. For example, successful teams showed more balanced patterns of support, initiative, and shared progress, whereas losing teams tended to exhibit more uneven behavioral interactions. This indicates that team success may depend less on the presence of individual behaviors and more on how these behaviors interact during gameplay.

In particular, behaviors related to guidance and environmental awareness revealed contrasting interaction patterns between groups. While both winning and losing teams engaged in information seeking and guidance-related actions, losing teams appeared to rely more on reactive or delayed coordination, whereas winning teams demonstrated more adaptive behavioral combinations. These findings suggest that winning teams are not simply more active but are better at aligning leadership, guidance, and collective awareness in response to evolving game demands.

Overall, the exploratory findings indicate that team success in SGs is associated with multivariate behavioral patterns rather than isolated actions. Winning teams appear to combine leadership, guidance, and communicative behaviors in a more integrated manner, supporting the view that team performance

emerges from multivariate behavioral dynamics rather than single behavioral indicators.

Interpretation of ML and Explainable AI Findings

The ML analyses further support this multivariate interpretation of team performance. Across all 4 models, `extern5` and `sturing5` are overall the 2 most important features to determine whether a team will win or not. `Extern5` refers to “team members celebrating progression with each other,” whereas `sturing5` refers to “someone taking the lead when the team is stuck.” The significance of these 2 features displays the importance of good leadership and exuberance. The importance of the other features varied from 1 model to another and had a moderate impact on the final prediction. For the LR model, these were `geschlacht` and `consensus4`, which reflect the “gender distribution of a team” and “the time it takes for a team to take action after consensus has been reached.” `Sturing3` and `sturing1` (which refer to “a team’s activity level” and “guidance provided”) were ranked third and fourth, respectively, with a moderate impact for the RF model. MLP was the only model for which `sturing5` was not one of the two most important features, ranking only fourth after `sturing8_` and `sturing1`, which also reflect a team’s activity level. Lastly, `geschlacht` and `sturing3` (which refer to “gender distribution of a team” and “guidance provided”) were the third and fourth most important features of the SVC model.

The SHAP explainable AI technique revealed that the most impactful features influencing team success were gender composition and escape room experience. Specifically, teams with a higher proportion of female members and greater experience were more likely to be classified as winners. The behavioral traits such as sharing positive emotions when progress was made (`extern5`) and guiding the team during uncertainty (`sturing5`) had a mild but positive influence on success. In contrast, passive or hesitant behaviors, such as waiting for others to act (`consensus4` and `sturing8_`), tended to decrease the likelihood of winning. The model also found that early enthusiasm (`extern7`) and collaborative monitoring (`sturing3`) were weak but consistent predictors of better performance. These results suggest that successful teams managed uncertainty not simply by dividing tasks, but by enabling collaborative leadership, expressing collective motivation, and establishing situational awareness.

It is important to note that the key variables identified through quantification of feature importances of the ML models and those highlighted by the SHAP explainability technique were not always identical. This divergence is due to the fundamental differences in how these methods evaluate the relationship between predictors and the target outcome. The ML models rank features based on their predictive contribution to classification performance, which the algorithm’s internal mechanics influences [33] (eg, regularization in LR, impurity reduction in RF, or learned weights in MLP).

In contrast, SHAP values rank features in a model-agnostic way. They provide an estimate of the marginal contribution of each feature by simulating its effect across all possible combinations of feature inputs. SHAP estimates how much each feature contributes to the predictions by considering its effect across different combinations of inputs. In the results, SHAP

highlighted demographic factors such as gender distribution and experience, whereas the ML models gave more weight to behavioral indicators such as `extern5` and `sturing5`. Together, these methods offer complementary perspectives, showing that both demographic and behavioral factors played a role in shaping team outcomes. Recognizing these methodological differences is critical for interpreting the results holistically, as each method offers complementary insights into what drives team success. In sum, predictive models compute outcomes, whereas SHAP explains the underlying contribution of features to those outcomes [33].

Comparison With Existing Literature

Prior work [34] emphasized the importance of team familiarity and communication across subgroup faultlines in shaping performance. Using survival models, they found that team familiarity could exert a negative influence on success, particularly when communication across subgroups was limited [34]. While this study did not explicitly model faultlines or social ties, it identifies behavioral and demographic features that contribute to team success, potentially through similar mechanisms of coordination, engagement, and leadership. For instance, `sturing5` (taking initiative when the team is stuck) and `sturing3` (guiding others and monitoring progress) reflect adaptive leadership and teamwork, which help teams stay on track, especially in unfamiliar or fast-paced situations [34]. Furthermore, a study [27] conducted one of the few observational studies into collaboration within escape rooms, emphasizing the central role of verbal communication, implicit leadership emergence, and social awareness. Their qualitative insights on players communicating loudly across the room, navigating spatial awareness through verbal cues, and different leadership styles [27] echo the behavioral constructs we identified as significant. For example, their observations that experienced players often assume leadership roles [27] align with the predictive power of `ervaring` (experience) and `sturing5` (leading when stuck) in our models. Their finding that teams converge when stuck [27] also aligns with the identification of `sturing1` (early active problem-solving) and `sturing3` (task-related guidance) as positive predictors.

Beyond the context of SGs, our findings are aligned with a broader trend in applied ML, where complex, multifactor systems are modeled using data-driven and often hybrid approaches. Recent studies in smart city energy forecasting, groundwater management, CO₂ emissions prediction for electric vehicles, and rainfall classification have combined traditional statistical models or neural networks with metaheuristic optimization and feature-selection techniques to handle high-dimensional, noisy, and context-dependent data while improving predictive accuracy and efficiency [35-39]. In a similar sense, our study uses EDA, supervised ML, and explainable AI to identify a set of behavioral and demographic predictors that capture the significant variables of team performance in escape rooms. Rather than focusing on metaheuristic optimization [35,37], we emphasize data preprocessing, feature selection, and interpretability as practical strategies for organizations that work with limited or expert-coded data. This positions our work within a growing body of research showing that ML techniques can reveal

meaningful patterns in complex systems, whether environmental or behavioral, when combined with domain knowledge and transparent model evaluation. These methodological parallels across domains illustrate a broader trend in ML toward transparent, data-driven analysis, which aligns with the approach taken in this study.

This study makes several contributions to research on team performance in SGs. Methodologically, it advances the field by combining EDA, supervised ML, and explainable AI to examine team success as a multivariate behavioral phenomenon rather than as the outcome of isolated factors. This approach differs from much of the existing serious-games literature [10-13,15,17,21], which has primarily relied on univariate statistical comparisons or qualitative observations [27] and has rarely integrated interpretable ML to explain why certain teams succeed. By identifying multivariate influential behavioral and demographic predictors across multiple models, the current study contributes empirical evidence on which observable team behaviors meaningfully distinguish successful from unsuccessful teams in escape-room settings, which can be translated in terms of improved productivity in business and organization settings.

Implications for Understanding Team Performance in SGs

These findings have several implications for the design, evaluation, and application of SGs as tools for team training and assessment. First, identifying which behavioral patterns meaningfully distinguish successful from unsuccessful teams can support the refinement of observer-coded behavioral frameworks. Rather than relying on a large number of overlapping indicators, SG designers and expert observers can focus on a smaller set of behaviors that capture leadership, guidance, and collective engagement more effectively. This may improve the efficiency of behavioral assessment in game-based training environments.

Second, the results can inform the selection of performance indicators used during and after gameplay. By highlighting behavioral combinations associated with success, SGs can move beyond simple outcome measures, such as winning or losing, toward richer evaluations of how teams function under pressure. This opens opportunities for developing feedback tools that emphasize team strengths and areas for improvement, either through debriefings or automated postgame reports. Such debriefings can support reflection, learning, and skill development by making abstract team dynamics more concrete and observable.

Finally, this study demonstrates the practical value of interpretable ML for translating complex behavioral data into actionable insights. By combining predictive models with explainable techniques, it becomes possible to generate transparent and understandable explanations of team performance rather than opaque scores or rankings. This is particularly important in educational and organizational settings, where trust, accountability, and learning are central concerns [40,41]. Interpretable ML thus offers a promising pathway for integrating data-driven analysis into SGs in a way that supports understanding, feedback, and responsible decision-making.

Limitations of Research

While the findings in this study provide some new insights into team performance in escape room–based serious gaming, several limitations need to be acknowledged.

First, the limited sample size could restrict the generalizability of our results. Additionally, with a significant class imbalance toward losing teams (141 losing teams and 92 winning teams), this could have introduced bias in the model for that specific group. This imbalance could bias model training because standard ML algorithms often assume that classes are equally represented. Prior research shows that when 1 class is larger, models tend to favor the majority class, leading to poorer detection of the minority class or even treating minority cases as noise [42,43]. Class imbalance is a well-known challenge across many applied domains, such as fraud detection, medical diagnosis, and intrusion detection, and can distort decision boundaries, reduce sensitivity to the minority group, and limit overall model fairness [42,43]. Although several methods have been proposed to manage imbalance, there is no universal solution suitable for all datasets [43]. In this study, the unequal distribution between winning and losing teams may therefore influence model performance and should be considered when interpreting the results.

Second, the limitation concerns the dataset size. Although machine-learning studies in fields such as image recognition often rely on datasets with tens of thousands of samples, behavioral research involving team performance rarely allows for such scale due to practical and resource constraints. Our dataset of 233 teams reflects these real-world limitations. As a result, model generalizability is reduced, and findings should be interpreted as proof of concept rather than fully generalizable predictions. Future research with larger, multi-institutional samples would be valuable for validating and extending these insights.

Third, the dataset is geographically and culturally limited to teams recruited in a specific demographic region. The cultural norms around communication, leadership, and emotional expression may influence gameplay behavior in ways that are not generalizable to teams in other countries or contexts. As a result, the applicability of the findings to international, intercultural, or remote teams may be constrained to 1 specific region only.

Additionally, this study focuses exclusively on a single game genre (escape rooms) within 1 specific serious gaming environment. While this provides insight into a certain domain of games, it might have limited transferability to other SGs involving different structures and dynamics, for example, simulations, role-playing, or strategy-based games, where behavioral markers and success criteria may differ substantially.

Finally, it is important to note that the conclusions drawn from this study are based on a cross-sectional retrospective design and therefore reflect associative and predictive relationships rather than causal effects. While the ML models identify behavioral and demographic patterns that are strongly associated with team success, they do not imply that modifying individual behaviors will necessarily lead to improved outcomes. Instead,

the findings should be interpreted as evidence of multivariate patterns that characterize successful teams within the observed serious gaming context. Within these methodological boundaries, this study provides a robust, data-driven foundation for understanding team dynamics and for informing the design of feedback, training, and adaptive systems in SGs.

Ethical and Interpretability Implications

This study relied on the secondary use of fully anonymized behavioral data collected by JGM Serious eXperiences as part of routine training activities. All participants provided written informed consent, all identifying information was removed before data storage and sharing, and the research team received only deidentified, team-level variables.

Predictive analytics can infer sensitive traits or behavioral tendencies from seemingly harmless data, a concern described as “predictive privacy” [44]. This occurs when models generate inferences about individuals or groups without their awareness, potentially leading to differential treatment or unintended consequences. While our study does not involve such high-stakes decisions, the concept remains relevant for future applications of ML-based team assessment in organizational settings.

Another key issue concerns fairness and bias. Predictive systems can unintentionally reproduce existing social inequalities if model outputs correlate with attributes such as gender, experience, or group composition [44,45]. Demographic variables such as gender composition and experience were influential predictors in this study, reflecting patterns present in the dataset rather than normative judgments. SHAP values were used in this study to provide transparent explanations of model outputs, improving clarity around why certain predictions were made. Prior research shows that explainable AI can

increase trust and informed decision-making in human-AI systems [40], whereas poorly explained feedback may negatively affect user engagement and performance [41]. As ML-based evaluation becomes more common, clear ethical safeguards and transparent design are needed to ensure that behavioral insights support learning rather than unfairly labeling participants.

Conclusion

This study demonstrates that team success in serious gaming environments is best understood as a multivariate phenomenon emerging from the interaction between behavioral dynamics and team composition, rather than from isolated actions or traits. By integrating exploratory analysis with ML and explainable AI, the findings show that leadership, guidance, collective engagement, and contextual team characteristics jointly shape performance outcomes under time pressure. More broadly, this study highlights the value of treating SGs not only as training tools but also as structured environments for observing and analyzing complex team behavior using transparent, data-driven methods.

Beyond prediction, the results have broader implications for how SGs and AI-supported assessment systems are designed and deployed. Explainable AI techniques (eg, SHAP) enable behavioral data to be translated into understandable and actionable insights, supporting feedback, reflection, and learning rather than opaque scoring. Consistent with prior work showing that transparency and explainability influence how humans engage with AI systems [40,41], this study suggests that AI-supported SGs can enhance team development when human behavioral insight and explainable analytics are combined responsibly. In this way, the work contributes to a broader understanding of how human-AI systems can support effective teamwork, not by replacing human judgment, but by augmenting it with interpretable, evidence-based insight.

Acknowledgments

The authors would like to thank JGM Serious eXperiences for providing data for this study. The authors declare the use of generative artificial intelligence (GenAI) was not used in the research and writing process. According to the GAIDeT [46] (Generative Artificial Intelligence Delegation Taxonomy; 2025), no generative AI tools were used in the preparation of this manuscript. Responsibility for the final paper lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the outcomes. Declaration submitted by: GGA-R.

Funding

The authors declare that no funds, grants, or other forms of financial support were received for the preparation of this paper.

Data Availability

The datasets analyzed during this study are available in the GitHub repository [47].

Conflicts of Interest

None declared.

Checklist 1

TRIPOD checklist.

[[DOCX File, 91 KB - games_v14i1e83478_app1.docx](#)]

References

<https://games.jmir.org/2026/1/e83478>

JMIR Serious Games 2026 | vol. 14 | e83478 | p.213
(page number not for citation purposes)

1. PV A, K P. Revolutionizing sports excellence through machine learning. *IJFMR* 2024;6(2). [doi: [10.36948/ijfmr.2024.v06i02.16581](https://doi.org/10.36948/ijfmr.2024.v06i02.16581)]
2. Schumaker RP, Solieman OK, Chen H. Predictive modeling for sports and gaming. In: *Sports Data Mining*: Springer US; 2010:55-63. [doi: [10.1007/978-1-4419-6730-5_6](https://doi.org/10.1007/978-1-4419-6730-5_6)]
3. Samuel AL. Some studies in machine learning using the game of checkers. II—recent progress. *IBM J Res Dev* 1967 Nov;11(6):601-617. [doi: [10.1147/rd.116.0601](https://doi.org/10.1147/rd.116.0601)]
4. Giannakas F, Troussas C, Krouska A, Sgouropoulou C, Voyiatzis I. Multi-technique comparative analysis of machine learning algorithms for improving the prediction of teams' performance. *Educ Inf Technol* 2022 Jul;27(6):8461-8487. [doi: [10.1007/s10639-022-10900-4](https://doi.org/10.1007/s10639-022-10900-4)]
5. Lamsal R, Choudhary A. Predicting outcome of Indian Premier League (IPL) matches using machine learning. arXiv. Preprint posted online on Sep 21, 2020. [doi: [10.48550/arXiv.1809.09813](https://doi.org/10.48550/arXiv.1809.09813)]
6. Liu H, Gomez M, Lago-Peñas C, Sampaio J. Match statistics related to winning in the group stage of 2014 Brazil FIFA World Cup. *J Sports Sci* 2015;33(12):1205-1213. [doi: [10.1080/02640414.2015.1022578](https://doi.org/10.1080/02640414.2015.1022578)] [Medline: [25793661](https://pubmed.ncbi.nlm.nih.gov/25793661/)]
7. Adhikari B, Ahamed M. Predicting BPL match winners: an empirical study using machine learning approach. Presented at: 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT); Jul 6-8, 2023; Delhi, India p. 1-7. [doi: [10.1109/ICCCNT56998.2023.10306698](https://doi.org/10.1109/ICCCNT56998.2023.10306698)]
8. Pramanik M, Suzan MH, Biswas AA, Rahman MZ, Kalaiarasi A. Performance analysis of classification algorithms for outcome prediction of T20 cricket tournament matches. In: 2022 International Conference on Computer Communication and Informatics (ICCCI): IEEE; 2022:01-07. [doi: [10.1109/ICCCI54379.2022.9740867](https://doi.org/10.1109/ICCCI54379.2022.9740867)]
9. Malagón-Selma P, Debón A, Ferrer A. Exploring essential variables for successful and unsuccessful football teams in the "Big Five" with multivariate supervised techniques. *Electron J Appl Stat Anal* 2022 May;15(1). [doi: [10.1285/i20705948v15n1p249](https://doi.org/10.1285/i20705948v15n1p249)]
10. Miljanovic MA, Bradbury JS. Engineering Adaptive Serious Games Using Machine Learning', in *Software Engineering for Games in Serious Contexts: Theories, Methods, Tools*: Springer Nature Switzerland; 2023:117-134. [doi: [10.1007/978-3-031-33338-5_6](https://doi.org/10.1007/978-3-031-33338-5_6)]
11. Frutos-Pascual M, Zapirain BG. Review of the use of AI techniques in serious games: decision making and machine learning. *IEEE Trans Comput Intell AI Games* 2017 Jun;9(2):133-152. [doi: [10.1109/TCIAIG.2015.2512592](https://doi.org/10.1109/TCIAIG.2015.2512592)]
12. Westera W, Prada R, Mascarenhas S, et al. Artificial intelligence moving serious gaming: presenting reusable game AI components. *Educ Inf Technol* 2020 Jan;25(1):351-380. [doi: [10.1007/s10639-019-09968-2](https://doi.org/10.1007/s10639-019-09968-2)]
13. Brisson A, Pereira G, Prada R, et al. Artificial intelligence and personalization opportunities for serious games. *AIIDE* 2012;8(5):51-57. [doi: [10.1609/aiide.v8i5.12576](https://doi.org/10.1609/aiide.v8i5.12576)]
14. Hare R, Tang Y. Player modeling and adaptation methods within adaptive serious games. *IEEE Trans Comput Soc Syst* 2023 Aug;10(4):1939-1950. [doi: [10.1109/TCSS.2022.3203926](https://doi.org/10.1109/TCSS.2022.3203926)]
15. Mitsea E, Drigas A, Skianis C. A systematic review of serious games in the era of artificial intelligence, immersive technologies, the metaverse, and neurotechnologies: transformation through meta-skills training. *Electronics (Basel)* 2025 Jan;14(4):649. [doi: [10.3390/electronics14040649](https://doi.org/10.3390/electronics14040649)]
16. Karavidas L, Apostolidis H, Tsiatsos T. Usability evaluation of an adaptive serious game prototype based on affective feedback. *Information* 2022 Sep;13(9):425. [doi: [10.3390/info13090425](https://doi.org/10.3390/info13090425)]
17. Pérez J, Castro M, López G. Serious games and AI: challenges and opportunities for computational social science. *IEEE Access* 2023;11:62051-62061. [doi: [10.1109/ACCESS.2023.3286695](https://doi.org/10.1109/ACCESS.2023.3286695)]
18. Carneiro D, Veloso P. Veloso, 'Ethics, Transparency, Fairness and the Responsibility of Artificial Intelligence', in *New Trends in Disruptive Technologies, Tech Ethics and Artificial Intelligence*: Springer International Publishing; 2022:109-120. [doi: [10.1007/978-3-030-87687-6_12](https://doi.org/10.1007/978-3-030-87687-6_12)]
19. Mansilla A, Ochoa A, Ponce J, Herrera M, Hernandez A, Cossio E. Implementation of an chatbot in a serious game associated with the acquisition of social skills and the promotion of collaborative tasks in children. Presented at: 2017 Twelfth Latin-American Conference on Learning Technologies (LACLO); Oct 9-13, 2017; La Plata p. 1-4. [doi: [10.1109/LACLO.2017.8120912](https://doi.org/10.1109/LACLO.2017.8120912)]
20. Zagal JP, Tomuro N, Shepitsen A. Natural language processing in game studies research. *Simul Gaming* 2012 Jun;43(3):356-373. [doi: [10.1177/1046878111422560](https://doi.org/10.1177/1046878111422560)]
21. Colado IJP, Colado VMP, Morata AC, Píriz RSC, Manjón BF. Using new AI-driven techniques to ease serious games authoring. In: 2023 IEEE Frontiers in Education Conference (FIE): IEEE; 2023:1-9. [doi: [10.1109/FIE58773.2023.10343021](https://doi.org/10.1109/FIE58773.2023.10343021)]
22. Rodela R, Ligtenberg A, Bosma R. Conceptualizing serious games as a learning-based intervention in the context of natural resources and environmental governance. *Water (Basel)* 2019;11(2):245. [doi: [10.3390/w11020245](https://doi.org/10.3390/w11020245)]
23. Greenblat CS. Group dynamics and game design: some reflections. *Simul Games* 1980 Mar;11(1):35-58. [doi: [10.1177/104687818001100106](https://doi.org/10.1177/104687818001100106)]
24. Bottino R, Ott M, Tavella M. Serious gaming at school: reflections on students' performance, engagement and motivation. *IJGBL* 2014 Jan;4(1):21-36. [doi: [10.4018/IJGBL.2014010102](https://doi.org/10.4018/IJGBL.2014010102)]

25. Casal CA, Losada JL, Barreira D, Maneiro R. Multivariate exploratory comparative analysis of LaLiga teams: principal component analysis. *Int J Environ Res Public Health* 2021 Mar 19;18(6):3176. [doi: [10.3390/ijerph18063176](https://doi.org/10.3390/ijerph18063176)] [Medline: [33808634](https://pubmed.ncbi.nlm.nih.gov/33808634/)]
26. Cohen TN, Griggs AC, Keebler JR, et al. Using escape rooms for conducting team research: understanding development, considerations, and challenges. *Simulation Gaming* 2020 Aug;51(4):443-460. [doi: [10.1177/1046878120907943](https://doi.org/10.1177/1046878120907943)]
27. Pan R, Lo H, Neustaedter C. Collaboration, awareness, and communication in real-life escape rooms. *Assoc Comput Mach* 2017 Jun 10;1353-1364. [doi: [10.1145/3064663.3064767](https://doi.org/10.1145/3064663.3064767)]
28. Flin R, Patey R, Glavin R, Maran N. Anaesthetists' non-technical skills. *Br J Anaesth* 2010 Jul;105(1):38-44. [doi: [10.1093/bja/aeq134](https://doi.org/10.1093/bja/aeq134)] [Medline: [20522911](https://pubmed.ncbi.nlm.nih.gov/20522911/)]
29. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. *Int J Endocrinol Metab* 2012;10(2):486-489. [doi: [10.5812/ijem.3505](https://doi.org/10.5812/ijem.3505)] [Medline: [23843808](https://pubmed.ncbi.nlm.nih.gov/23843808/)]
30. Pedregos F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;12(85):2825-2830 [FREE Full text]
31. An introduction to explainable AI with shapley values. SHAP. URL: https://shap.readthedocs.io/en/latest/example_notebooks/overviews/An%20introduction%20to%20explainable%20AI%20with%20Shapley%20values.html#An-introduction-to-explainable-AI-with-Shapley-values [accessed 2026-03-16]
32. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024 Apr 16;385:e078378. [doi: [10.1136/bmj-2023-078378](https://doi.org/10.1136/bmj-2023-078378)] [Medline: [38626948](https://pubmed.ncbi.nlm.nih.gov/38626948/)]
33. Ekanayake IU, Meddage DPP, Rathnayake U. A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley Additive Explanations (SHAP). *Case Stud Constr Mater* 2022 Jun;16:e01059. [doi: [10.1016/j.cscm.2022.e01059](https://doi.org/10.1016/j.cscm.2022.e01059)]
34. O Szabo R, Chowdhary S, Deritei D, Battiston F. The anatomy of social dynamics in escape rooms. *Sci Rep* 2022 Jun 22;12(1):10498. [doi: [10.1038/s41598-022-13929-0](https://doi.org/10.1038/s41598-022-13929-0)] [Medline: [35732634](https://pubmed.ncbi.nlm.nih.gov/35732634/)]
35. Elkenawy ESM, Alhussan AA, Eid MM, Ibrahim A. Rainfall classification and forecasting based on a novel voting adaptive dynamic optimization algorithm. *Front Environ Sci* 2024 Jun;12:1417664. [doi: [10.3389/fenvs.2024.1417664](https://doi.org/10.3389/fenvs.2024.1417664)]
36. Elkenawy ESM, Alhussan AA, Khafaga DS, Tarek Z, Elshewey AM. Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification. *Sci Rep* 2024 Oct;14(1):23784. [doi: [10.1038/s41598-024-72013-x](https://doi.org/10.1038/s41598-024-72013-x)]
37. Ali Alhussan A, El-Kenawy ESM, Sami Khafaga D, Alharbi AH, Eid MM. Groundwater resource prediction and management using comment feedback optimization algorithm for deep learning. *IEEE Access* 2025;13:169554-169593. [doi: [10.1109/ACCESS.2025.3614168](https://doi.org/10.1109/ACCESS.2025.3614168)]
38. Saqr AES, Saraya MS, El-Kenawy ESM. Enhancing CO2 emissions prediction for electric vehicles using Greylag Goose Optimization and machine learning. *Sci Rep* 2025 May 13;15(1):16612. [doi: [10.1038/s41598-025-99472-0](https://doi.org/10.1038/s41598-025-99472-0)] [Medline: [40360691](https://pubmed.ncbi.nlm.nih.gov/40360691/)]
39. El-Kenawy ESM, Ibrahim A, Alhussan AA, Khafaga DS, Ahmed AEM, Eid MM. Smart city electricity load forecasting using Greylag Goose Optimization-enhanced time series analysis. *Arab J Sci Eng* 2025 Sep 25. [doi: [10.1007/s13369-025-10647-3](https://doi.org/10.1007/s13369-025-10647-3)]
40. Rai A. Explainable AI: from black box to glass box. *J Acad Mark Sci* 2020 Jan;48(1):137-141. [doi: [10.1007/s11747-019-00710-5](https://doi.org/10.1007/s11747-019-00710-5)]
41. Chen A, Lyu A, Lu Y. Member's performance in human-AI hybrid teams: a perspective of adaptability theory. *Inf Technol People* 2026 Jan 14;39(1):157-177. [doi: [10.1108/ITP-05-2023-0513](https://doi.org/10.1108/ITP-05-2023-0513)]
42. Niaz NU, Shahariar KMN, Patwary MJA. Class imbalance problems in machine learning: a review of methods and future challenges. *Assoc Comput Mach* 2022 Mar 10;485-490. [doi: [10.1145/3542954.3543024](https://doi.org/10.1145/3542954.3543024)]
43. Elrahman SMA, Abraham A. A review of class imbalance problem. *J Netw Innov Comput* 2013 Oct;1:9 [FREE Full text]
44. Mühlhoff R. Predictive privacy: towards an applied ethics of data analytics. *Ethics Inf Technol* 2021 Dec;23(4):675-690. [doi: [10.1007/s10676-021-09606-x](https://doi.org/10.1007/s10676-021-09606-x)]
45. Barocas S, Selbst AD. Big Data's disparate impact. *SSRN J* 2016:671-732. [doi: [10.2139/ssrn.2477899](https://doi.org/10.2139/ssrn.2477899)]
46. Suchikova Y, Tsybuliak N, Teixeira da Silva JA, Nazarovets S. GAIDeT (Generative AI Delegation Taxonomy): a taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. *Account Res* 2026 Apr;33(3):1-27. [doi: [10.1080/08989621.2025.2544331](https://doi.org/10.1080/08989621.2025.2544331)] [Medline: [40781729](https://pubmed.ncbi.nlm.nih.gov/40781729/)]
47. Gruyff101/machine-learning-multivariate-analysis. GitHub. URL: <https://github.com/Gruyff101/Machine-Learning-Multivariate-Analysis> [accessed 2026-03-17]

Abbreviations

- AI:** artificial intelligence
- ANTS:** Anesthetists' Non-Technical Skills
- EDA:** exploratory data analysis
- KNN:** k-nearest neighbor

LR: logistic regression

ML: machine learning

MLP: multilayer perceptron

NB: Naive Bayes

RF: random forest

SG: serious game

SHAP: Shapley Additive Explanations

SVC: support vector classifier

TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by S Brini; submitted 03.Sep.2025; peer-reviewed by MM Eid, R Mamidipally, S Xiong; revised version received 20.Jan.2026; accepted 21.Jan.2026; published 13.Apr.2026.

Please cite as:

Abdul-Rahman GG, de Lange F, Zwitter A, Haleem N

Analysis of Machine Learning–Based Investigation Into Multivariate Factors of Team Performance in Serious Games: Cross-Sectional Retrospective Study

JMIR Serious Games 2026;14:e83478

URL: <https://games.jmir.org/2026/1/e83478>

doi: [10.2196/83478](https://doi.org/10.2196/83478)

© Gruyff Germain Abdul-Rahman, Freark de Lange, Andrej Zwitter, Noman Haleem. Originally published in JMIR Serious Games (<https://games.jmir.org>), 13.Apr.2026. 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.

Original Paper

The Effects of Dance-Based Exergaming on Mental Rotation, General Motor Coordination, and Math Achievement in Adolescent Students: Nonrandomized Controlled Pilot Study

Patrick Fargier^{1,2}, PhD, Prof Dr; Valérian Cece¹, PhD; Vanessa Lentillon-Kaestner^{1*}, PhD, Prof Dr; Cédric Roure^{1*}, PhD, Prof Dr

¹University of Teacher Education, State of Vaud (HEP VD), Lausanne, State of Vaud, Switzerland

²University of Lyon, University Claude Bernard Lyon 1, Inter-University Laboratory on Human Movement Biology (LIBM), Villeurbanne, France

*these authors contributed equally

Corresponding Author:

Patrick Fargier, PhD, Prof Dr

University of Teacher Education, State of Vaud (HEP VD)

Avenue de Cour 33

Lausanne, State of Vaud, 1014

Switzerland

Phone: 33 (0)618271153

Email: patrick.fargier@hepl.ch

Abstract

Background: Exergaming can promote adolescent health by encouraging repetition of game-related tasks, potentially contributing to academic success by developing motor and cognitive skills. Studies have highlighted the influence of exergaming on motor skill learning but have not clarified its effects on general motor coordination. Other studies suggest that exergames may influence math success, particularly with nonmathematical, dance-based exergames, possibly through mental rotation or general motor coordination training. However, the influence of a single exergaming sequence on these abilities and on math achievement in the same participants has not been studied.

Objective: This study aimed to determine whether nonmathematical dance-based exergaming (DEX) improves mental rotation, general motor coordination, and math achievement in adolescent students.

Methods: An experimental group (EG: 15 girls and 15 boys; mean age 14.0, SD 0.7 years) and a control group (CG: 14 boys and 12 girls; mean age 14.2, SD 0.9 years), with no contraindications to physical activity or special educational needs, participated in this nonrandomized controlled study. EG performed DEX involving varied locomotion and interlimb coordination, while CG performed precision ball-throwing exergaming (TEX), requiring catching and throwing. Only DEX involved mental rotation; neither of the two involved mathematical content. Both consisted of 5 weekly 45-minute sessions, implemented using the Lü platform. A pretest-posttest design compared effects on (1) a mental rotation test, (2) locomotion and sprint tests, and (3) quantity comparison plus simple and complex addition and multiplication tests. Each session led to noninstrumented observation, and sessions 3 and 5 included instrumented tracking of physical activity and situational interest.

Results: Monitoring of the sequences highlighted a significant between-group difference in triggered situational interest in session 5 (multivariate analysis of variance: $F_{1,54}=9.15$; $P=.004$; $\eta_p^2=0.14$). This variable was added as a covariate in the analyses of covariance (ANCOVAs), with generalized linear model approach performed further. ANCOVA results showed an advantage for EG over CG in terms of mental rotation ($F_{1,52}=6.17$; $P=.02$; $\eta_p^2=0.11$), number of correct simple additions ($F_{1,52}=8.26$; $P=.006$; $\eta_p^2=0.14$), and error rate in complex addition ($F_{1,52}=8.40$; $P=.005$; $\eta_p^2=0.14$). No other significant results were found.

Conclusions: This study helps clarify the previously unexplored influence of the same exergaming sequence on mental rotation and math skills by showing the positive influence of DEX on mental rotation and only on calculations involving mental rotation, according to the literature. Further research on exergaming is needed to clarify whether such improvement in calculation is linked to mental rotation training itself and might be amplified by the development of general motor coordination. However, the study encourages consideration of integrating exergames into active learning approaches in schools.

(JMIR Serious Games 2026;14:e82610) doi:[10.2196/82610](https://doi.org/10.2196/82610)

KEYWORDS

dance-based exergaming; exergaming; general motor coordination; math achievement; mental rotation

Introduction

Problem

The digital revolution belongs to a new techno-economic paradigm that has led to the development of a gaming sector with a commercial and entertainment aim [1]. Since the 1970s, this has led to the creation of different genres of video games played using an audiovisual device to interact with a virtual environment under defined rules, conditions, and complexity levels [2]. Such games became part of popular culture, and researchers gradually saw them as potentially useful for academic learning [3]. Studies reported positive effects of practicing entertainment video games on cognitive functions and abilities that promote school achievement [4,5]. The literature also emphasized that video games that embed a pedagogical aim, thus considered serious games, proved useful for sustaining learning activity over time and promoting learning success [6]. Moreover, manufacturers have developed *active* video games, also known as exergames, for entertainment and educational purposes. Although sharing characteristics with conventional video games, exergames require locomotion and movements of bodily segments, not limited to the use of keyboards, joysticks, or other portable controllers [7]. These video games are thus likely to encourage physical activity, and as a result, to have a significant influence on health but also school achievement [8], which remains to be further examined.

Review of Relevant Scholarship

From the outset, authors have argued that exergames may influence health by favoring physical training and a positive attitude toward physical activity [9]. Consistently, studies showed that exergaming generally elicits, at least in the short term, positive emotions [10-12] and well-being [13,14] in children and adolescents. Other studies showed that the eventual decline over time in the appeal of a given exergame [15] can be avoided by specific practice modalities [10]. Furthermore, the literature reported that such influence on interest may favor energy expenditure [16]. Accordingly, studies found that exergaming may reduce perceived exertion [17], promote intense efforts [18], and have a positive impact on physical fitness [19,20]. On the other hand, the attractiveness of practicing an exergame might favor the repetition of the tasks to be performed when exergaming, thus learning these tasks, as authors have long considered repetition to be a key element of learning [21,22].

Most studies examined the possible influence of exergaming on motor learning outside the school context, and especially fundamental motor skills learning. These encompass “balance/stability skills” (eg, 1-foot balance), “object control skills” (eg, throwing), and “locomotor skills” (eg, jumping), rooted in physical fitness (eg, coordination), and form a basis for learning and performing different types of physical activity [23-25]. Mastery of these motor skills is thus likely to encourage long-term commitment to physical activity, with all its attendant health benefits [26,27]. Interestingly, reviews and meta-analyses

highlighted that exergaming may develop balance and postural stability [23,28]. A meta-analysis also found a positive influence of exergames involving actions, such as catching or kicking, on performance in a test assessing object control skills [25]. On the other hand, although a few experimental results suggested the possible influence of exergaming on locomotor skills in children and adolescents [29,30], reviews and meta-analyses found that such findings require confirmation [23]. Especially, whether exergaming might develop general motor coordination (ie, the ability to coordinate several limbs and the whole body [31] as a basis for locomotor skills [23]) remains, to our knowledge, to be clarified.

Furthermore, the influence of exergaming on school achievement has been little studied to date, although studies emphasized that active learning, which links movement and learning, may be a promising way to reduce sedentary behavior while improving academic performance [32,33]. Two recent studies reported a positive influence of practicing exergames with math content on motivation and learning in the areas involved by the exergames, which is in line with the literature on serious games [6]. One of these studies found a positive influence on motivation in math, but not on math achievement, and on physical literacy level [34]. The other study found that exergaming, based on ball throws to hit points after identifying them in a Cartesian coordinate system, increased the performance of a paper-and-pencil task of identifying points in such a coordinate system, as well as the accuracy of ball throws at a target [35]. A 2013 study also reported the influence of practicing a nonmathematical exergame, based on aerobic dance, not only on children’s performance in a 1-mile run, but also on math scores on the Utah Criterion-Referenced Test [36]. The last result thus suggests that practicing an exergame might positively influence school achievement beyond the areas that this exergame targets, which requires further investigation, as such a possibility would be of great pedagogical interest.

This result might be due to the influence of exergaming on executive functions reported in the literature [13,37], as these functions include working memory, attentional flexibility, inhibitory control, and higher-level functions such as reasoning, planning, and problem-solving [38]. However, the explanation seems insufficient, as meta-analyses found that training a cognitive function such as working memory generally leads to progress only in tasks similar to those used in the training [39,40]. Another explanation might lie in the development, thanks to exergaming, of spatial ability, that is, a component of intellectual ability [41] and a predictor of math achievement [42]. A meta-analysis found a correlation between video game skill and spatial ability, but no training influence of video game playing on spatial ability [43]. However, another meta-analysis found that spatial ability may be improved by training, with possible transfer effects [44]. Interestingly, studies found that the practice of physical activities in which the processing of spatial information is decisive (eg, wrestling [45] and dance [46]) may improve mental rotation, that is, a component of spatial ability that allows one to rotate the mental image of an

object [47,48]. Other studies showed that mental rotation training may positively influence math achievement [49]. More results suggested the existence of links among mental rotation, general motor coordination, and math achievement, particularly in arithmetic [50-52].

Hypotheses, Aim, and Objectives

The literature led us to consider that dance-based exergaming (DEX), without explicit math content, could influence mental rotation, general motor coordination, and math achievement. Our study, therefore, aimed to clarify this possible influence in adolescent students.

This study differed from previous studies that examined whether exergames based on math problems and ball-throwing may favor math achievement, physical activity, and ball-throwing accuracy [34,35]. Instead, our study focused on DEX as potentially practiced during physical education (PE) lessons, but also during school periods without formal teaching, or as part of the preparation for a school event. This aimed to examine whether such practice may favor school achievement, as suggested by a previous study that showed a positive influence of exergaming based on aerobic dance on the performance of both an endurance run and a math test [36]. Compared with this study, our study had the distinction of focusing on a dance-based exergame selected as being well-suited to train mental rotation and whose practice was organized to scaffold this training.

The main objective of this study was to determine the influence of this DEX practice on mental rotation efficiency, general motor coordination, and math achievement in an experimental group (EG). This was done by comparison with a control group (CG) involved in precision ball-throwing exergaming (TEx), using a pretest-posttest design. Students' intensity levels of physical activity during exergaming and the situational interest [53] that this practice generated in them were also measured to monitor their involvement.

Methods

Participant Characteristics

This study focused on 2 groups, EG and CG, comprising students aged 13-16 years from a regular secondary school based in Switzerland (Canton of Vaud). A total of 56 students (27 girls and 29 boys) participated in the entire study; EG consisted of 30 students (15 girls and 15 boys) with a mean age of 14.0 (SD 0.7) years, and CG consisted of 26 students (12 girls and 14 boys) with a mean age of 14.2 (SD 0.9) years.

Two qualified PE teachers, with experience in teaching (10 years) and familiarity with their school's Lü exergaming system, participated in the study (implementation of the exergaming sequences). One was the regular PE teacher for the 2 classes forming EG, and the other for the 2 classes forming CG.

Inclusion and Exclusion Criteria

We constituted the EG from 2 classes of the school where the experiment took place, and the CG from 2 other classes of this school. On this basis, the inclusion criteria were no known contraindications to physical activity and, for harmonization purposes, no known or presumed special educational needs;

thus, not meeting one or more of these criteria led to the a priori exclusion from the study.

Sampling Procedures

In the enrollment stage (January 2023), 72 students were eligible to participate in the study, 37 as potential EG individuals, and 35 as potential CG individuals. These numbers were likely to decrease depending on consent to participate and then, due to a series of requirements relating to the different phases of the experiment (from February to April 2023): pretests (presence), exergaming (participation in the 5 planned exergaming sessions), and posttests (presence). Among the individuals meeting these criteria, compliance (vs noncompliance) with instructions during testing was the final selection criterion.

Sample Size, Power, and Precision

Based on preliminary data, a sample size calculation, using the *pwr* package in R software (version 4.5.2; R Foundation for Statistical Computing), indicated that 21 participants per group would be required to achieve 90% power ($\alpha=.05$, effect size of 0.5) to detect differences in the primary outcomes of the study. Consequently, EG (n=30) and CG (n=26) were deemed acceptable for the study. However, it should be noted that since power calculations were performed for the primary outcomes, other outcomes may have been underpowered.

Ethical Considerations

The study was conducted in accordance with the Helsinki Declaration of 1975, as revised in 2000. In accordance with local procedure, the project for this pilot interventional study was submitted to the research coordination committee of the host university (University of Teacher Education, Lausanne, Canton of Vaud, Switzerland). This committee approved the project (RCC-pro ID 767) without requiring further submission to the Cantonal Commission for Ethics in Human Research. Subsequently, this project has been integrated into the university's research project repository.

We also obtained official authorization to implement the study in schools located in the Canton of Vaud and selected 1 secondary school for the study due to its exergaming facilities. We presented the project to the school's headteacher, who agreed to its implementation. The teachers involved in the experiment, the parents of the students, and the students completed written informed consent forms gathered prior to data collection, which gave them the option to withdraw their consent at any time. No specific compensation was provided for participation in the research. The data were anonymized and stored on a secure university computer. In addition, we have verified that no images that could identify the study participants were used in the study (including in the manuscript and multimedia appendices).

Design and Conditions

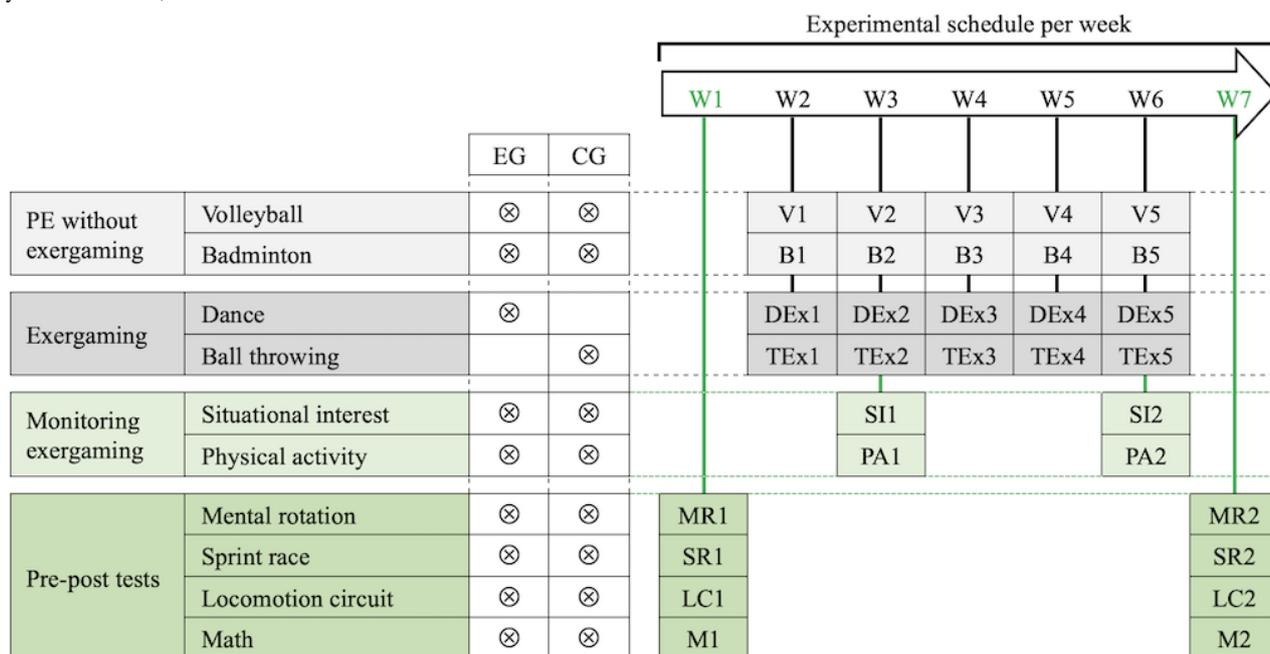
Experimental Design

The sequences of DEX in EG and TEx in CG were implemented for 5 weeks during the regular PE timetable of the classes involved in the experiment. The participants attended 2 weekly PE lessons: one 90-minute lesson ("2-period lesson") and one 45-minute lesson ("1-period lesson"). Only the first period of

each 2-period lesson differed between EG and CG, with DEx sessions in EG and TEx sessions in CG; each of these weekly exergaming sessions lasted 45 minutes. No student declared previous experience of the exergames practiced during these

sequences. Both EG and CG practiced badminton during the second period of each 2-period lesson and volleyball during each 1-period lesson (Figure 1).

Figure 1. Experimental design to study the influence of a dance-based exergaming sequence, compared with a precision ball-throwing exergaming sequence, on mental rotation, general motor coordination, and math achievement in adolescent students. B1-B5: badminton lessons 1-5; CG: control group; DEx1-DEx5: dance-based exergaming sessions 1-5; EG: experimental group; LC: locomotor circuit, performed during pretests 1 and posttests 2; M1: math tests performed during pretests; M2: math tests performed during posttests; MR1: mental rotation test performed during pretests; MR2: mental rotation test performed during posttests; PA: physical activity monitoring; PE: physical education; SI: situational interest assessments; SR1: sprint race performed during pretests; SR2: sprint race performed during posttest; TEx1-TEx5: precision ball-throwing exergaming sessions 1-5; V1-V5: volleyball lessons 1-5; W1-W7: week 1-7.



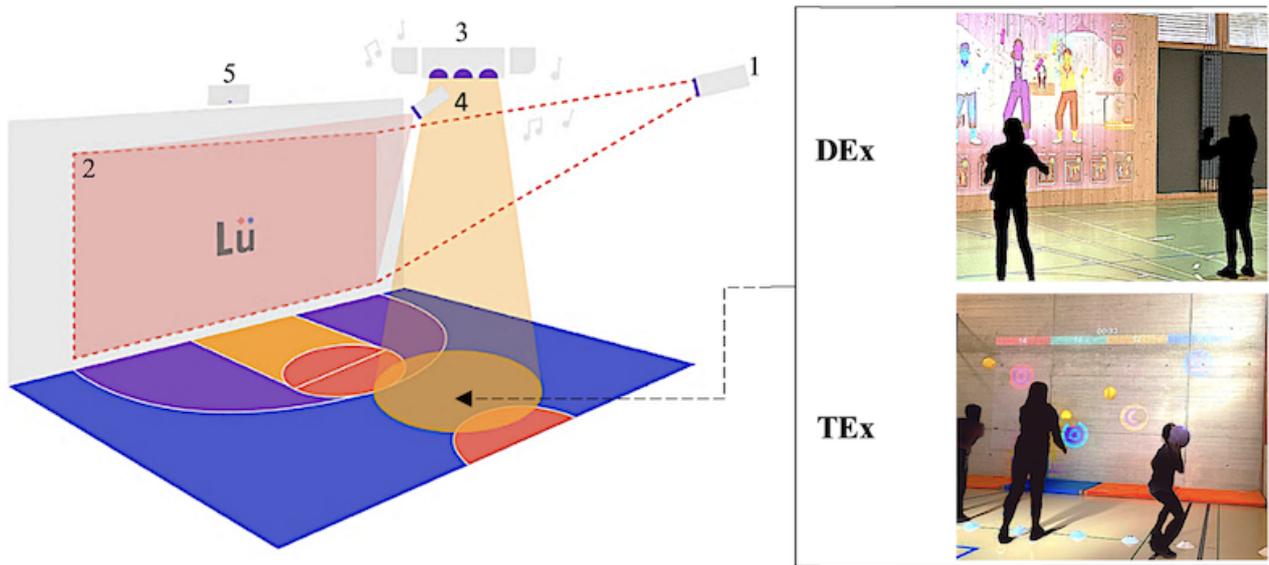
The same experimenters observed each session of DEx and TEx to check whether the teacher had implemented it as planned. Furthermore, the intensity levels of the EG and CG participants' physical activity were measured to check whether these were in line with the respective exergaming demands (Figure 1). This also allowed an estimation of the exergaming contribution to the moderate to vigorous physical activity (MVPA) recommended daily by the World Health Organization [54]. Measurements were done during sessions 3 and 5. The experimenters monitored physical activity, as in previous studies [55], during the sessions, at the end of which situational interest was assessed to control for possible differences between EG and CG likely to have influenced activity during exergaming (Figure 1). These sessions were chosen to avoid novelty bias (session 3) and allow time for possible evolution of situational interest (session 5, ie, the last of DEx and TEx), as in previous studies [35]. Instrumented monitoring of physical activity and situational interest was carried out with each EG and CG participant.

The students completed pretests during the week preceding the beginning of DEx and TEx and posttests during the week after DEx and TEx. During both pretests and posttests, EG and CG performed a mental rotation test, a sprint race, and a locomotion circuit used to assess general motor coordination, and math tests.

Exergaming System

The teachers participating in the study implemented DEx and TEx using a Lü Üno system, that is, an audiovisual infrastructure devoted to exergames practice and coming with gaming applications. In this system, the so-called "heart" module (comprising a PC with a webcam activated by remote control) is connected to an image projection module, a light and sound module, and a camera module for motion detection, for example, to detect the contact of a ball with the wall serving as a projection screen (Figure 2).

Figure 2. A diagram of the Lü Üno system used to study the influence of dance-based exergaming (DEx), compared with precision ball-throwing exergaming (TEEx), on mental rotation, general motor coordination, and math achievement in adolescent students. This system consists of a single-wall configuration with (1) high-definition laser projection (5200 lumens), (2) image (6 m × 3 m), (3) multicolored LED lighting system with moving lights and sound system with dual speakers (2400 Watts), (4) movement detection camera allowing interaction with the projected images, and (5) internal computer with a webcam. The photographs show examples of practice during the experiment with Gröëve and Target.



DEx in EG

During DEx, EG practiced the exergame Gröëve (Figure 2) involving no overt math content. Gröëve leads to the practice of various dancing steps that drawings indicate, and avatars

demonstrate. Gröëve also allows one to progressively learn an entire dance using different practice modalities (Figure 3). The choice of this application was in line with the local PE program (“CM32: Develop one’s coordination skills and use one’s own body as a means of expression and communication”).

Figure 3. Using the exergame Gröëve to implement a dance-based sequence and study its influence on mental rotation, general motor coordination, and math achievement in adolescent students. The screenshot shows an example of movements included in the dance learned during the experiment that were visualizable on a screen. The available options to adapt the practice difficulty to the participants are also shown; these allowed one to control the movements’ speed (x1.00, x0.75, or x0.50) and/or the sound played (metronome only, metronome and music, or music only). The application also allows one to pause the video (“Pause” button) and to save the progress of the participants (“Mastered” button).



EG practiced Gröëve with the explicit aim of learning a dance based on a song including 4 verses and a chorus. The EG participants had to learn the dance by observing and imitating

avatars’ movements (Figures 2 and 3). Sessions 1, 2, and 3 focused respectively on the first and second verses, the refrain, and the third and fourth verses, while sessions 4 and 5 were

devoted to reviewing the entire dance ([Multimedia Appendix 1](#)). During sessions 1-3, the teacher organized movement learning using part practice [56] and progressively combining the learned parts; during sessions 4-5, the participants reviewed the dance part by part and then entirely, leading to whole practice [56]. Consequently, the content of the sessions varied from one session to another, while ensuring continuity and progressivity in the learning. The teacher also used available options on Grööve ([Figure 3](#)) to allow progressive learning from lessons 1-3, as well as repetition under facilitating conditions and in final conditions during lessons 4 and 5, respectively ([Multimedia Appendix 1](#)).

Each part of the dance was largely composed of asymmetric movements that avatars showed facing the EG participants; however, the avatars were also turning, so that the EG participants were seeing them in profile ([Multimedia Appendix 1](#)). Thus, mental rotation was useful to observe and imitate the avatars' movements. However, it was possible to imitate correctly part of these movements without using mental rotation. For example, when the participants viewed an avatar from the front, arms uncrossed, they could imitate an arm elevation seen on their left by deciding, without using mental rotation, to elevate their own right arm.

Thus, to maximize the use of mental rotation, the teacher had to instruct the EG participants to imagine themselves dancing in the position of one of the avatars facing them, that is, to rotate their own body mentally to get into this position. The teacher had to give this instruction before each dancing phase, to reiterate it during learning, and when providing feedback. Each session, the teacher also had to film the dance movements of the EG participants, with the Lü camera used facing them ([Figure 2](#)). At the end of the session, the teacher had to use the videos for formative evaluation and to train the participants to analyze freeze-frames with the same instruction as when imitating an avatar (eg, to determine which arm of the viewed participant was outstretched).

TEx in CG

During TEx, exergaming in CG was based on precision ball-throwing ([Figure 2](#)), which was compliant with the local PE program ("CM32: Coordination training [...] using balls"). TEx involved no overt math content and no relevant use of mental rotation.

The teacher implemented TEx using 8 Lü applications, each leading to throwing a ball at a target, in addition to catching it and performing simple walk or run locomotion ([Multimedia Appendix 2](#)). These applications were (1) Target (throw a ball at a circular target, with the target location changing automatically as the game progresses), (2) Galactic (throw a ball at asteroids to protect a spaceship), (3) Relé (relay race with a run to reach a floor line, a ball-throw at the team zone on the screen, then a ball catch to throw it to a partner), (4) Spörts (perform basketball shots at a virtual basket), (5) Germ (throw a ball to create obstacles preventing a virus to reach own zone and allowing to direct this virus to the adverse zone), (6) Vikä (throw a ball at targets to answer questions, then propel a boat toward a village, and protect villagers from a monster), (7) Lüvia (puzzle solving: throw a ball at the correct answer displayed on

the screen among other possible answers), and (8) Brüşh (throw a ball at candy before it hits an avatar's teeth, while avoiding the floss heading toward the avatar).

Throughout TEx, the common thread in terms of learning consisted of practicing precision ball-throwing. On this basis, the experimenters and the teacher involved in TEx designed the sessions to ensure a mix of continuity, variety, and progressivity in practice. Each session involved the use of 3 of the applications described above in the following chronological order: Target, Galactic, and Relé during session 1, Target, Galactic, and Spörts during session 2, Germ, Vikä, and Relé during session 3, Lüvia, Brüşh, and Spörts during session 4, and Lüvia, Spörts, and Vikä during session 5. In addition to this mix of novelty and reuse in applications from one session to another, the teacher initially had to adjust the task requirements to suit the participants' capacities, then gradually increase these requirements, using the options available for each application. For example, the participants had to play Target with targets shown in the same screen column and then scattered across the entire screen ([Multimedia Appendix 2](#)).

Volleyball and Badminton Sequences

During the experiment, EG and CG separately took part in the same weekly 45-minute badminton and volleyball lessons ([Figure 1](#)). Each lesson began with a conventional warm-up (running, stretching, and muscle strengthening), followed by tasks focused on a sport-specific technique, and then regular matches. The content of the lessons was in line with the local PE program ("CM34: Master basic sports techniques, adapt to the game situation").

Measures and Covariates, Data Collection, Quality of Measurements

Before the experiment, each sequence was tested with other students than those forming the EG and CG. This enabled the teachers involved in the study to perfect their use of the Lü Üno system, as DEx and TEx included elements (content, practice options) that they needed to explore further. This also led to adjusting the sequences for subsequent implementation.

The experimenters conducted a noninstrumented observation of each of the 5 DEx and TEx sessions implemented ([Figure 1](#)) to enable comparison between the observed session and the planned session regarding content, teacher intervention, and participant activity.

For confirmation purposes, during sessions 3 and 5 of DEx and TEx, the experimenters measured the intensity levels of physical activity achieved by each EG and CG participant ([Figure 1](#)). The examined variables were (1) the moderate and vigorous intensity levels, expressed in seconds and percentage of time activity, and (2) MVPA, expressed similarly to the moderate and vigorous intensity levels. Given the exergaming characteristics in DEx and TEx described above, we expected EG to be more involved than CG in vigorous-intensity activity, and CG more involved than EG in moderate-intensity activity, as the musical tempo in DEx was to be increased during session 3 and remain at peak in session 5, whereas no such activity peaks were planned in TEx.

Directly after sessions 3 and 5, the experimenters assessed the situational interest of each EG and CG participant to control for possible between-group differences likely to have influenced participants' activity (Figure 1). The measurement of interest, as conceived in the field of educational psychology [57], is well-suited to examine the closeness and sustainability of a person's relationship to learning depending on the content to be learned [58]. The triggered situational interest in a person depends on the present situation and personal characteristics [53]. This may lead to maintained situational interest, eventually allowing the gradual stabilization of interest. Interestingly, studies also showed that situational interest may influence physical activity intensity levels and energy expenditure during physical activity [59].

During both pretests and posttests (Figure 1), the EG and CG participants performed a mental rotation test and math tests at the beginning of 2 separate math lessons. On the days these tests were administered, no specific physical activity was performed beforehand (PE classes or otherwise) to promote harmonization of testing conditions. The participants performed a sprint race during the 1-period PE lesson (45 minutes) of the week, and a locomotion circuit during the 2-period PE lesson (90 minutes).

The experimenters also asked science and math teachers in classes of the EG and CG participants not to provide any specific training on mental rotation and the mathematical content covered in the pretests and posttests during the experiment.

Instrumentation

Monitoring Intensity Levels of Physical Activity and Situational Interest

During sessions 3 and 5 of DEx and TEx, the experimenters measured the intensity levels of physical activity (moderate, vigorous, and MVPA) achieved by each EG and CG participant using hip-worn GT3X+ accelerometers (Actigraph LLC), providing a receivable estimation of physical activity in free-living conditions [60]. The accelerometer data were processed using ActiLife6 (Actigraph LLC). This led to converting the data into activity counts that were summed and recorded using a specific time interval known as an "epoch" (selectable duration from 5 seconds to 1 minute). The selected epoch was 5 seconds, as the study was focused on acute and intermittent exercise bouts.

In this study, situational interest was assessed using the French situational interest questionnaire [61] that contains three factors: (1) triggered situational interest (participant's perception of the game demanding full attention), (2) maintained situational interest feeling (to what extent the activities were found enjoyable), and (3) maintained situational interest value (importance allocated to success during the game). To complete the questionnaire, the participants had to rate statements on a 5-point Likert scale (Multimedia Appendix 3 [53,57,58,61]); for each factor, this led to scores ranging from 1 (lowest possible score) to 5 (highest possible score).

Pretests and Posttests

The EG and CG participants completed the French version [62] of Vandenberg and Kuse's mental rotations test [63] based on a comparison of images of geometrical structures made of 10 cubes (Multimedia Appendix 4 [52,62-70]), the scores ranging from 0 to 40 points, with 40 corresponding to a perfect performance.

These participants also performed a locomotion circuit called the Harre circuit test in the literature [64] to assess general motor coordination [65], with reaction to an audible signal, a forward roll, running phases with direction changes, and phases with jumping over a bench, then crawling under it (Multimedia Appendix 4). In line with previous studies [52,66], a sprint race (20 meters) was performed separately to control for the influence of sprinting capacities on the circuit performance. The performance of the locomotion circuit and of the sprint race was timed using the same timing system (Race Link Racing System; Brower Timing Systems).

Moreover, the participants performed 2 separate paper-and-pencil series of quantity comparisons, based on a previous study [67] (Multimedia Appendix 4), that is, a series of comparisons between 2 sets of points and a series of comparisons between 2 single-digit numbers. The participants also performed 4 separate paper-and-pencil series of mental calculations, based on a previous study [68] (Multimedia Appendix 4), that is, a series of simple additions (2 single-digit numbers), complex additions (three 2-digit numbers), simple multiplication (2 single-digit numbers), and complex multiplication (a 2-digit number and a single-digit number). For each task series listed above, we considered the total number of processed items, the number of correct results, and the error rate.

Masking

The EG and CG students, as well as the teachers involved in the study, were informed of the main lines of the study but were left unaware of any expected results. DEx and TEx sessions have been designed separately, each by a group including the researchers and the PE teacher who would teach the sequence (either to EG or to CG).

Data Diagnostic, Psychometrics, and Analytic Strategy

Statistical analyses were performed without missing data, due to sampling procedures (see "Sampling Procedures" subsection above) using IBM SPSS Statistics (version 30.0.0.0; 171). On this basis, the normality of the numerical variables was preliminarily checked, and the internal reliability of the situational interest scales was evaluated using Cronbach α coefficients.

Multivariate analysis of variance (MANOVA) tests were performed to compare the intensity levels of physical activity in EG and CG during sessions 3 and 5 of DEx and TEx. The scores for each intensity level considered were specified as dependent variables and the group condition (EG vs CG) as a fixed factor. Similarly, MANOVA tests were done to examine possible differences between EG and CG in scores for each of

the 3 situational interest factors measured after sessions 3 and 5.

For each variable measured at pretest and posttest, analysis of covariance (ANCOVA) tests were performed using a generalized linear model approach [71]. The model included basically the posttest score as the dependent variable, the group (EG vs CG) as a fixed factor, and the pretest score as a covariate; in the event of a between-group difference regarding a situational interest variable, it was planned to add the variable as a covariate. Beyond data normality, variance homogeneity, heteroskedasticity, and regression homogeneity were checked; on such a basis, ANCOVA was computed either on untransformed data or on ranks, using the same method for the computations [71,72].

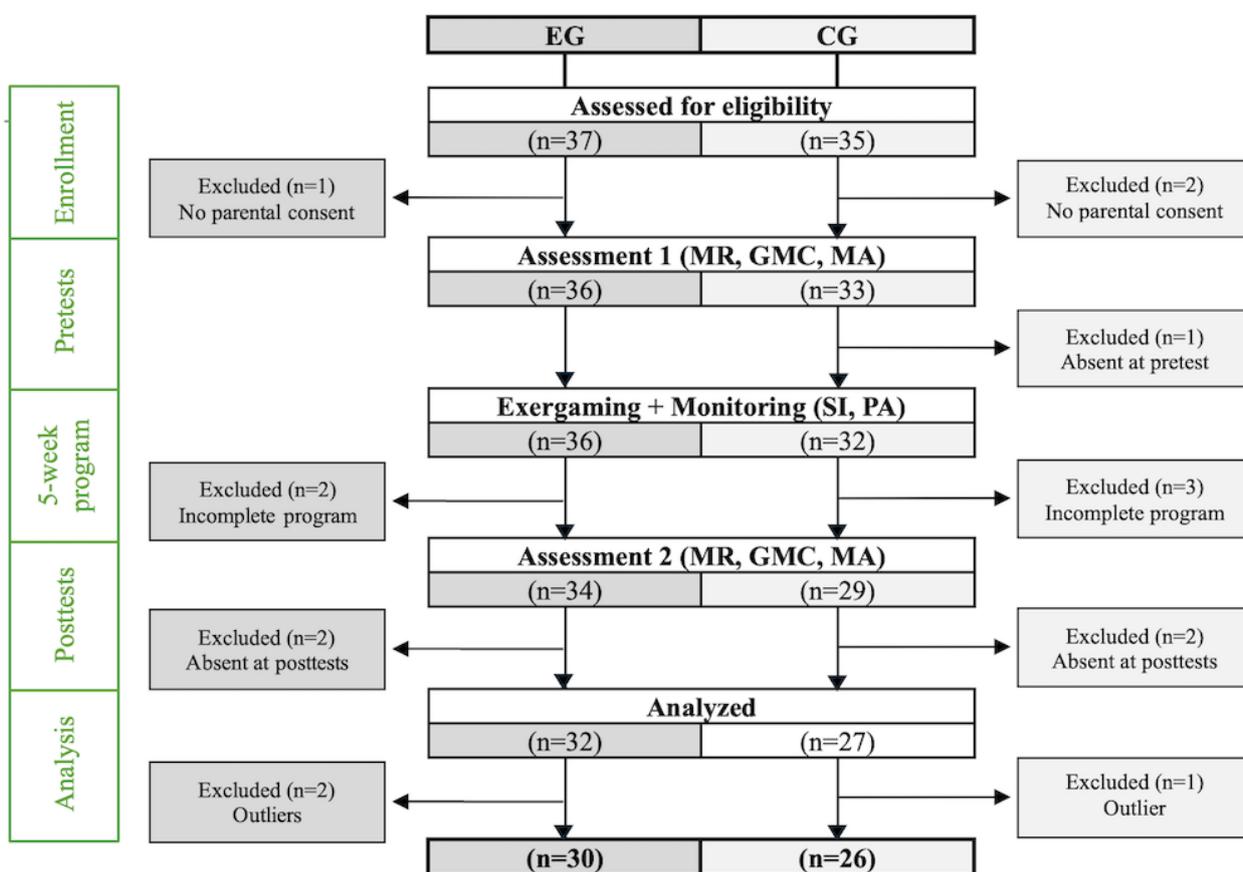
For all the analyses, partial eta squared (η_p^2) provided an index of effect size characterizing the effect as small, medium, or large, depending on whether the value of η_p^2 was between 0.01 and 0.06, between 0.07 and 0.14, or >0.14, respectively [73]. Contrast analyses were applied in the case of significant results.

Results

Recruitment and Participant Flow

The different stages involved in forming the groups, which resulted in the selection of 30 students (15 girls and 15 boys) for EG (mean age 14.0, SD 0.7 years) and 26 students (12 girls and 14 boys) for CG (mean age 14.2, SD 0.9 years), are presented in Figure 4.

Figure 4. Study flowchart. Steps to form experimental group (EG; dance-based exergaming) and control group (CG; precision ball-throwing-based exergaming) to study the influence of dance-based exergaming, compared to precision ball-throwing exergaming, on mental rotation (MR), general motor coordination (GMC), and math achievement (MA) in adolescent students: enrollment stage (January 2023), pretest, 5-week program, and posttests (from February to April 2023). Three participants, among those having participated in all pretests, posttests, and exergaming sessions, were excluded due to failure to follow instructions during testing (random responses). PA: physical activity intensity levels; SI: situational interest.



Statistics and Data Analysis

Intensity Levels of Physical Activity and Situational Interest

The observation of each DEx and TEx session led to the conclusion that the sessions were taught as planned and that the students completed the exergaming tasks. Such observation was

supplemented by the results of analyses of physical activity intensity levels and situational interest during sessions 3 and 5.

The data relating to physical activity intensity levels were normally distributed (skewness values ranging from -0.64 to 0.79, SE 0.32; kurtosis values ranging from -0.58 to 2.56, SE 0.63). For each MANOVA (Table 1), Pillai trace values ranged from $F_{3,52}=10.44$ to $F_{3,52}=63.63$ ($P<.001$).

Table 1. Results of multivariate analysis of variance (MANOVA) tests for physical activity intensity levels. Physical activity intensity levels were monitored during sessions 3 and 5 of 2 exergaming sequences (5 weekly sessions of 45 minutes each) as part of a study of dance-based exergaming (DEX) influence on mental rotation, general motor coordination, and math achievement in adolescent students. The influence of DEX (experimental group [EG]) was compared to that of precision ball-throwing exergaming (control group [CG]).

	EG (n=30), mean (SD)	CG (n=26), mean (SD)	MANOVA tests			Contrasts
			<i>F</i> test (<i>df</i>)	<i>P</i> value	η_p^2	
Session 3 (seconds)						
Moderate	739.7 (122.7)	1060.8 (154.4)	75.10 (1, 54)	<.001	0.58	EG < CG
Vigorous	336.7 (82.2)	283.6 (139.4)	3.10 (1, 54)	.08	0.05	N/A ^b
MVPA ^a	1180.7 (173.5)	1419.2 (290.3)	14.16 (1, 54)	<.001	0.21	EG < CG
Session 3 (%^c)						
Moderate	35.9 (5.3)	42.0 (5.5)	17.75 (1, 54)	<.001	0.25	EG < CG
Vigorous	16.4 (4.0)	12.5 (5.9)	8.46 (1, 54)	.005	0.13	EG > CG
MVPA	57.4 (8.0)	56.1 (3.3)	0.30 (1, 54)	.59	0.005	N/A
Session 5 (seconds)						
Moderate	724.7 (102.5)	1035.0 (164.3)	73.94 (1, 54)	<.001	0.58	EG < CG
Vigorous	322.7 (63.3)	221.0 (115.3)	17.34 (1, 54)	<.001	0.24	EG > CG
MVPA	1189.3 (154.1)	1311.9 (291.7)	4.01 (1, 54)	.05	0.07	N/A
Session 5 (%^d)						
Moderate	32.1 (4.4)	40.9 (5.6)	42.01 (1, 54)	<.001	0.44	EG < CG
Vigorous	14.3 (2.9)	10.0 (5.5)	14.72 (1, 54)	<.001	0.21	EG > CG
MVPA	52.8 (6.9)	51.7 (10.3)	0.17 (1, 54)	.68	0.003	N/A

^aMVPA: moderate to vigorous physical activity.

^bN/A: not applicable.

^cIntensity levels of physical activity in percentage of the total physical activity time during session 3.

^dIntensity levels of physical activity in percentage of the total physical activity time during session 5.

On average, during session 3, the EG participants spent 1180.7 (SD 173.5) seconds in MVPA, that is, 57.4% (35420/61770 seconds) of total physical activity time, and the CG participants, 1419.2 (SD 290.3) seconds, that is, 56.1% (36900/64745 seconds) of total physical activity time. During session 5, the values were 1189.3 (SD 154.1) seconds and 52.8% (34110/64765 seconds) in EG, 1311.9 (SD 291.7) seconds and 51.7% (35680/67680 seconds) in CG (Table 1).

CG had more MVPA time than EG in sessions 3 ($P<.001$; large effect size) and 5 ($P=.05$; small effect size), but no between-group difference was found when considering MVPA as a percentage of total physical activity time in sessions 3 ($P=.59$) and 5 ($P=.68$). As expected, MVPA relied on (1) significantly higher moderate-intensity values in CG than in EG, in seconds and percentage of total physical activity time during sessions 3 and 5 ($P<.001$; large effect sizes), and (2) higher vigorous-intensity values in EG than in CG, in seconds

and percentage of total physical activity time during sessions 3 and 5, with significant differences ($P<.001$; medium to large effect sizes), except for values in seconds during session 3 ($P=.08$).

On the other hand, data on situational interest in sessions 3 and 5 were normally distributed (skewness values from -0.50 to 0.47 , SE 0.32 ; kurtosis values from -1.19 to -0.83 , SE 0.63). In addition, the Cronbach α coefficients were good for each situational interest factor measured after sessions 3 (from 0.89 to 0.92) and 5 (from 0.90 to 0.92). Results of MANOVA tests reached statistical significance for session 5 (Pillai trace with $F_{3,52}=3.93$; $P=.01$), but not for session 3 ($F_{3,52}=1.09$; $P=.36$), due to a higher score for triggered situational interest in EG than in CG ($P=.004$; large effect size) in session 5 (Table 2). Consequently, triggered situational interest during session 5 was added as a covariate to perform ANCOVA tests in further analyses.

Table 2. Results of multivariate analysis of variance (MANOVA) tests for situational interest. Situational interest factors were assessed directly after sessions 3 and 5 of 2 exergaming sequences (5 weekly 45-minute sessions each) as part of a study of dance-based exergaming influence on mental rotation, general motor coordination, and math achievement in adolescent students. The influence of dance-based exergaming (experimental group [EG]) was compared to that of precision ball-throwing-based exergaming (control group [CG]).

	EG (n=30), mean (SD)	CG (n=26), mean (SD)	MANOVA tests			Contrasts
			<i>F</i> test (<i>df</i>)	<i>P</i> value	η_p^2	
Situational interest in session 3						
TSI ^a	3.52 (1.39)	2.89 (1.21)	3.23 (1,54)	.08	0.06	N/A ^b
MSIF ^c	2.63 (1.17)	2.21 (1.05)	1.98 (1,54)	.16	0.03	N/A
MSIV ^d	2.59 (1.13)	2.16 (1.10)	2.04 (1,54)	.16	0.04	N/A
Situational interest in session 5						
TSI	3.49 (1.23)	2.47 (1.29)	9.15 (1,54)	.004	0.14	EG > CG
MSIF	2.34 (1.05)	2.02 (1.12)	1.23 (1,54)	.27	0.02	N/A
MSIV	2.42 (1.10)	2.08 (1.14)	1.34 (1,54)	.25	0.02	N/A

^aTSI: triggered situational interest.

^bN/A: not applicable.

^cMSIF: maintained situational interest feeling.

^dMSIV: maintained situational interest value.

For all MANOVA results, the estimated marginal means and corresponding 95% CIs are given in [Multimedia Appendix 5](#).

Pretests, Posttests, and Exergaming Influence

ANCOVAs were performed on ranks for the total number and the number of correct comparisons of dot ensembles performed (heteroskedasticity test with $F_{1,54}=4.88$; $P=.03$ and $F_{1,54}=4.18$; $P=.046$), all data series for number comparison (kurtosis values from 2.60 to 4.90, SE 0.63), the total number of calculations in simple addition (Levene's test with $F_{1,54}=7.94$; $P<.001$), error rate in simple (kurtosis values of 3.02 and 14.82, SE 0.63) and complex (Levene's test with $F_{1,54}=18.72$; $P<.001$) additions,

and in simple multiplication (kurtosis values of 3.26 and 15.50, SE 0.63). Untransformed data were used in any other case. No significant regression homogeneity issue was found (results ranging from $F_{1,50}=0.003$; $P=.95$ to $F_{1,50}=3.05$; $P=.08$).

The results showed that EG had a statistically significant advantage (medium to large effect sizes) over CG regarding the Vandenberg and Kuse mental rotations test scores ($P=.02$), the number of correct results of simple additions ($P=.006$), and the error rate in complex additions ($P=.005$). No other significant differences were found, either in terms of additions or in measurements relating to the sprint race, the locomotion circuit, quantity comparisons, or multiplications ([Table 3](#)).

Table 3. Results of analysis of covariance (ANCOVA) tests. Effects of dance-based exergaming on test scores of mental rotation, general motor coordination, and math in adolescent students. Comparison of test scores obtained after a dance-based exergaming sequence of five 45-minute weekly sessions (experimental group [EG]) with those obtained after a similar sequence based on precision ball-throwing (control group [CG]), controlling for the scores before the first exergaming sessions and the triggered situational interest levels during exergaming.

	Pretests, mean (SD)		Posttests, mean (SD)		ANCOVA tests			Contrast
	EG (n=30)	CG (n=26)	EG (n=30)	CG (n=26)	F test (df)	P value	η_p^2	
VMRT^a								
Score	18.40 (7.98)	13.77 (7.51)	22.23 (8.95)	13.88 (8.35)	6.17 (1, 52)	.02	0.11	EG > CG
GMC^b								
Sprint	3.60 (.41)	3.62 (.34)	3.65 (.30)	3.63 (.38)	0.18 (1, 52)	.68	0.003	N/A ^c
LC ^d	22.41 (4.15)	22.42 (3.44)	21.72 (4.39)	21.20 (3.41)	0.86 (1, 52)	.36	0.02	N/A
DC^e								
N ^f (r ^g)	59.43 (9.03)	60.81 (7.27)	64.63 (8.76)	67.23 (6.22)	0.66 (1, 52)	.42	0.01	N/A
C ^h (r)	55.20 (5.23)	56.85 (5.26)	58.73 (8.14)	61.27 (4.57)	0.01(1, 52)	.93	0.0001	N/A
ER ⁱ	6.23 (7.03)	6.15 (4.41)	8.92 (7.36)	8.61 (4.24)	0.08 (1, 52)	.77	0.002	N/A
NC^j								
N (r)	65.53 (7.08)	66.15 (6.42)	68.96 (6.10)	68.81 (5.51)	0.13 (1, 52)	.71	0.003	N/A
C (r)	64.97 (6.90)	65.08 (6.52)	67.83 (5.80)	67.92 (5.51)	2.11 (1, 52)	.15	0.04	N/A
ER (r)	.83 (1.24)	1.63 (2.21)	1.65 (1.32)	1.29 (1.20)	2.17 (1, 52)	.15	0.04	N/A
SA^k								
N (r)	39.47 (11.08)	42.92 (12.59)	47.30 (13.65)	45.27 (11.22)	3.11 (1, 52)	.08	0.06	N/A
C	38.97 (11.21)	41.35 (12.19)	46.47 (13.33)	43.46 (12.21)	8.26 (1, 52)	.006	0.14	EG > CG
ER (r)	1.43 (2.99)	3.53 (3.88)	1.72 (2.46)	4.63 (7.51)	1.85 (1, 52)	.18	0.03	N/A
CA^l								
N	7.90 (2.19)	7.46 (2.55)	9.07 (2.66)	8.65 (3.21)	0.09 (1, 52)	.76	0.002	N/A
C	6.73 (2.11)	5.42 (2.84)	8.07 (2.46)	6.35 (3.26)	1.45 (1, 52)	.23	0.03	N/A
ER (r)	14.57 (16.24)	30.32 (25.50)	10.70 (10.59)	27.08 (22.93)	8.40 (1, 52)	.005	0.14	EG < CG
SM^m								
N	28.60 (9.47)	33.04 (11.00)	31.63 (9.30)	35.27 (10.41)	0.04 (1, 52)	.85	0.001	N/A
C	27.70 (9.81)	31.31 (10.81)	30.47 (9.10)	33.19 (11.11)	0.01 (1, 52)	.92	0.0002	N/A
ER (r)	3.76 (5.98)	5.51 (6.07)	3.69 (4.75)	6.46 (11.15)	0.01 (1, 52)	.93	0.0001	N/A
CMⁿ								
N	13.43 (2.61)	13.42 (3.05)	14.27 (3.05)	13.92 (2.53)	0.01 (1, 52)	.91	0.0002	N/A
C	12.30 (2.58)	11.85 (2.92)	12.80 (2.71)	12.31 (2.82)	0.21 (1, 52)	.64	0.004	N/A
ER	8.54 (8.71)	11.25 (10.54)	9.65 (9.64)	11.86(9.64)	0.04 (1, 52)	.84	0.001	N/A

^aVMRT: Vandenberg and Kuse mental rotations test.

^bGMC: general motor coordination (all measurements in seconds).

^cN/A: not applicable.

^dLC: locomotion circuit (Harre circuit test).

^eDC: comparison of dot ensembles.

^fN: total number of items performed.

^gr: ANCOVA computed on ranks.

^hC: number of correct responses.

ⁱER: error rate.

^jNC: number comparison.

^kSA: simple addition.

^lCA: complex addition.

^mSM: simple multiplication.

ⁿCM: complex multiplication.

For all ANCOVA results, the estimated marginal means and corresponding 95% CIs are given in [Multimedia Appendix 5](#).

Discussion

Principal Findings

This pilot interventional study aimed to verify whether DEX, involving mental rotation to perform different types of locomotion and interlimb coordination, may influence (1) mental rotation efficiency, (2) general motor coordination, and (3) math achievement (quantity comparisons, simple and complex additions, and multiplications) in an EG of adolescent students ([Multimedia Appendix 6](#)). This was done using a pretest-posttest design, including a CG involved in TEx and monitoring physical activity intensity levels and situational interest.

Noninstrumented observation suggested that the sessions proceeded as planned. This was corroborated by physical activity monitoring ([Table 1](#)), which also showed that MVPA timed values approached one-third of the daily 3600 seconds recommended for beneficial effects on adolescent health [54], that is, from 1137 (31.6%) seconds to 1372 (38.1%) seconds. The monitored amounts of moderate and vigorous physical activity were consistent with the expectations based on the planned sequences, which suggested that situational interest did not impede the physical activity required by DEX and TEx [59]. Only the scores for the triggered situational interest factor in session 5 significantly differed between groups, with a higher score in EG than in CG (large effect size), possibly due to the choice of the DEX and TEx activities and how they were perceived by the participants [58] ([Table 2](#)). ANCOVAs were performed controlling for this variable, considering the possible influence of situational interest on energy expenditure [59], thus exergaming task repetition and related learning [22].

The main ANCOVA results showed an advantage for EG after DEX over CG after TEx on mental rotation efficiency, the number of correct simple additions performed, and the error rate in complex addition (medium to large effect sizes). No other statistically significant ANCOVA results were found.

Interpretation

Exergaming Influence on Mental Rotation Test Scores

Previous meta-analysis found that spatial ability (including mental rotation) may be improved by training [44], whereas in our study, DEX in EG, but not TEx in CG, was likely to elicit mental rotation training. The advantage for EG over CG regarding mental rotation found, once DEX and TEx were completed, may thus be due to mental rotation training in EG during DEX [44]. This influence of DEX on mental rotation in

adolescents is consistent with the results of previous studies showing the influence on mental rotation of exergaming based on leg movements, but in adult patients with chronic stroke [74], and of certain physical activities (eg, wrestling vs endurance running) in healthy students, but which were practiced outside the exergaming framework [45].

Furthermore, in this study, DEX influence on mental rotation was found using a test [63] with tasks and stimuli (geometrical structures to be compared to a reference) different from mental rotation training tasks and stimuli (images of avatars and humans to be related to one's own body). This supports that mental rotation training may elicit transfer to tasks and stimuli different from those used for training. Such a result, in line with those of previous studies obtained without physical activity during mental rotation training [75], contributes to fuel a long-standing and still running debate in the literature [76].

Exergaming Influence on General Motor Coordination

This study showed no significant difference regarding general motor coordination between EG and CG, once DEX and TEx were completed ([Table 3](#)). No difference was expected for sprinting, as neither DEX nor TEx trained sprinting capacities. Sprint performance times recorded in EG and CG during pretests and posttests ([Table 3](#)) showed that EG and CG had similar and stable potential to achieve the run parts of the locomotion circuit used to assess general motor coordination [52]. Thus, DEX and TEx showed similar influence on the performance of the other parts of the circuit, and the averaged performance improved by 3% (0.69/22.41 seconds) in EG and by 5% (1.22/22.42 seconds) in CG.

It remains that these variations did not lead to significant between-group differences in terms of ANCOVA results, while movement requirements differed between DEX and TEx. Considering that general motor coordination is a multidimensional construct [77], both DEX and TEx might have developed abilities influential on the performance of the locomotion circuit used for assessment. For example, DEX influence on mental rotation might have impacted this performance [52], but TEx may have developed other abilities with such an impact, particularly those related to the ability to react to a signal involved in moving to catch a ball [78]. On such a basis and considering the existing literature [23], further research remains needed to clarify whether, and under which conditions, exergaming might influence general motor coordination.

Exergaming Influence on Math Achievement

A previous study reported the positive influence of exergaming based on aerobic dance on the *global* score of a test assessing

math achievement [36]. Our results confirmed such an influence of DEx when compared with TEx. However, this influence was found to be limited to peculiar mathematical tasks (Table 3). No significant differences between EG and CG were found regarding the “number tasks” [79] of quantity comparison (ie, with numbers and quantities but without requirements in formal calculation). On the other hand, mixed results were obtained regarding the “calculation tasks” [79] (ie, requiring the application of operation rules to numbers), with (1) no significant between-group differences in multiplications, and (2) a significant advantage for EG over CG regarding the number of simple additions correctly performed and the error rate in complex additions. Calculation of simple additions was thus faster in EG than in CG, without detrimental effect on the error rate, and the error rate in complex additions was lower in EG than in CG, without detrimental effect on calculation speed.

According to the literature, this advantage for EG might be limited to tasks primarily requiring procedural strategies for calculating, rather than strategies for retrieving memorized results [80]. In support of this possibility, additions, even simple (eg, $7+5$), have been found to rely on procedural strategies more than simple multiplications (eg, 7×5), mainly based on retrieval strategies [80]. This explanation appears plausible, though we found no differences between EG and CG in terms of complex multiplications. Such multiplications may involve a mixture of simple multiplication and addition (eg, for adding the retrieved results of 2×7 and of 10×7 , when calculating 12×7), that is, retrieval strategies, combined with calculation procedures [81,82]. Although EG had an advantage over CG in additions (Table 3), this might have elicited a weak impact on efficacy in complex multiplications. Overall, this is also consistent with the advantage in mental rotation efficiency for EG over CG found in our study (Table 3). Indeed, studies showed that mental rotation may improve calculation efficiency by promoting appropriate access to a representation of numbers as a mental number line [83], particularly involved during addition, but not multiplication [84].

Limitations

Although our study helps to fill a gap in the literature concerning the possible influence of exergaming on academic performance, it is not without its limitations. One limitation is linked to the implementation of scientific research in a school environment. Notably, this did not allow us to allocate students randomly between EG and CG. Possible generalization of the results should also be questioned, especially as these results might depend on features of the exergaming sequences implemented, of the school of implementation, and of the students involved in the sequences. Confirmation of our results should also be sought using larger samples. Especially, in this pilot study, no formal correction for multiple comparisons was applied to the ANCOVA tests since applying strict corrections could have increased the risk of type II errors, given the sample size, the number of factors, and related outcomes [85]. Larger samples

would also be needed to clarify the extent to which the selected DEx and TEx activities may modulate our results, by controlling for physical activity (eg, varying the levels of physical activity intensity in 2 or more DEx sequences, as well as in 2 or more TEx sequences), and situational interest [59]. Furthermore, larger samples would allow for proper examination of possible gender differences in the influence of exergaming on spatial abilities and math performance, ideally taking age into account [86,87]. Longer sequences than those used in this study could also be implemented for confirmation purposes [22].

Perspectives

The results of this study encourage us to clarify whether mental rotation training through exergaming may influence math achievement. These results also encourage further investigation to determine if the influence of exergaming on the performance of arithmetical tasks is limited to tasks requiring procedural strategies for calculation. This also raises the question of whether such influence might depend on age, gender, or math experience. Analysis of task repetition and participants' experience when exergaming might be useful in clarifying these points. Moreover, this study found no evidence of an effect of exergaming on general motor coordination, but a slight improvement in general motor coordination scores was observed in both exergaming sequences considered. This encourages further research to clarify the possible influence of exergaming on general motor coordination and its potential links to mental rotation training during exergaming. It would also be of interest to assess the potential added value of specific exergaming features (eg, avatar demonstration or options for managing exercise variety and progression) in achieving the effects observed in our study [88]. Teacher interventions directly linked to such features could complementarily be considered [23]. This assessment might be done by comparison with a conventional physical activity.

Conclusion

The possible benefits of exergaming in schools have been little studied to date, except in relation to promoting physical activity and health [20]. Especially, no studies have yet been conducted to determine whether a given exergaming sequence may positively influence mental rotation, general motor coordination, and math achievement. For the first time, to our knowledge, our study provides support for the possibility that the same DEx sequence without math content may have a positive influence, in the same adolescent students, on performance in a mental rotation test and in tasks of addition. The results of this study encourage complementary investigations, especially to delimit the possible influence of practicing an exergame on math achievement and determine to what extent such influence might depend on mental rotation training, and possibly motor coordination training, during this practice. However, these results suggest that it would be appropriate to integrate exergames into an active learning approach at school, to promote both PE and math outcomes.

Acknowledgments

The authors would like to thank all the students and teachers who took part in the study.

Data Availability

For follow-up purposes, the datasets analyzed during this study are available from the corresponding author upon reasonable request [89].

Funding

This work was supported by the Reserve and Innovation Fund of the University of Teacher Education (HEP-VD), Lausanne, Switzerland (grant 10_FRI_A21_Fargier).

The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Authors' Contributions

PF contributed to conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, and writing (original draft, review, and editing). VC contributed to formal analysis, investigation, methodology, writing (original draft, review, and editing). VL-K and CR contributed to conceptualization, funding acquisition, investigation, and writing (review and editing).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Dance-based exergaming sequence (DEx).

[PDF File (Adobe PDF File), 399 KB - [games_v14i1e82610_app1.pdf](#)]

Multimedia Appendix 2

Exergaming sequence based on precision ball-throwing (TEEx).

[PDF File (Adobe PDF File), 811 KB - [games_v14i1e82610_app2.pdf](#)]

Multimedia Appendix 3

Situational interest monitoring.

[PDF File (Adobe PDF File), 147 KB - [games_v14i1e82610_app3.pdf](#)]

Multimedia Appendix 4

Pretests and posttests.

[PDF File (Adobe PDF File), 341 KB - [games_v14i1e82610_app4.pdf](#)]

Multimedia Appendix 5

Estimated marginal means and 95% CIs.

[PDF File (Adobe PDF File), 182 KB - [games_v14i1e82610_app5.pdf](#)]

Multimedia Appendix 6

CONSORT-EHEALTH (V 1.6.1) checklist.

[PDF File (Adobe PDF File), 2430 KB - [games_v14i1e82610_app6.pdf](#)]

References

1. Quwaider M, Alabed A, Duwairi R. The impact of video games on the players behaviors: a survey. *Procedia Computer Sci* 2019;151:575-582. [doi: [10.1016/j.procs.2019.04.077](#)]
2. Granic I, Lobel A, Engels RCME. The benefits of playing video games. *Am Psychol* 2014 Jan;69(1):66-78. [doi: [10.1037/a0034857](#)] [Medline: [24295515](#)]
3. Cole C, Parada RH, Mackenzie E. A scoping review of video games and learning in secondary classrooms. *J Res Technol Educ* 2023;56(5):544-577. [doi: [10.1080/15391523.2023.2186546](#)]
4. Martinez L, Gimenes M, Lambert E. Entertainment video games for academic learning: a systematic review. *J Educ Comput Res* 2022;60(5):1083-1109. [doi: [10.1177/07356331211053848](#)]
5. Martinez L, Gimenes M, Lambert E. Video games and board games: effects of playing practice on cognition. *PLoS One* 2023;18(3):e0283654 [FREE Full text] [doi: [10.1371/journal.pone.0283654](#)] [Medline: [36972271](#)]
6. Brandl LC, Schrader A. Serious games in higher education in the transforming process to education 4.0—systematized review. *Educ Sci* 2024;14(3):281. [doi: [10.3390/educsci14030281](#)]

7. 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 (Lausanne)* 2019;6:321 [FREE Full text] [doi: [10.3389/fmed.2019.00321](https://doi.org/10.3389/fmed.2019.00321)] [Medline: [32047751](https://pubmed.ncbi.nlm.nih.gov/32047751/)]
8. Barbosa A, Whiting S, Simmonds P, Scotini Moreno RS, Mendes R, Breda J. Physical activity and academic achievement: an umbrella review. *Int J Environ Res Public Health* 2020;17(16):5972 [FREE Full text] [doi: [10.3390/ijerph17165972](https://doi.org/10.3390/ijerph17165972)] [Medline: [32824593](https://pubmed.ncbi.nlm.nih.gov/32824593/)]
9. Prot S, McDonald KA, Anderson CA, Gentile DA. Video games: good, bad, or other? *Pediatr Clin North Am* 2012;59(3):647-658. [doi: [10.1016/j.pcl.2012.03.016](https://doi.org/10.1016/j.pcl.2012.03.016)] [Medline: [22643171](https://pubmed.ncbi.nlm.nih.gov/22643171/)]
10. Andrade A, Correia CK, da Cruz WM, Bevilacqua GG. Acute effect of exergames on children's mood states during physical education classes. *Games Health J* 2019;8(4):250-256. [doi: [10.1089/g4h.2018.0083](https://doi.org/10.1089/g4h.2018.0083)] [Medline: [30730230](https://pubmed.ncbi.nlm.nih.gov/30730230/)]
11. Gao Z, Chen S. Are field-based exergames useful in preventing childhood obesity? A systematic review. *Obes Rev* 2014;15(8):676-691. [doi: [10.1111/obr.12164](https://doi.org/10.1111/obr.12164)] [Medline: [24602072](https://pubmed.ncbi.nlm.nih.gov/24602072/)]
12. Marques LM, Uchida PM, Barbosa SP. The impact of exergames on emotional experience: a systematic review. *Front Public Health* 2023;11:1209520 [FREE Full text] [doi: [10.3389/fpubh.2023.1209520](https://doi.org/10.3389/fpubh.2023.1209520)] [Medline: [37744509](https://pubmed.ncbi.nlm.nih.gov/37744509/)]
13. Kolovelonis A, Papastergiou M, Samara E, Goudas M. Acute effects of exergaming on students' executive functions and situational interest in elementary physical education. *Int J Environ Res Public Health* 2023;20(3):1902 [FREE Full text] [doi: [10.3390/ijerph20031902](https://doi.org/10.3390/ijerph20031902)] [Medline: [36767269](https://pubmed.ncbi.nlm.nih.gov/36767269/)]
14. Pérez-Jorge D, Martínez-Murciano MC, Contreras-Madrid AI, Alonso-Rodríguez I. The relationship between gamified physical exercise and mental health in adolescence: an example of open innovation in gamified learning. *Healthcare (Basel)* 2024;12(2):124 [FREE Full text] [doi: [10.3390/healthcare12020124](https://doi.org/10.3390/healthcare12020124)] [Medline: [38255013](https://pubmed.ncbi.nlm.nih.gov/38255013/)]
15. Sun H. Exergaming impact on physical activity and interest in elementary school children. *Res Q Exerc Sport* 2012;83(2):212-220. [doi: [10.1080/02701367.2012.10599852](https://doi.org/10.1080/02701367.2012.10599852)] [Medline: [22808707](https://pubmed.ncbi.nlm.nih.gov/22808707/)]
16. Pasco D, Roure C. Situational interest impacts college students' physical activity in a design-based bike exergame. *J Sport Health Sci* 2022;11(2):172-178 [FREE Full text] [doi: [10.1016/j.jshs.2021.03.003](https://doi.org/10.1016/j.jshs.2021.03.003)] [Medline: [33722758](https://pubmed.ncbi.nlm.nih.gov/33722758/)]
17. Haddock B, Siegel S, Costa P, Jarvis S, Klug N, Medina E, et al. Fitness assessment comparison between the "Jackie Chan Action Run" videogame, 1-mile run/walk, and the PACER. *Games Health J* 2012;1(3):223-227 [FREE Full text] [doi: [10.1089/g4h.2012.0007](https://doi.org/10.1089/g4h.2012.0007)] [Medline: [26193440](https://pubmed.ncbi.nlm.nih.gov/26193440/)]
18. Moholdt T, Weie S, Chorianopoulos K, Wang AI, Hagen K. Exergaming can be an innovative way of enjoyable high-intensity interval training. *BMJ Open Sport Exerc Med* 2017;3(1):e000258 [FREE Full text] [doi: [10.1136/bmjsem-2017-000258](https://doi.org/10.1136/bmjsem-2017-000258)] [Medline: [29021913](https://pubmed.ncbi.nlm.nih.gov/29021913/)]
19. Marsigliante S, My G, Mazzotta G, Muscella A. The effects of exergames on physical fitness, body composition and enjoyment in children: a six-month intervention study. *Children (Basel)* 2024;11(10):1172 [FREE Full text] [doi: [10.3390/children11101172](https://doi.org/10.3390/children11101172)] [Medline: [39457137](https://pubmed.ncbi.nlm.nih.gov/39457137/)]
20. Moller AC, Sousa CV, Lee KJ, Alon D, Lu AS. A comprehensive systematic review and content analysis of active video game intervention research. *Digit Health* 2023;9. [doi: [10.1177/20552076231171232](https://doi.org/10.1177/20552076231171232)]
21. Ebbinghaus H. Über Das Gedächtnis: Untersuchungen Zur Experimentellen Psychologie. Leipzig: Dunker & Humbolt; 1885.
22. Musfeld P, Souza AS, Oberauer K. Repetition learning is neither a continuous nor an implicit process. *Proc Natl Acad Sci USA* 2023;120(16):e2218042120 [FREE Full text] [doi: [10.1073/pnas.2218042120](https://doi.org/10.1073/pnas.2218042120)] [Medline: [37040406](https://pubmed.ncbi.nlm.nih.gov/37040406/)]
23. Liu W, Zeng N, McDonough DJ, Gao Z. Effect of active video games on healthy children's fundamental motor skills and physical fitness: a systematic review. *Int J Environ Res Public Health* 2020;17(21):8264 [FREE Full text] [doi: [10.3390/ijerph17218264](https://doi.org/10.3390/ijerph17218264)] [Medline: [33182327](https://pubmed.ncbi.nlm.nih.gov/33182327/)]
24. Logan SW, Ross SM, Chee K, Stodden DF, Robinson LE. Fundamental motor skills: a systematic review of terminology. *J Sports Sci* 2018;36(7):781-796. [doi: [10.1080/02640414.2017.1340660](https://doi.org/10.1080/02640414.2017.1340660)] [Medline: [28636423](https://pubmed.ncbi.nlm.nih.gov/28636423/)]
25. Oppici L, Stell FM, Utesch T, Woods CT, Foweather L, Rudd JR. A skill acquisition perspective on the impact of exergaming technology on foundational movement skill development in children 3-12 years: a systematic review and meta-analysis. *Sports Med Open* 2022;8(1):148 [FREE Full text] [doi: [10.1186/s40798-022-00534-8](https://doi.org/10.1186/s40798-022-00534-8)] [Medline: [36525118](https://pubmed.ncbi.nlm.nih.gov/36525118/)]
26. Gao Z, Wen X, Fu Y, Lee JE, Zeng N. Motor skill competence matters in promoting physical activity and health. *Biomed Res Int* 2021;2021:9786368 [FREE Full text] [doi: [10.1155/2021/9786368](https://doi.org/10.1155/2021/9786368)] [Medline: [35028311](https://pubmed.ncbi.nlm.nih.gov/35028311/)]
27. Barnett LM, Lai SK, Veldman SLC, Hardy LL, Cliff DP, Morgan PJ, et al. Correlates of gross motor competence in children and adolescents: a systematic review and meta-analysis. *Sports Med* 2016;46(11):1663-1688 [FREE Full text] [doi: [10.1007/s40279-016-0495-z](https://doi.org/10.1007/s40279-016-0495-z)] [Medline: [26894274](https://pubmed.ncbi.nlm.nih.gov/26894274/)]
28. Sousa CV, Lee K, Alon D, Sternad D, Lu AS. A systematic review and meta-analysis of the effect of active video games on postural balance. *Arch Phys Med Rehabil* 2023;104(4):631-644 [FREE Full text] [doi: [10.1016/j.apmr.2023.01.002](https://doi.org/10.1016/j.apmr.2023.01.002)] [Medline: [36669637](https://pubmed.ncbi.nlm.nih.gov/36669637/)]
29. McGann J, Issartel J, Hederman L, Conlan O. Hop.Skip.Jump.Games: the effect of "principled" exergameplay on children's locomotor skill acquisition. *Brit J Educ Tech* 2019;51(3):798-816. [doi: [10.1111/bjet.12886](https://doi.org/10.1111/bjet.12886)]

30. Li S, Song Y, Cai Z, Zhang Q. Are active video games useful in the development of gross motor skills among non-typically developing children? A meta-analysis. *BMC Sports Sci Med Rehabil* 2022;14(1):140 [FREE Full text] [doi: [10.1186/s13102-022-00532-z](https://doi.org/10.1186/s13102-022-00532-z)] [Medline: [35870986](https://pubmed.ncbi.nlm.nih.gov/35870986/)]
31. Trecroci A, Cavaggioni L, Caccia R, Alberti G. Jump rope training: balance and motor coordination in preadolescent soccer players. *J Sports Sci Med* 2015;14(4):792-798 [FREE Full text] [Medline: [26664276](https://pubmed.ncbi.nlm.nih.gov/26664276/)]
32. Daly-Smith A, Morris JL, Norris E, Williams TL, Archbold V, Kallio J, et al. Behaviours that prompt primary school teachers to adopt and implement physically active learning: a meta synthesis of qualitative evidence. *Int J Behav Nutr Phys Act* 2021;18(1):151 [FREE Full text] [doi: [10.1186/s12966-021-01221-9](https://doi.org/10.1186/s12966-021-01221-9)] [Medline: [34801039](https://pubmed.ncbi.nlm.nih.gov/34801039/)]
33. Chalkley AE, Mandelid MB, Thurston M, Daly-Smith A, Singh A, Huiberts I, et al. "Go beyond your own comfort zone and challenge yourself": a comparison on the use of physically active learning in Norway, the Netherlands and the UK. *Teach Teach Educ* 2022;118:103825 [FREE Full text] [doi: [10.1016/j.tate.2022.103825](https://doi.org/10.1016/j.tate.2022.103825)]
34. Goncalves A, Lespiau F, Briet G, Vaillant-Coindard E, Palermo A, Decobert E, et al. Exploring the use of a learning-based exergame to enhance physical literacy, soft skills, and academic learning in school-age children: pilot interventional study. *JMIR Serious Games* 2024;12:e53072 [FREE Full text] [doi: [10.2196/53072](https://doi.org/10.2196/53072)] [Medline: [38393767](https://pubmed.ncbi.nlm.nih.gov/38393767/)]
35. Cece V, Fargier P, Roure C, Lentillon-Kaestner V. Multidisciplinary teaching with an active video game: the effect on learning in mathematics and physical education. *Technol Pedagog Educ* 2025;34(1):121-136. [doi: [10.1080/1475939x.2024.2407380](https://doi.org/10.1080/1475939x.2024.2407380)]
36. Gao Z, Hannan P, Xiang P, Stodden DF, Valdez VE. Video game-based exercise, Latino children's physical health, and academic achievement. *Am J Prev Med* 2013;44(3 Suppl 3):S240-S246 [FREE Full text] [doi: [10.1016/j.amepre.2012.11.023](https://doi.org/10.1016/j.amepre.2012.11.023)] [Medline: [23415189](https://pubmed.ncbi.nlm.nih.gov/23415189/)]
37. Benzing V, Heinks T, Eggenberger N, Schmidt M. Acute cognitively engaging exergame-based physical activity enhances executive functions in adolescents. *PLoS One* 2016;11(12):e0167501 [FREE Full text] [doi: [10.1371/journal.pone.0167501](https://doi.org/10.1371/journal.pone.0167501)] [Medline: [28030542](https://pubmed.ncbi.nlm.nih.gov/28030542/)]
38. Diamond A. Executive functions. *Annu Rev Psychol* 2013;64:135-168 [FREE Full text] [doi: [10.1146/annurev-psych-113011-143750](https://doi.org/10.1146/annurev-psych-113011-143750)] [Medline: [23020641](https://pubmed.ncbi.nlm.nih.gov/23020641/)]
39. Melby-Lervåg M, Redick TS, Hulme C. Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": evidence from a meta-analytic review. *Perspect Psychol Sci* 2016;11(4):512-534 [FREE Full text] [doi: [10.1177/1745691616635612](https://doi.org/10.1177/1745691616635612)] [Medline: [27474138](https://pubmed.ncbi.nlm.nih.gov/27474138/)]
40. Sala G, Aksayli ND, Tatlidil KS, Tatsumi T, Gondo Y, Gobet F. Near and far transfer in cognitive training: a second-order meta-analysis. *Collabra Psychol* 2019;5(1):18 [FREE Full text] [doi: [10.1525/collabra.203](https://doi.org/10.1525/collabra.203)]
41. Linn MC, Petersen AC. Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Dev* 1985;56(6):1479-1498. [doi: [10.2307/1130467](https://doi.org/10.2307/1130467)]
42. Wai J, Lubinski D, Benbow CP. Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J Educ Psychol* 2009;101(4):817-835. [doi: [10.1037/a0016127](https://doi.org/10.1037/a0016127)]
43. Sala G, Tatlidil KS, Gobet F. Video game training does not enhance cognitive ability: a comprehensive meta-analytic investigation. *Psychol Bull* 2018;144(2):111-139. [doi: [10.1037/bul0000139](https://doi.org/10.1037/bul0000139)] [Medline: [29239631](https://pubmed.ncbi.nlm.nih.gov/29239631/)]
44. Uttal DH, Meadow NG, Tipton E, Hand LL, Alden AR, Warren C, et al. The malleability of spatial skills: a meta-analysis of training studies. *Psychol Bull* 2013;139(2):352-402. [doi: [10.1037/a0028446](https://doi.org/10.1037/a0028446)] [Medline: [22663761](https://pubmed.ncbi.nlm.nih.gov/22663761/)]
45. Moreau D, Clerc J, Mansy-Dannay A, Guerrien A. Enhancing spatial ability through sport practice. Evidence for an effect of motor training on mental rotation performance. *J Ind Diff* 2012;33(2):83-88. [doi: [10.1027/1614-0001/a000075](https://doi.org/10.1027/1614-0001/a000075)]
46. Jansen P, Kellner J, Rieder C. The improvement of mental rotation performance in second graders after creative dance training. *Creat Educ* 2013;04(06):418-422. [doi: [10.4236/ce.2013.46060](https://doi.org/10.4236/ce.2013.46060)]
47. Shepard RN, Metzler J. Mental rotation of three-dimensional objects. *Science* 1971;171(3972):701-703. [doi: [10.1126/science.171.3972.701](https://doi.org/10.1126/science.171.3972.701)] [Medline: [5540314](https://pubmed.ncbi.nlm.nih.gov/5540314/)]
48. Voyer D, Voyer S, Bryden MP. Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychol Bull* 1995;117(2):250-270. [doi: [10.1037/0033-2909.117.2.250](https://doi.org/10.1037/0033-2909.117.2.250)] [Medline: [7724690](https://pubmed.ncbi.nlm.nih.gov/7724690/)]
49. Hawes ZCK, Gilligan-Lee KA, Mix KS. Effects of spatial training on mathematics performance: a meta-analysis. *Dev Psychol* 2022;58(1):112-137 [FREE Full text] [doi: [10.1037/dev0001281](https://doi.org/10.1037/dev0001281)] [Medline: [35073120](https://pubmed.ncbi.nlm.nih.gov/35073120/)]
50. Fargier P, Champely S, Massarelli R, Ammary L, Hoyek N. Modelling response time in a mental rotation task by gender, physical activity, and task features. *Sci Rep* 2022;12(1):15559 [FREE Full text] [doi: [10.1038/s41598-022-19054-2](https://doi.org/10.1038/s41598-022-19054-2)] [Medline: [36114235](https://pubmed.ncbi.nlm.nih.gov/36114235/)]
51. Geertsen SS, Thomas R, Larsen MN, Dahn IM, Andersen JN, Krause-Jensen M, et al. Motor skills and exercise capacity are associated with objective measures of cognitive functions and academic performance in preadolescent children. *PLoS One* 2016;11(8):e0161960 [FREE Full text] [doi: [10.1371/journal.pone.0161960](https://doi.org/10.1371/journal.pone.0161960)] [Medline: [27560512](https://pubmed.ncbi.nlm.nih.gov/27560512/)]
52. Hoyek N, Champely S, Collet C, Fargier P, Guillot A. Is mental rotation ability a predictor of success for motor performance? *J Cogn Dev* 2014;15(3):495-505. [doi: [10.1080/15248372.2012.760158](https://doi.org/10.1080/15248372.2012.760158)]
53. Hidi SE, Renninger KA. The four-phase model of interest development. *Educ Psychol* 2006;41(2):111-127. [doi: [10.1207/s15326985ep4102_4](https://doi.org/10.1207/s15326985ep4102_4)]

54. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med* 2020;54(24):1451-1462 [[FREE Full text](#)] [doi: [10.1136/bjsports-2020-102955](https://doi.org/10.1136/bjsports-2020-102955)] [Medline: [33239350](https://pubmed.ncbi.nlm.nih.gov/33239350/)]
55. Pasco D, Roure C, Kermarrec G, Pope Z, Gao Z. The effects of a bike active video game on players' physical activity and motivation. *J Sport Health Sci* 2017 Mar;6(1):25-32 [[FREE Full text](#)] [doi: [10.1016/j.jshs.2016.11.007](https://doi.org/10.1016/j.jshs.2016.11.007)] [Medline: [30356595](https://pubmed.ncbi.nlm.nih.gov/30356595/)]
56. Chan JSY, Luo Y, Yan JH, Cai L, Peng K. Children's age modulates the effect of part and whole practice in motor learning. *Hum Mov Sci* 2015;42:261-272 [[FREE Full text](#)] [doi: [10.1016/j.humov.2015.06.002](https://doi.org/10.1016/j.humov.2015.06.002)] [Medline: [26112404](https://pubmed.ncbi.nlm.nih.gov/26112404/)]
57. Renninger KA, Hidi SE. *The Power of Interest for Motivation and Engagement*. New York: Routledge; 2017.
58. Schiefele U. Situational and individual interest. In: Wentzel KR, Wigfield A, editors. *Handbook of Motivation at School*. New York: Routledge; 2009:197-222.
59. Huang C, Gao Z. Associations between students' situational interest, mastery experiences, and physical activity levels in an interactive dance game. *Psychol Health Med* 2013;18(2):233-241. [doi: [10.1080/13548506.2012.712703](https://doi.org/10.1080/13548506.2012.712703)] [Medline: [22881580](https://pubmed.ncbi.nlm.nih.gov/22881580/)]
60. Aadland E, Ylvisåker E. Reliability of the ActiGraph GT3X+ accelerometer in adults under free-living conditions. *PLoS One* 2015;10(8):e0134606 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0134606](https://doi.org/10.1371/journal.pone.0134606)] [Medline: [26274586](https://pubmed.ncbi.nlm.nih.gov/26274586/)]
61. Roure C. Clarification du construit de l'intérêt en situation en éducation physique. *STAPS* 2020;130(4):61-77. [doi: [10.3917/sta.130.0061](https://doi.org/10.3917/sta.130.0061)]
62. Albaret JM, Aubert E. Étalonnage 15-19 ans du test de rotation mentale de Vandenberg. *Evol Psychom* 1996;8(34):205-215 [[FREE Full text](#)]
63. Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. *Percept Mot Skills* 1978;47(2):599-604. [doi: [10.2466/pms.1978.47.2.599](https://doi.org/10.2466/pms.1978.47.2.599)] [Medline: [724398](https://pubmed.ncbi.nlm.nih.gov/724398/)]
64. Dalloio L, Ceciliani A, Sanna T, Garulli A, Leoni E. Proposal for an enhanced physical education program in the primary school: evaluation of feasibility and effectiveness in improving physical skills and fitness. *J Phys Act Health* 2016;13(10):1025-1034 [[FREE Full text](#)] [doi: [10.1123/jpah.2015-0694](https://doi.org/10.1123/jpah.2015-0694)] [Medline: [27172612](https://pubmed.ncbi.nlm.nih.gov/27172612/)]
65. Harre D. *Principles of Sports Training: Introduction to the Theory and Methods of Training*. Berlin: Sportverlag; 1982.
66. Vandorpe B, Vandendriessche J, Lefevre J, Pion J, Vaeyens R, Matthys S, et al. The körperkoordinations test für kinder: reference values and suitability for 6-12-year-old children in flanders. *Scand J Med Sci Sports* 2011;21(3):378-388 [[FREE Full text](#)] [doi: [10.1111/j.1600-0838.2009.01067.x](https://doi.org/10.1111/j.1600-0838.2009.01067.x)] [Medline: [20136753](https://pubmed.ncbi.nlm.nih.gov/20136753/)]
67. Nosworthy N, Bugden S, Archibald L, Evans B, Ansari D. A two-minute paper-and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS One* 2013;8(7):e67918 [[FREE Full text](#)] [doi: [10.1371/journal.pone.0067918](https://doi.org/10.1371/journal.pone.0067918)] [Medline: [23844126](https://pubmed.ncbi.nlm.nih.gov/23844126/)]
68. Vogel SE, Haigh T, Sommerauer G, Spindler M, Brunner C, Lyons IM, et al. Processing the order of symbolic numbers: a reliable and unique predictor of arithmetic fluency. *J Numer Cogn* 2017;3(2):288-308. [doi: [10.5964/jnc.v3i2.55](https://doi.org/10.5964/jnc.v3i2.55)]
69. Thevenot C, Castel C, Fanget M, Fayol M. Mental subtraction in high- and lower skilled arithmetic problem solvers: verbal report versus operand-recognition paradigms. *J Exp Psychol Learn Mem Cogn* 2010;36(5):1242-1255. [doi: [10.1037/a0020447](https://doi.org/10.1037/a0020447)] [Medline: [20804294](https://pubmed.ncbi.nlm.nih.gov/20804294/)]
70. Tschentscher N, Hauk O. How are things adding up? Neural differences between arithmetic operations are due to general problem solving strategies. *Neuroimage* 2014;92:369-380 [[FREE Full text](#)] [doi: [10.1016/j.neuroimage.2014.01.061](https://doi.org/10.1016/j.neuroimage.2014.01.061)] [Medline: [24525170](https://pubmed.ncbi.nlm.nih.gov/24525170/)]
71. Barrett TJ. Computations using analysis of covariance. *WIREs Comp Stats* 2011;3(3):260-268 [[FREE Full text](#)] [doi: [10.1002/wics.165](https://doi.org/10.1002/wics.165)]
72. Conover WJ, Iman RL. Analysis of covariance using the rank transformation. *Biometrics* 1982;38(3):715 [[FREE Full text](#)] [doi: [10.2307/2530051](https://doi.org/10.2307/2530051)]
73. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. New York: Routledge; 1988.
74. Huber SK, Held JPO, de Bruin ED, Knols RH. Personalized motor-cognitive exergame training in chronic stroke patients-a feasibility study. *Front Aging Neurosci* 2021;13:730801 [[FREE Full text](#)] [doi: [10.3389/fnagi.2021.730801](https://doi.org/10.3389/fnagi.2021.730801)] [Medline: [34744688](https://pubmed.ncbi.nlm.nih.gov/34744688/)]
75. Wallace B, Hofelich BG. Process generalization and the prediction of performance on mental imagery tasks. *Mem Cognit* 1992;20(6):695-704. [doi: [10.3758/bf03202719](https://doi.org/10.3758/bf03202719)] [Medline: [1435272](https://pubmed.ncbi.nlm.nih.gov/1435272/)]
76. Wright R, Thompson WL, Ganis G, Newcombe NS, Kosslyn SM. Training generalized spatial skills. *Psychon Bull Rev* 2008;15(4):763-771 [[FREE Full text](#)] [doi: [10.3758/pbr.15.4.763](https://doi.org/10.3758/pbr.15.4.763)] [Medline: [18792502](https://pubmed.ncbi.nlm.nih.gov/18792502/)]
77. Di Cagno A, Battaglia C, Giombini A, Piazza M, Fiorilli G, Calcagno G, et al. Time of day-effects on motor coordination and reactive strength in elite athletes and untrained adolescents. *J Sports Sci Med* 2013;12(1):182-189 [[FREE Full text](#)] [Medline: [24149743](https://pubmed.ncbi.nlm.nih.gov/24149743/)]
78. Sui Y, Cui L, Jia B, Ding X, He M, Da Y, et al. Correlation between gross motor coordination and basic coordination capacities in normal-weight and overweight/obese children aged 9-10 years. *PeerJ* 2024;12:e17865 [[FREE Full text](#)] [doi: [10.7717/peerj.17865](https://doi.org/10.7717/peerj.17865)] [Medline: [39135953](https://pubmed.ncbi.nlm.nih.gov/39135953/)]

79. Arsalidou M, Pawliw-Levac M, Sadeghi M, Pascual-Leone J. Brain areas associated with numbers and calculations in children: meta-analyses of fMRI studies. *Dev Cogn Neurosci* 2018;30:239-250 [FREE Full text] [doi: [10.1016/j.dcn.2017.08.002](https://doi.org/10.1016/j.dcn.2017.08.002)] [Medline: [28844728](https://pubmed.ncbi.nlm.nih.gov/28844728/)]
80. Metcalfe AWS, Campbell JID. Adults' strategies for simple addition and multiplication: verbal self-reports and the operand recognition paradigm. *J Exp Psychol Learn Mem Cogn* 2011;37(3):661-672 [FREE Full text] [doi: [10.1037/a0022218](https://doi.org/10.1037/a0022218)] [Medline: [21261421](https://pubmed.ncbi.nlm.nih.gov/21261421/)]
81. Isoda M, Olfos R. Problematics for conceptualization of multiplication. In: Isoda M, Olfos R, editors. *Teaching Multiplications with Lesson Study, Japanese and Ibero-American theories for International Mathematics Education*. Cham: Springer; 2021:37-64.
82. Suárez-Pellicioni M, Prado J, Booth JR. Lack of improvement in multiplication is associated with reverting from verbal retrieval to numerical operations. *Neuroimage* 2018;183:859-871 [FREE Full text] [doi: [10.1016/j.neuroimage.2018.08.074](https://doi.org/10.1016/j.neuroimage.2018.08.074)] [Medline: [30189338](https://pubmed.ncbi.nlm.nih.gov/30189338/)]
83. Yang X, Yu X. The relationship between mental rotation and arithmetic: do number line estimation, working memory, or place-value concept matter? *Br J Educ Psychol* 2021;91(3):793-810 [FREE Full text] [doi: [10.1111/bjep.12403](https://doi.org/10.1111/bjep.12403)] [Medline: [33368175](https://pubmed.ncbi.nlm.nih.gov/33368175/)]
84. Li M, Liu D, Li M, Dong W, Huang Y, Chen Q. Addition and subtraction but not multiplication and division cause shifts of spatial attention. *Front Hum Neurosci* 2018;12:183 [FREE Full text] [doi: [10.3389/fnhum.2018.00183](https://doi.org/10.3389/fnhum.2018.00183)] [Medline: [29773985](https://pubmed.ncbi.nlm.nih.gov/29773985/)]
85. Perneger TV. What's wrong with Bonferroni adjustments. *BMJ* 1998;316(7139):1236-1238 [FREE Full text] [doi: [10.1136/bmj.316.7139.1236](https://doi.org/10.1136/bmj.316.7139.1236)] [Medline: [9553006](https://pubmed.ncbi.nlm.nih.gov/9553006/)]
86. Neuburger S, Jansen P, Heil M, Quaiser-Pohl C. Gender differences in pre-adolescents' mental-rotation performance: do they depend on grade and stimulus type? *Pers Individ Dif* 2011;50(8):1238-1242 [FREE Full text] [doi: [10.1016/j.paid.2011.02.017](https://doi.org/10.1016/j.paid.2011.02.017)]
87. Hyde JS, Fennema E, Lamon SJ. Gender differences in mathematics performance: a meta-analysis. *Psychol Bull* 1990;107(2):139-155. [doi: [10.1037/0033-2909.107.2.139](https://doi.org/10.1037/0033-2909.107.2.139)] [Medline: [2138794](https://pubmed.ncbi.nlm.nih.gov/2138794/)]
88. Li BJ, Ratan R, Lwin MO. Virtual game changers: how avatars and virtual coaches influence exergame outcomes through enactive and vicarious learning. *Behav Inform Technol* 2022;41(7):1529-1543 [FREE Full text] [doi: [10.1080/0144929x.2021.1884290](https://doi.org/10.1080/0144929x.2021.1884290)]
89. Data-MVPA-Situational-Interest-Pretests-Postests. Zenodo. URL: <https://doi.org/10.5281/zenodo.18016022> [accessed 2026-02-24]

Abbreviations

ANCOVA: analysis of covariance
CG: control group
DEx: dance-based exergaming
EG: experimental group
MANOVA: multivariate analysis of variance
MVPA: moderate to vigorous physical activity
PE: physical education
TEx: precision ball-throwing exergaming

Edited by S Brini; submitted 20.Aug.2025; peer-reviewed by A Goncalves, H Namba; comments to author 05.Dec.2025; accepted 09.Feb.2026; published 19.Mar.2026.

Please cite as:

Fargier P, Cece V, Lentillon-Kaestner V, Roure C
The Effects of Dance-Based Exergaming on Mental Rotation, General Motor Coordination, and Math Achievement in Adolescent Students: Nonrandomized Controlled Pilot Study
JMIR Serious Games 2026;14:e82610
URL: <https://games.jmir.org/2026/1/e82610>
doi: [10.2196/82610](https://doi.org/10.2196/82610)
PMID:

©Patrick Fargier, Valérian Cece, Vanessa Lentillon-Kaestner, Cédric Roure. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 19.Mar.2026. 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.

User Experiences of Behavioral and Psychological Change Techniques in a Walking-Based Mobile Exergame: Cross-Sectional Qualitative Study

Marianna Antoniadou^{1,2*}, MSc; Helena Orädd^{3*}, MSc; Aseel Berglund³, PhD; Erik Berglund³, PhD; Anna Strömberg^{1,4}, PhD; Tiny Jaarsma¹, PhD; Leonie Klompstra¹, PhD

¹Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

²Post Graduate School for Integrated Care, Örebro University, Örebro, Sweden

³Department of Computer and Information Science, Linköping University, Linköping, Sweden

⁴Department of Cardiology, Linköping University, Linköping, Sweden

*these authors contributed equally

Corresponding Author:

Leonie Klompstra, PhD

Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden

Abstract

Background: Physical activity plays a central role in the course and progression of chronic conditions in older adults. However, individuals within this population tend to have an inactive lifestyle. Exergaming, which is defined as the integration of physical activity with game-based elements, offers a promising approach to promote physical activity in individuals with chronic conditions. Despite its potential, limited evidence exists on how specific game elements influence behavioral and psychological outcomes in this population.

Objective: The aim of this study is to explore the behavioral and psychological change techniques experienced by individuals with chronic conditions using a walking-based mobile exergame, called Heart Farming.

Methods: A cross-sectional qualitative design was used based on a gamification framework, using data from semistructured interviews with 14 participants aged 67 to 92 years who used the Heart Farming exergame for 3 months. Participants with chronic conditions, including heart failure, Parkinson disease, type 2 diabetes, stroke, or rheumatic disease, were recruited from 2 ongoing studies. Data were analyzed using deductive content analysis and presented based on the gamification framework.

Results: The data analysis revealed increased motivation to walk, which was facilitated by game elements such as goals, rewards, feedback, and planning. Participants valued not only the sense of progression and achievement within the game but also the real-world benefits, such as spending time in nature or feeling a sense of community with others. Exergaming was integrated into daily routines by supporting the planning and structuring of daily activities. It was also perceived as enjoyable, especially due to its farming theme and visual design. Behavioral change techniques such as goal setting, feedback, and social support were commonly experienced, whereas focus on past success (as described in the gamification framework) was not used as a technique by the exergame. Psychological techniques, including self-monitoring and stress management, were mentioned. A feeling of discomfort due to playing in public was reported, and participants varied in terms of the levels of digital literacy. Social interaction features were not adequately used, even though some participants created informal support groups to exergame.

Conclusions: This study contributes to the exergaming literature by examining behavioral and psychological change techniques from the perspective of individuals with chronic conditions. Exergaming incorporates a variety of behavioral and psychological change techniques, which were experienced by the participants in various ways. Individuals' previous knowledge of exergaming, interests, and illness progression were mentioned as factors that can influence their experiences of exergaming. By increasing the understanding of how game elements are experienced and how they influence health-related behavior, the findings of this study may inform the development of exergames that are better tailored to users' needs. Further research is needed to refine behavioral change techniques and assess condition-specific adaptations to maximize engagement and health outcomes.

Trial Registration: ClinicalTrials.gov NCT05641662; <https://clinicaltrials.gov/study/NCT05641662>

(*JMIR Serious Games* 2026;14:e78776) doi:[10.2196/78776](https://doi.org/10.2196/78776)

KEYWORDS

exergames; physical activity; technology; behavioral change techniques; psychological change techniques; qualitative study

Introduction

Background

Approximately one-third of adults worldwide are living with at least 1 chronic condition [1]. Being physically active can have a positive impact on the course and progression of different chronic conditions, such as chronic heart failure [2], stroke [3], or Parkinson disease [4]. However, despite the significance of being physically active, individuals with chronic conditions tend to have an inactive lifestyle. This is characterized by low levels of physical activity and sedentary behavior like sitting or lying down [5].

Gamification is a well-known and widely used approach to increase motivation and user engagement by implementing game design elements in nongame contexts [6], including health care and behavioral change [7]. Gamification emphasizes identifiable behavioral and psychological change techniques, such as goal setting, feedback, and self-monitoring, from which psychological and behavioral change can emerge [8]. Although these techniques can be layered onto existing activities or systems, they can also be part of a unified game structure as a *serious game*. Serious games are games with a primary nonentertainment objective, such as promoting learning, health, or behavioral change, that fundamentally shape the design and function of a game [9]. In their book on serious games, Dörner et al [9] introduced the term “characterizing goal” to describe these primary nonentertainment objectives of serious games. When the characterizing goal refers to exercises or other physical activities, the game is typically called an *exergame*.

Exergaming, which combines gameplay with physical activity in real life, is a relatively new approach to increasing physical activity for individuals with chronic conditions. Exergames are serious games for health that can possibly increase motivation to physical activity [10] and improve physical function, balance, exercise capacity, and energy expenditure [2,11,12]. Although many traditional exergames are designed for indoor use in front of a screen, some individuals might prefer outdoor exergames rather than being confined to screen-based play indoors [13-15]. Mobile exergaming is a suitable option for remote physical activity interventions, particularly when in-person meetings with health care professionals are not possible. Smartphones are equipped with built-in technologies, such as accelerometers, cameras, and sensors [16].

Mobile exergames employ a range of technologies to support and motivate physical activity, including indoor, outdoor, and hybrid solutions. GPS is a technology that relies on satellite positioning to track movement over larger areas. GPS is used in location-based exergames that are primarily suited for outdoor use [17,18]. An example of a game that uses this technology is Pokémon Go [19]. However, location-based exergames, including Pokémon Go, have previously been criticized as inappropriate or even dangerous. Pokémon Go has been shown to contribute to accidents caused by players' inattention to their surroundings while walking or even driving [20,21]. In contrast, augmented reality is a technology that integrates digital content with the real-world environment using cameras and sensors, enabling interactive and context-aware experiences that can be

used both indoors and outdoors, while also supporting precise movement and spatial tracking [16].

Exergames use various techniques that aim to achieve behavioral change and can thus be useful in establishing healthy behaviors in individuals with chronic conditions [22]. The combination of key elements (eg, reward, challenge, feedback, information, goals, and planning) with the main aims of rehabilitation programs, such as motor skills learning, long-term retention, and transfer of skills, makes exergaming a valuable tool for health care professionals [23,24]. Although the physiological component (ie, the physical activity) is important for an exergame to be effective, it is also important to keep the psychological attractiveness (ie, the “fun”) of exergames in mind [25,26]. The attractive nature of exergames might enhance individuals' motivation to engage in physical activity, which may in turn lead to positive outcomes in terms of dose response, adherence to, and long-term effectiveness of the rehabilitation programs. Although exergaming can help distract from unpleasant feelings like pain, there is a negative correlation between perceived exertion and enjoyment [27,28], and it is important to adapt the intensity of the exercise to avoid feelings of overexertion.

There are several previous studies indicating the health benefits of exergaming in individuals with chronic conditions. A recent randomized controlled trial (RCT) on the home-based and tailored exergaming interventions has shown them to be feasible and safe for individuals with movement disorders, with adjustable difficulty levels, real-time feedback, and balance-focused tasks being particularly important for addressing gait and postural control in Parkinson disease [29]. Findings from studies conducted on people with heart failure demonstrate that exergaming, alongside other exercise modalities such as yoga, can positively influence exercise capacity as well as physical and mental health, highlighting the importance of designing exergames that are safe, motivating, and adaptable to varying functional abilities to support sustained participation among older adults with chronic conditions [30]. In addition, the findings of a systematic review show that exergaming can improve quality of life, physical functioning, and psychological well-being among people living with a wide range of chronic conditions [22]. Gamification-based interventions have also been found to enhance self-management, motivation, and long-term engagement in chronic disease care, further supporting the potential of exergames as a complementary tool in chronic condition management [31].

Study Objectives

Although there is much research to support the potential health benefits of exergaming, there is a gap in evidence regarding the game elements that impact health behavior and how individuals experience them. To fully understand the potential of exergames as a tool to increase physical activity in people with chronic conditions, further exploration of this population's experiences of exergaming is required. By gaining insight into how gamification techniques are experienced, researchers and developers may be able to better tailor and adapt exergaming interventions to the user's needs. Therefore, the purpose of this study is to explore the experienced behavioral and psychological

change techniques in individuals with heart failure, Parkinson disease, type 2 diabetes, stroke, or rheumatic disease when using a walking-based mobile exergame.

Methods

Research Design Overview

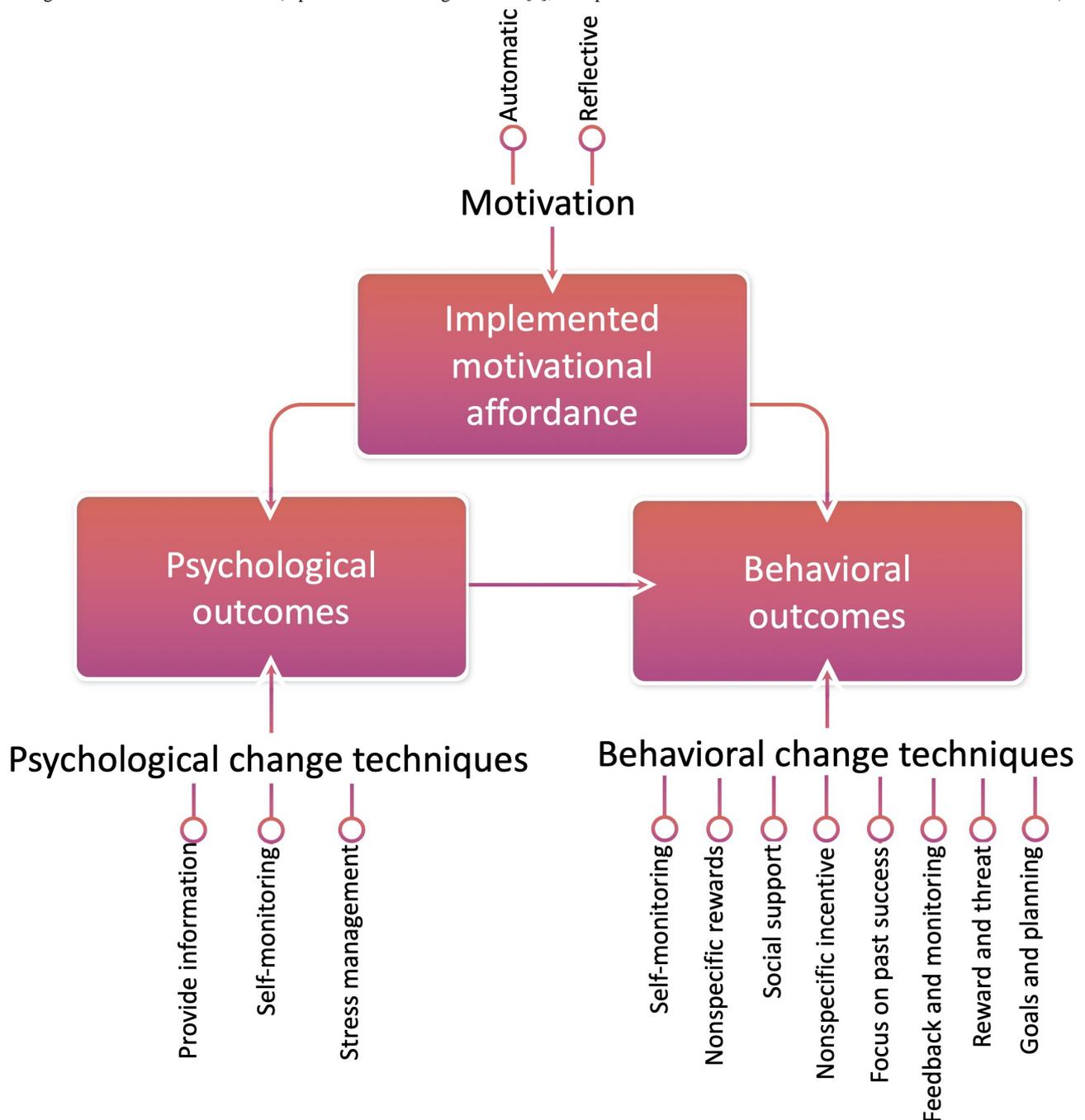
This study adopted a qualitative descriptive cross-sectional study design and was conducted within the Heart-eXgame (Heart-eXg) study, which is an international RCT, and the Heart Farming Chronic Conditions (HF-CC) study. The study was reported in accordance with the Journal Article Reporting Standards for Qualitative Research guidelines [32].

The Gamification Framework

Gamification may be used in exergaming to maintain behavioral change and encourage physical exercise [33]. Gamification is

defined as the use of game elements in a nongaming context, for example, behavioral change [6]. Gamification consists of 3 main components (Figure 1): implemented motivational affordances (properties of the gamified system that allow the user to experience competence, relatedness, and autonomy), psychological outcomes (induced by motivational affordance and psychological change techniques), and behavioral outcomes (induced by motivational affordances, psychological outcomes, and behavioral change techniques) [8]. The psychological change techniques are *providing information, self-monitoring, and stress management*, while the included behavioral change techniques are *self-monitoring, nonspecific rewards, social support, nonspecific incentives, focus on past success, feedback and monitoring, reward and threat, and goals and planning*.

Figure 1. The gamification framework for exergaming by Berglund et al [8], showcasing how behavioral and psychological change techniques influence psychological and behavioral outcomes (reproduced from Berglund et al [8], with permission from the authors under the CC BY-NC 4.0 license).



Specific game elements used to impact health behaviors can also be sorted as different principles of gamification [8]. These principles state that the gamified system should provide *meaningful purpose* by aligning goals with the user's goal or interest, include *meaningful choice* by giving users agency over how they reach the goals, *support player archetypes* by leveraging individual characteristics of the user, clearly communicate *feedback* on how the user's actions affect their progress, and provide *visibility* over the user's progress. These principles can be used to sort specific game elements according to their purpose. For example, digital rewards and game levels can be used to communicate progress (visibility) and which actions cause progression (feedback).

The Heart Farming Exergame

The mobile exergame Heart Farming used in this study was developed by a research team of game developers and health care professionals [16,34]. It was developed using the gamification framework described earlier [8], with the characterizing goal of motivating individuals with heart failure to walk 10 minutes more per day [16].

The exergame uses walking as the main interaction to drive the game forward. To allow for precise tracking both outside and inside, the exergame uses augmented reality technology to track the phone's, and therefore the player's, movements by analyzing the view from the mobile phone's camera [16]. The theme of the exergame is a farm where the player can buy fields and

manage them by sowing, watering, and harvesting crops. All actions in the game have a set number of meters the player must walk in real life to complete the action. The distance needed to complete actions is based on the player's individual walking goals to balance the effort needed for people with different abilities.

To progress in the exergame, the player must create a plan with actions to perform and walk until all tasks are completed. The exergame also included an option called "autoplay," where a plan is created automatically, allowing the player to walk without actively planning their actions. The created plan will yield random crops and is not as effective in progressing in the game as actively planning. Multiple gamification elements are also used in the exergame as displayed in [Table 1](#).

Table 1. Descriptions of key game elements in the Heart Farming walking-based exergame and how they utilize the gamification elements described by Berglund et al [8].

Gamification element	Game elements in Heart Farming
Leaderboards	A leaderboard displays the 3 players who gained the most experience points the previous day.
Levels of achievements or ranks	Levels are based on the player's experience points, and leveling up will also give access to new game contents.
Score systems	The number of meters the player has walked is used to reach their daily goal, set to how many meters would be needed for that player to walk for 10 minutes.
Experience points reward systems	All actions performed in the game are associated with a number of experience points the player earns when completing them. The amount of experience points earned for an action is the same for all players, even though the meters needed is determined by each player's own walking ability.
Item-granting system rewards	Players receive a gift each day when they open the exergame. The gift contains a random crop they can sell or store to use later.
Resources	Crops can be sold for money or stored to use for specific achievements, for example, to fulfill requirements for unlocking an animal or a neighbor. The money is used to buy new fields to sow new types of crops and expand the farm and progress in the game.
Achievement systems	Heart Farming has a mission book with animals that can be unlocked and placed on the farm by doing all the tasks specified by that animal. There are also neighboring houses placed on the farm that request certain crops, and selling the requested crops fills up a series of hearts that resets at the start of every week. If all hearts are filled before the end of the week, the player gains a decorative item on that neighbor's house.
Feedback messages	Sound effects are used to communicate completion of tasks and plans while the player is walking, as well as let the player know if the game is registering walking or not. Animations are used when achieving goals, such as a trophy appearing when reaching the daily goal or a heart expanding to show a new item when completing the requests from a neighbor.
Unlocking mechanisms	The crops the player can grow are limited by their current level and are unlocked when leveling up. Leveling up also unlocks some available actions in the game, such as being able to fish in the river or do squat exercises for extra crops.
Social or peer pressure	Players can send messages to other players in the form of emojis and can help each other by sending crops from their own farm.

Study Participants

The Heart-eXg study is registered on clinicaltrials.gov NCT05641662. In the Heart-eXg study, participants had to be diagnosed with heart failure by a physician, while participants in the HF-CC study had to self-report at least one of the following chronic conditions: heart failure, Parkinson disease, type 2 diabetes, stroke, or rheumatic disease. In both studies, participants had to be older than 18 years of age and in a stable condition. Participants who were unable to perform a 6-minute walk test or unwilling to use a smartphone to exergame were not included. Further exclusion criteria were severe cognitive

dysfunction, anticipated short-term survival, or difficulties in understanding or reading the Swedish language. All participants provided informed consent before participating in the study. Participants in the intervention group of the Heart-eXg study and the participants in the HF-CC study were advised to walk 10 minutes more every day for 3 months using a mobile exergame. For further information, see the Heart-eXg study website [35].

All participants were introduced to the exergame during introductory sessions in the university, outpatient clinics, or offices of the patient organizations, where members of the research team provided oral instructions and assistance to install

the exergame on their mobile phones. When needed, the research team offered individual help to participants, including more detailed explanations or providing practical assistance with the exergame. In addition, participants were given step-by-step written instructions (manual) with QR codes to short instructional videos. The manual was provided in Swedish with simple vocabulary and illustrations of the exergame to facilitate the understanding of the different functions/steps in the exergame. During the first month of the intervention, participants also received follow-up calls weekly, where they could also raise questions or issues regarding the exergame. Participants in the intervention group of the Heart-eXg study and the participants in the HF-CC study were all advised to walk 10 minutes more every day for 3 months using a mobile exergame called Heart Farming. Interviews were conducted at the 3-month follow-up.

A purposeful sampling strategy was used to recruit participants with variation in their playing patterns (eg, participants who played extensively and others who primarily used the autoplay function), symptoms, age, and sex. This strategy was chosen to ensure variability within the sample and to capture rich, information-dense data that could illuminate different experiences relevant to the study aim.

Data Collection

The semistructured interviews were conducted by 3 researchers: one of the authors of this paper (HO), who was actively involved in the exergame's development as a game developer, and 2 health care professionals (a nurse and an occupational therapist) who were not involved as authors in the study. To mitigate bias, the interviewers did not have any pre-established relationship with study participants. If, for instance, one interviewer conducted follow-up phone calls, another researcher was assigned to perform the interview to maintain objectivity. Participants were not informed of HO's involvement in the development of the game to minimize the risk of response bias. To further minimize potential bias related to HO's involvement in developing the exergame, interviews were conducted using a semistructured guide, and reflexive discussions were held within the research team throughout data collection and analysis. In addition, data analysis involved multiple researchers who were not involved in the game's development, ensuring analytical triangulation and critical appraisal of interpretations.

The interview guide was divided into 3 different components, including the participants' experiences of playing the Heart Farming exergame, its influence on daily life and habits, and its different game elements. The development of the interview guide was based on the gamification framework developed by Berglund et al [8] to focus on how the gamification elements were experienced by the participants. The interview guide was also informed by Klompstra et al [13] to include the overall experience of using the Heart Farming exergame, including daily life and habits. The interview started with an introductory question ("Can you tell me about your experiences with playing this game for the last 3 months?") and continued with more specific questions about the exergame. Probing questions, such as "Could you elaborate?" "Could you provide an example?" or "What do you mean by saying that?" were used to facilitate

the conversation and provide more accurate descriptions and rich information. A translated version of the interview guide is attached as [Multimedia Appendix 1](#).

A pilot interview was first conducted with a participant to assess the clarity of the questions. After discussions with the research team, no further modifications were needed to the interview guide, based on this pilot interview. Due to the relevance of the information, the pilot interview was included in the data analysis. No field notes were kept by the interviewers. After conducting 12 interviews, information power was achieved because of clear redundancy with repetition of patterns in the data without novel additions [36]. To confirm this, 2 additional interviews were conducted. No additional new information came from the last 2 interviews.

The interviews took place at the participants' homes, one of the university campuses, or over the telephone, depending on the participants' preferences and the feasibility of traveling to and from the interview. In 2 phone interviews, the participants' spouses were present during the interview. During the analysis, the researchers made sure not to include any of the spouses' comments during the analysis, and only the answers from the participants were included. The interviews were recorded and transcribed for analysis, with the participants' consent. The length of the interviews varied between 18 and 69 minutes, with an average length of 35 (SD 15) minutes. All interviews were conducted in Swedish, which is the native language of both the participants and the interviewers.

Data Analysis

The data analysis was performed after the completion of data collection. A deductive content analysis following the steps suggested by Elo and Kyngäs [37] was applied. The framework introduced by Berglund et al [8] was used in a descriptive way to structure the analysis. A categorization matrix was developed based on the behavioral and psychological change techniques introduced in the framework and visualized in [Figure 1](#).

During the preparation phase, all authors read through the transcribed interviews to familiarize themselves with the data and gain an in-depth understanding of the content. In the organization phase, the first 3 interviews were analyzed in parallel by 4 of the authors (MA, HO, LK, and AB), and the meaning units were allocated to the categorization matrix based on the gamification framework for exergaming. Thereafter, the remaining interviews were independently analyzed by the first 2 authors (MA and HO), who equally divided the work. To facilitate the coding process, the NVivo 15 (Lumivero) software was used. Finally, in the reporting phase, the findings were presented in alignment with the structured categories in a way that ensured clarity.

Methodological rigor was ensured by following the guidelines established for qualitative studies, introduced by Guba and Lincoln [38]. Analysis of triangulation was ensured through regular discussions between the authors, aiming at reaching a consensus considering the categorization of the data. The interviews were read multiple times to get an overall understanding of the content, and the parts of the interviews were constantly compared with the whole. Authentic quotations

were translated into English and presented as part of the results to increase the trustworthiness of the findings. The translation of the quotations was performed by the 2 first authors (MA and HO), and another researcher (LK) verified the accuracy of the translations.

Researcher Reflexivity

Two female researchers (MA and HO) conducted the biggest part of the data analysis presented in this paper. Both were PhD students employed at Linköping University. MA had a background in occupational therapy with a master's education, and HO was a game developer with a master's education as well. Their personal presumptions were that physical activity can be promoted by the Heart Farming exergame and that it can motivate participants to walk more on a daily basis. Both presumed that exergaming can influence daily life and habits.

Their contribution to developing the interview guide was equal. MA did not participate in data collection, while HO performed the interviews with 2 other health care professionals. The transcription of the interviews and the data analysis were mainly performed by MA and HO. To mitigate biases, part of the data analysis was performed by LK and AB, and a discussion took place considering the data analysis. All the authors gained closeness to the data by repeatedly reading the interview transcripts and listening to the original interview recordings. Reflexive discussions with co-authors AB, EB, AS, TJ, and LK supported MA and HO in enacting researcher reflexivity. AB and EB had a background in game development, AS and TJ had a nursing background, and LK had a background in public health. They were all experienced researchers in the field of exergaming and individuals with chronic conditions.

Ethical Considerations

The study is registered on clinicaltrials.gov with the ID NCT05641662. The Heart-eXg study and the HF-CC study were conducted in line with the Declaration of Helsinki. The Heart-eXg received ethical approval from the Regional Ethical Review Board in Linköping (Dnr 2021 - 03314), as did the HF-CC study (Dnr 2023-06398-01). Participants received written and verbal information about the study's aim before providing written informed consent to participate.

Participant privacy and confidentiality were carefully protected throughout the study. All collected data were anonymized prior to analysis, and any potentially identifiable information was removed from the transcribed interviews. The anonymized data were stored securely in a separate folder, accessible only to the research team. No financial or other forms of compensation were provided to the participants in this study. The manuscript and all supplementary materials do not include images or other content that could enable the identification of individual

participants. As such, additional consent for publication of identifiable materials was not required.

Due to the sensitive nature of personal data and study materials, data cannot be made freely available. However, by contacting the corresponding author, procedures for sharing data, analytic methods, and study materials for reproducing the results or replicating the procedure can be arranged in accordance with Swedish legislation.

Results

Participant Details

A total of 14 participants were included in this study. Study participants were individuals with chronic conditions who had access to exergaming in the Heart-eXg or the HFCC study for 3 months. There were 8 (57.1%) women and 6 (42.9%) men with ages ranging from 67 to 92 (mean age 75.3, SD 8.9) years. Five (35.7%) participants were diagnosed with heart failure, 4 (28.6%) with Parkinson disease, 4 (28.6%) with rheumatic disease, and 1 (7.1%) with stroke.

Deductive Analysis

The data were structured based on the gamification framework by Berglund et al [8], which was partly confirmed by the data from the participants. Regarding the behavioral change techniques, 5 out of 8 could be fully confirmed in the data. The technique "focus on past success" was not used in the design of the exergame, and therefore, participants did not mention it. Participants experienced rewards but not threats as described in "Rewards and threats." "Nonspecific rewards" was also not used as a technique in the exergame, but participants experienced the social aspects of exergaming, and thus, it is described under social support. "Self-monitoring" could be found but was experienced as psychological rather than behavioral and was thus only included under the psychological change techniques.

Behavioral Change Techniques

Overview

These techniques have been identified as crucial to changing the behavioral outcomes of users in real life. Following the framework, behavioral change techniques include goals and planning, reward and threat, feedback and monitoring, focus on past success, nonspecific incentive, social support, nonspecific reward, and self-monitoring [8]. The exergame used multiple gamification elements, which could be mapped to the techniques according to Table 2. The analysis of the interviews revealed additional experiences connected to behavioral techniques. Although these experiences were not linked to specific gamification elements and are not included in the table, they are discussed further in the text.

Table . The behavioral change techniques experienced by people with chronic conditions after playing the Heart Farming walking-based exergame for 3 months and the gamification elements facilitating the techniques based on the deductive analysis using the gamification framework by Berglund et al [8].

Behavioral change techniques as described in the framework	Heart Farming gamification elements
Goals and planning	The daily goal, one of the exergame's <i>score systems</i> , gave participants a goal to reach and plan around. This assisted in structuring daily life and developing a gaming strategy.
Reward and threat	The <i>experience points</i> , <i>reward system</i> , and <i>achievement systems</i> (eg, collecting animals) worked as rewards that motivated participants to walk more. No threats were mentioned by the participants.
Feedback and monitoring	<i>Feedback messages</i> assisted in planning their walk (walking progress) and motivated participants to continue walking to finish their plans (audio feedback).
Focus on past success	Not used as a technique and not mentioned.
Nonspecific incentive	The graphics and farming theme increased motivation.
Social support	The <i>leaderboard</i> was used to compare scores with the top 3 players.
Nonspecific rewards	<i>Social or peer pressure</i> (via the experience of an exergame community) is described under social support.
Self-monitoring	This was experienced as a psychological, not behavioral, technique.

Goals and Planning

Participants described how exergaming assisted in planning and structuring their daily routine. Participants preferred to exergame in the mornings, and they mentioned that they usually have coffee and breakfast directly afterward, as a form of reward for accomplishing their daily goal, which is one of the exergame's reward systems. Participants also combined their everyday activities such as visiting the doctor or doing grocery shopping with exergaming. Sometimes participants played inside the house, especially when the weather did not allow them to play outside.

I use it every time I go to the health care center and when I get a blood test, and I also use it when I go to my children and my grandchildren. [67-year-old man]

They also mentioned the gaming strategy they used while exergaming and the use of the option to autoplay. Participants who used the game emphasized that they chose to prepare their plan for the game at home, so they did not have to interrupt their walking during the journey. The variety of options within the game provided was appreciated, and participants mentioned that they chose to harvest the farm and collect animals. Autoplay was used when participants wanted a simpler option or they did not want to interrupt their walking.

Reward and Threat

Participants mentioned that the digital rewards they received motivated them to walk more. Collecting animals was considered enjoyable, and players chose what actions to take in the exergame according to what was required (eg, harvest 3 fields) to get that animal.

And then I look at what I am missing, for example, do I need to get a new dog, horse or elephant? Then I work towards gaining them. [67-year-old man]

Sometimes participants expressed the need to receive more rewards and money when selling to neighbors and more points

when they walked further (after having achieved their daily goal). They mentioned that receiving more money for selling to neighbors might increase their motivation. They felt that if they wanted to walk more daily after achieving their daily goal, that would lead to more rewards.

Feedback and Monitoring

The feedback provided from the game assisted the participants in monitoring their behavior. Participants mentioned that knowing how much they walked during the day helped them in planning their walk for the rest of the day.

Participants generally liked the sound and noted that the sound effects helped to understand what happened in the game and made the process more interactive. The footsteps and farming sounds helped confirm if the camera was tracking the distance and that the sounds for reaching a goal or the end of a plan acted as both notification and reward.

You heard gravel when you walked and it also because a bit of a whip [referencing the expression "carrot and stick," in Swedish literally "carrot and whip"], and a little carrot as well, to have it there and know that it checks how much you walk. [80-year-old woman]

If participants continued exergaming 3 times more after having reached their daily goal, which was 10 minutes of walking daily, the game reminded them to rest. Participants adhered to this suggestion.

Nonspecific Incentive

Increased physical activity and perceived improved health were considered motivating factors to continue exergaming. Participants mentioned that they noticed an increase in their walking speed, endurance, and the number of daily steps. Some even mentioned that they noticed weight loss because of increased physical activity, and thus, they felt more motivated to exergame.

It was not that I was a big fan of this specific game but measuring the daily physical activity and comparing it with how improved I felt [regarding physical activity level], that was what I liked. [77-year-old man]

Participants highlighted that exergaming not only motivated them to walk more but also increased their appetite to go outside and enjoy nature. Exergaming provided them with the “little push,” as mentioned by a 71-year-old woman, and thus, it gave them a purpose to go out and walk with the game, instead of being physically inactive inside.

Participants liked the graphics used in the game. A 68-year-old woman stated that “it’s nice with these fields and the barns in the middle of it all.” They also mentioned that they liked the option provided by the exergame to do squats instead of walking as an alternative that provided some variation. Participants also thought that the graphics were clear and made the game easier to understand, providing an incentive to continue with exergaming. However, participants also mentioned that the graphics were too childish. A 71-year-old man characterized the exergaming as “a game for children.”

Participants generally liked the farming theme of the game. They referred to a sense of nostalgia while exergaming, and specifically, a 68-year-old woman mentioned “[It] can be fun to have a farm in your imagination.” Participants mentioned that they liked this particular exergame, despite a lack of general interest in games, due to its graphics.

I don’t have any interest in [games] but this, when I saw that auto-walking man out and walking, that I thought was a little fun, so I was immediately hooked on that. [85-year-old man]

Social Support

Even though playing with others was an option within the exergame, some participants were not that interested in this function, while others did not have anyone to play with or did not know how to use this function.

Participants were interested in using the leaderboard as a form of competition by comparing their own scores with the top 3 players displayed on the leaderboard. They also suggested ways to make it easier to see your own score compared to those of the other players, as they did not always make it to the top 3.

Yeah or end up in the top three on the leaderboard. It could maybe be a top ten list. So, you could see roughly where you are—or yeah, if you’re not even in the top ten, then depending on how many individuals are playing, maybe you’d feel like making an effort, or not? [67-year-old man]

Socializing and creating a community around exergaming was a nondigital reward mentioned during the interviews. Some participants who met during the introduction session reported that they took the initiative to organize group lunch meetings to discuss the game and walk together.

I think it has been very fun and therefore I think that the other ladies will also continue exergaming, and then I thought we could maybe go for a walk. We can walk, and after we can go and have a coffee, and the ones who can’t walk that fast can rest more often. [71-year-old woman]

Psychological Change Techniques

Overview

These techniques in the exergame aim to change the psychological outcomes of the users in real life. According to the framework, psychological change techniques include providing information, self-monitoring, and stress management. Even though the exergame was not designed to use the technique to provide information, participants experienced different levels of understanding of the exergame based on the information they received from the research team. Self-monitoring was used by the exergame and was thus mentioned by the participants. Stress management was not used in the design of the exergame; however, participants still mentioned stress and pain relief from getting immersed in the exergame. The gamification elements used in the exergame can be mapped to the techniques according to Table 3.

Table . The psychological change techniques experienced by people with chronic conditions after playing the Heart Farming walking-based exergame for 3 months and the gamification elements facilitating the techniques based on the deductive analysis using the gamification framework by Berglund et al [8].

Psychological change techniques as described in the framework	Heart Farming gamification elements
Provide information	The exergame did not use this technique. However, different levels of familiarity with the exergame were reported as an important factor in the experience of the exergame.
Self-monitoring	<i>Levels of achievements or ranks</i> increased motivation to progress in the exergame.
Stress management	Although stress management was not used as a technique, participants reported experiencing stress relief when exergaming. Some factors (eg, exergaming in public) unintentionally created stress for the participants.

Provide Information

The exergame did not use a technique to provide any health information, nor was this mentioned by the participants.

However, different levels of understanding of the exergame were reported as a factor affecting their exergame experience and are thus presented under this section. Participants with previous experience with games found the exergame to be

intuitive and easy to get into, while participants who were less familiar with technology found it to be complicated and hard to understand. A participant referred to her old age as a possible reason for not understanding the exergame.

Self-Monitoring

Participants found enjoyment in making progress in the game. The different levels they could reach in the game served as motivation and a goal to reach when playing.

I probably like these different levels the best, it has been the most fun to go after them. [77-year-old man]

Participants also liked receiving a trophy for reaching their daily goals. Participants mentioned sometimes reaching their daily goal 2 or 3 times.

Yes, I remember, it has happened a few times maybe that I have gotten many of these flags like when I have walked far now and then, and then I feel “wow, that was many of those flags.” [80-year-old woman]

Stress Management

The exergame did not use a technique to aid in stress management. However, playing the exergame could serve as stress relief, and specifically, an 80-year-old woman mentioned that *“it’s a relaxation at the same time that I need it.”* Participants highlighted the role of exergaming in relieving pain, as immersion in the exergame made it easier to distract from pain.

I absolutely feel like it has only felt good, and then it can also be good if you have a day with a lot of pain, then you also forget a little, when you are like busy with these animals [...] then I actually forget the rest a bit. [77-year-old man]

Participants also reported factors that unintentionally led to increased stress. Sometimes they thought that the game did not accurately measure how far they walked or that the exergame did not react to finger presses. Some worries were also expressed about what other individuals thought when they were exergaming in public.

So, I have to walk like this, and then I have gotten the question “are you recording me?” from individuals nearby. [68-year-old woman]

Discussion

Main Contributions

This study aimed to explore the experiences of behavioral and psychological change techniques used by a walking-based exergame. Principal findings indicated that exergaming provided motivation to walk by using goals and rewards, and participants found value not only by progressing and getting rewarded in the exergame but also through experiences of getting out in nature or finding community with others. Participants found that exergaming assisted them in planning and structuring their daily routines, often playing in the mornings and integrating it with activities, such as visiting the doctor or grocery shopping. Several challenges were mentioned, including holding the phone and handling the camera or playing in public. Participants

appreciated having support from the research team when being introduced to the exergame. Health-related impairments, such as knee problems, influenced the experience of exergaming. However, exergaming could also provide a distraction from pain.

Individuals’ capacities can influence their experience with exergaming, which includes their attitude toward exergaming, exergaming skills, and health-related impairments [27]. This was emphasized by the participants in our study, as they believed that their old age might have affected their capacity for exergaming. Previous qualitative studies indicated that exergaming can be experienced as a fun, engaging, and enjoyable activity for individuals with chronic conditions [13,39,40]. Participants in our study also enjoyed exergaming, although some were not as interested in exergaming and reported limited experiences with games prior to participating in this study. For an exergame to be experienced as an enjoyable activity, the skills and previous knowledge of participants should be evaluated [41]. Participants’ level of knowledge and skills with technology should be taken into consideration, since familiarization with technology can differ greatly for older adults [41]. For example, individuals may worry about their ability to play if they find the technology to be very advanced [14,41,42]. To overcome this challenge, individuals in our study received a structured familiarization with the exergame, including an introductory session and follow-up calls from the members of the research team who aimed to provide sufficient support, instruction, and feedback. Similar findings have been reported in other studies, highlighting the importance of familiarization when introducing exergames to older adults [27].

Social connection is a strong motivator for physical activity, especially in older adults [42-45], and could provide motivation to increase physical activity [41]. The exergame in our study included some support for social interaction within the exergame, where players could send emojis and crops to each other. However, participants in our study did not use the social functions in the exergame due to a lack of interest, difficulty understanding these functions, or not having anyone to play with. A previous study found a positive correlation between the number of messages players send to each other and their physical activity, highlighting the potential of social interaction within an exergame [46].

Participants in our study created social groups outside of the exergame where they walked together using the exergame but did not use the social interaction within the exergame. In another study, they also found that observing each other while exergaming can motivate players to play, even when playing single-player exergames [40]. Culture could also have an impact on how socialization affects exergame opportunities, as individualistic cultures are less likely to feel embarrassed, while older adults in other countries might experience discomfort from feeling judged by others [27]. In our study, there were also participants who stated that they are not interested in playing with others, as well as participants who expressed discomfort from playing the exergame in public, where other individuals could see them. This should also be considered when designing an exergame.

The participants of our study emphasized that exergaming is easy to incorporate into their daily lives, and they could combine daily activities, such as grocery shopping or doctor visits, with exergaming. In addition to walking 10 minutes more each day, study participants also mentioned that they integrated their daily activities with exergaming. Creating new routines is associated with increased adherence to exergaming [13]. To avoid interrupting their daily activities by stopping to play on the phone, participants of our study preferred to complete all their planning before they started walking and to handle all their rewards at the end of the walk. This indicates a preference for separating exercise from gameplay, which other exergames for older adults have utilized before [47].

Having the opportunity to exergame outdoors was considered an asset by the participants in our study. Exergaming motivated participants to go outside and engage with nature, exploring their surroundings instead of remaining inactive at home. However, individuals whose hands were unsteady or shaking while holding the phone could cause inaccurate measurements. Although participants in our study were offered armbands that could hold their phones without blocking the camera, this was not seen as ideal, and participants expressed a desire for another solution that would allow them to keep the phone in their pocket while walking, something that was not possible, as the technology used requires the phone's camera. The alternative solution of using GPS instead introduces the issue of imprecise tracking in indoor environments, especially when walking short distances [16]. As participants used the exergame indoors under certain circumstances, for example, during bad weather, the ability to do so should still be considered when deciding what technology to use for this type of exergame.

Health-related impairments can impact individuals' experience of exergaming [27]. The participants in our study were diagnosed with different chronic conditions (including stroke, chronic heart failure, Parkinson disease, and rheumatic disease) and varied in terms of symptom severity and disease progression. In our study, exergaming was considered an effective distraction from pain. This finding aligns with previous studies that have highlighted the significant role of exergaming in distracting from physical exertion and everyday life [10,27,38]. An RCT that compared brisk walking with exergaming in older adults also indicated that exergaming produced a lower level of perceived exertion, possibly due to distractions provided through exergame graphics or sound effects [47].

The results of our study show that future exergame design for individuals with chronic conditions should prioritize accessibility, adaptability, and meaningful integration into daily life. A systematic review by Cugusi et al [22] shows that exergames combining physical activity with cognitive engagement can enhance physical, cognitive, and mental health outcomes in older adults and people with chronic diseases while also supporting motivation and adherence. Tailored exergaming solutions that can be played at home, with adjustable difficulty levels and clear, intuitive feedback, are particularly relevant for addressing mobility, balance, and long-term engagement needs in clinical populations such as those with heart failure or dementia [30,48]. In line with these findings, it is essential to

assess individuals' skills and prior experience before developing or implementing exergames. Designing exergames that are safe, person-centered, and adaptable to varying functional abilities is crucial for supporting sustained participation among individuals living with chronic conditions [34].

Strengths and Limitations

The strengths and limitations of this study ought to be taken into consideration when interpreting the results. The target sample included individuals with diverse chronic conditions, which differ in terms of symptom severity and illness progression. This provides a deeper understanding of the topic under investigation. In terms of transferability, the results may mirror the Swedish older population with certain chronic conditions, but there may be cultural differences, age differences, and differences if the game were to be used in other chronic conditions. Future studies are required to determine whether condition-specific exergames should be developed.

The use of this gamification framework assisted in structuring the analysis of the data. However, its use might have led to the absence of some important aspects discussed during the interviews. A few parts of the framework were not confirmed by the data, including the focus on past success, nonspecific rewards, threats, and self-monitoring as behavioral techniques. Although this framework underpinned the development of the Heart Farming exergame, it can be considered a broad, overarching model for game design. Therefore, it is understandable that not all the techniques described in this framework are employed in the Heart Farming exergame. It is also important to consider that the interviews were conducted by one of the game developers, which might have introduced potential biases. However, the data analysis was conducted by researchers with diverse academic backgrounds and clinical experience, which may have strengthened the trustworthiness of the findings.

While participants made promising statements regarding the effect the exergame had on their health, this study did not collect or analyze any potential health effects. To draw any conclusions about the actual health benefits of this type of exergame, future studies should include quantitative data on measurable health indicators to analyze how they are affected by exergaming.

Conclusions and Implications

This study contributes to the exergaming literature by examining behavioral and psychological change techniques from the perspective of individuals with chronic conditions. To our knowledge, this is the first qualitative study to apply this gamification framework within this population. Considering the results of this study in combination with previous findings, exergaming can be a promising approach to improving physical activity and achieving behavioral change in individuals with chronic conditions. Considering individuals' needs, previous knowledge of exergaming, interests, and illness progression is critical when developing and using exergames in health care. By increasing understanding of how game elements are experienced and how they influence health-related behavior, the findings of this study may inform the development of exergames that are better tailored to users' needs.

Looking ahead, future studies with more diverse populations and larger sample sizes should focus on the evaluation of the impact of behavior-change techniques used in exergames to maximize the positive effects of exergaming in individuals with

chronic conditions. It would also be valuable to explore whether different behavior-change techniques are required depending on the symptoms of diverse chronic conditions.

Acknowledgments

According to the GAIDeT (Generative AI Delegation Taxonomy; 2025) [49], the following tasks were delegated to generative artificial intelligence (GAI) tools under full human supervision: proofreading and editing. The GAI tool used was Grammarly. Responsibility for the final manuscript lies entirely with the authors. The authors would like to thank everyone who participated in this study.

Funding

This work is supported by the Swedish National Science Council (2020-01109), the Swedish Heart and Lung Foundation (20200190 and 20240456), the Kamprad Foundation (20210074), and the Research Council in Southeast Sweden (FORSS-940933 and FORSS-995300).

Authors' Contributions

Conceptualization: MA, HO, AB, EB, AS, TJ, LK

Methodology: MA, HO, AB, EB, AS, TJ, LK

Data analysis: MA, HO, LK, AB

Writing – original draft: MA, HO

Writing – review and editing: MA, HO, AB, EB, AS, TJ, LK

Supervision: LK, TJ, AS, AB, EB

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide used in the study.

[DOCX File, 25 KB - [games_v14i1e78776_app1.docx](#)]

References

1. Hajat C, Stein E. The global burden of multiple chronic conditions: a narrative review. *Prev Med Rep* 2018 Dec;12:284-293. [doi: [10.1016/j.pmedr.2018.10.008](#)] [Medline: [30406006](#)]
2. Klompstra L, Jaarsma T, Strömberg A. Exergaming to increase the exercise capacity and daily physical activity in heart failure patients: a pilot study. *BMC Geriatr* 2014 Nov 18;14:119. [doi: [10.1186/1471-2318-14-119](#)] [Medline: [25407612](#)]
3. Givon N, Zeilig G, Weingarden H, Rand D. Video-games used in a group setting is feasible and effective to improve indicators of physical activity in individuals with chronic stroke: a randomized controlled trial. *Clin Rehabil* 2016 Apr;30(4):383-392. [doi: [10.1177/0269215515584382](#)] [Medline: [25952586](#)]
4. Ribas CG, Alves da Silva L, Corrêa MR, Teive HG, Valderramas S. Effectiveness of exergaming in improving functional balance, fatigue and quality of life in Parkinson's disease: a pilot randomized controlled trial. *Parkinsonism Relat Disord* 2017 May;38:13-18. [doi: [10.1016/j.parkreldis.2017.02.006](#)] [Medline: [28190675](#)]
5. Duran AT, Ewing Garber C, Cornelius T, Schwartz JE, Diaz KM. Patterns of sedentary behavior in the first month after acute coronary syndrome. *J Am Heart Assoc* 2019 Aug 6;8(15):e011585. [doi: [10.1161/JAHA.118.011585](#)] [Medline: [31364434](#)]
6. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining "gamification". Presented at: MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Sep 28-30, 2011 URL: <https://dl.acm.org/doi/abs/10.1145/2181037.2181040> [accessed 2026-02-28] [doi: [10.1145/2181037.2181040](#)]
7. Al-Rayes S, Al Yaqoub FA, Alfayez A, et al. Gaming elements, applications, and challenges of gamification in healthcare. *Inform Med Unlocked* 2022;31:100974. [doi: [10.1016/j.imu.2022.100974](#)]
8. Berglund A, Jaarsma T, Berglund E, Strömberg A, Klompstra L. Understanding and assessing gamification in digital healthcare interventions for patients with cardiovascular disease. *Eur J Cardiovasc Nurs* 2022 Aug 29;21(6):630-638. [doi: [10.1093/eurjcn/zvac048](#)] [Medline: [35709297](#)]
9. Dörner R, Göbel S, Effelsberg W, Wiemeyer J. *Serious Games: Foundations, Concepts and Practice*: Springer; 2016. [doi: [10.1007/978-3-319-40612-1](#)]

10. Schättin A, Häfliger S, Meyer A, et al. Correction: design and evaluation of user-centered exergames for patients with multiple sclerosis: multilevel usability and feasibility studies. *JMIR Serious Games* 2021 May 13;9(2):e30326. [doi: [10.2196/30326](https://doi.org/10.2196/30326)] [Medline: [33983893](https://pubmed.ncbi.nlm.nih.gov/33983893/)]
11. Jaarsma T, Klompstra L, Ben Gal T, 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. [doi: [10.1002/ejhf.305](https://doi.org/10.1002/ejhf.305)] [Medline: [26139585](https://pubmed.ncbi.nlm.nih.gov/26139585/)]
12. García-Bravo S, Cuesta-Gómez A, Campuzano-Ruiz R, et al. Virtual reality and video games in cardiac rehabilitation programs. A systematic review. *Disabil Rehabil* 2021 Feb;43(4):448-457. [doi: [10.1080/09638288.2019.1631892](https://doi.org/10.1080/09638288.2019.1631892)] [Medline: [31258015](https://pubmed.ncbi.nlm.nih.gov/31258015/)]
13. Klompstra L, Jaarsma T, Mårtensson J, Strömberg A. Exergaming through the eyes of patients with heart failure: a qualitative content analysis study. *Games Health J* 2017 Jun;6(3):152-158. [doi: [10.1089/g4h.2016.0087](https://doi.org/10.1089/g4h.2016.0087)] [Medline: [28422516](https://pubmed.ncbi.nlm.nih.gov/28422516/)]
14. Cacciata MC, Stromberg A, Klompstra L, et al. Facilitators and challenges to exergaming: perspectives of patients with heart failure. *J Cardiovasc Nurs* 2022;37(3):281-288. [doi: [10.1097/JCN.0000000000000825](https://doi.org/10.1097/JCN.0000000000000825)]
15. Ravani E, Ali F, Albuainain S, et al. Exploring the barriers and facilitators experienced by patients with heart failure when using popular exergaming platforms for self-management-a systematic review. *Digit Health* 2024;10:20552076241249823. [doi: [10.1177/20552076241249823](https://doi.org/10.1177/20552076241249823)] [Medline: [39070896](https://pubmed.ncbi.nlm.nih.gov/39070896/)]
16. Berglund A, Klompstra L, Orädd H, et al. The rationale behind the design decisions in an augmented reality mobile eHealth exergame to increase physical activity for inactive older people with heart failure. *JMIR Serious Games* 2024 Aug 21;12:e50066. [doi: [10.2196/50066](https://doi.org/10.2196/50066)] [Medline: [39185820](https://pubmed.ncbi.nlm.nih.gov/39185820/)]
17. Boulos MNK, Yang SP. Exergames for health and fitness: the roles of GPS and geosocial apps. *Int J Health Geogr* 2013 Apr 5;12:18. [doi: [10.1186/1476-072X-12-18](https://doi.org/10.1186/1476-072X-12-18)] [Medline: [23561306](https://pubmed.ncbi.nlm.nih.gov/23561306/)]
18. LeBlanc RG, Jacelon CS. Self-care among older people living with chronic conditions. *Int J Older People Nurs* 2018 Sep;13(3):e12191. [doi: [10.1111/opn.12191](https://doi.org/10.1111/opn.12191)] [Medline: [29573333](https://pubmed.ncbi.nlm.nih.gov/29573333/)]
19. Pokémon GO brings video games outside. American Heart Association. 2016. URL: <https://web.archive.org/web/20160801000000/http://news.heart.org/pokemon-go-brings-video-games-outside/> [accessed 2026-02-28]
20. Serino M, Cordrey K, McLaughlin L, Milanaik RL. Pokémon GO and augmented virtual reality games: a cautionary commentary for parents and pediatricians. *Curr Opin Pediatr* 2016 Oct;28(5):673-677. [doi: [10.1097/MOP.0000000000000409](https://doi.org/10.1097/MOP.0000000000000409)] [Medline: [27479151](https://pubmed.ncbi.nlm.nih.gov/27479151/)]
21. Ayers JW, Leas EC, Dredze M, Allem JP, Grabowski JG, Hill L. Pokémon GO-a new distraction for drivers and pedestrians. *JAMA Intern Med* 2016 Dec 1;176(12):1865-1866. [doi: [10.1001/jamainternmed.2016.6274](https://doi.org/10.1001/jamainternmed.2016.6274)] [Medline: [27635638](https://pubmed.ncbi.nlm.nih.gov/27635638/)]
22. Cugusi L, Prosperini L, Mura G. Exergaming for quality of life in persons living with chronic diseases: a systematic review and meta-analysis. *PM R* 2021 Jul;13(7):756-780. [doi: [10.1002/pmrj.12444](https://doi.org/10.1002/pmrj.12444)] [Medline: [32592238](https://pubmed.ncbi.nlm.nih.gov/32592238/)]
23. Hasselmann V, Oesch P, Fernandez-Luque L, Bachmann S. Are exergames promoting mobility an attractive alternative to conventional self-regulated exercises for elderly people in a rehabilitation setting? Study protocol of a randomized controlled trial. *BMC Geriatr* 2015 Sep 7;15:108. [doi: [10.1186/s12877-015-0106-0](https://doi.org/10.1186/s12877-015-0106-0)] [Medline: [26346751](https://pubmed.ncbi.nlm.nih.gov/26346751/)]
24. Kramer A, Dettmers C, Gruber M. Exergaming with additional postural demands improves balance and gait in patients with multiple sclerosis as much as conventional balance training and leads to high adherence to home-based balance training. *Arch Phys Med Rehabil* 2014 Oct;95(10):1803-1809. [doi: [10.1016/j.apmr.2014.04.020](https://doi.org/10.1016/j.apmr.2014.04.020)] [Medline: [24823959](https://pubmed.ncbi.nlm.nih.gov/24823959/)]
25. Sinclair J, Hingston P, Masek M. Considerations for the design of exergames. Presented at: GRAPHITE '07: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia; Dec 1-4, 2007 URL: <https://dl.acm.org/doi/proceedings/10.1145/1321261> [accessed 2026-02-28] [doi: [10.1145/1321261.1321313](https://doi.org/10.1145/1321261.1321313)]
26. Berglund A, Orädd H. Exploring the psychological effects and physical exertion of using different movement interactions in casual exergames that promote active microbreaks: quasi-experimental study. *JMIR Serious Games* 2024;12:e55905. [doi: [10.2196/55905](https://doi.org/10.2196/55905)] [Medline: [39186776](https://pubmed.ncbi.nlm.nih.gov/39186776/)]
27. Ning H, Jiang D, Du Y, et al. Older adults' experiences of implementing exergaming programs: a systematic review and qualitative meta-synthesis. *Age Ageing* 2022 Dec 5;51(12):afac251. [doi: [10.1093/ageing/afac251](https://doi.org/10.1093/ageing/afac251)] [Medline: [36571772](https://pubmed.ncbi.nlm.nih.gov/36571772/)]
28. Bravo-Aparicio J, Domínguez-López P, Díaz-González C, Martín-Caro Álvarez D, Álvarez DMC, Beltran-Alacreu H. Effects of different exercise doses through an augmented reality exergame in older adults: a pilot randomized controlled trial. *Appl Sci* 2024;14(22):10592. [doi: [10.3390/app142210592](https://doi.org/10.3390/app142210592)]
29. Nuic D, van de Weijer S, Cherif S, et al. Home-based exergaming to treat gait and balance disorders in patients with Parkinson's disease: a phase II randomized controlled trial. *Eur J Neurol* 2024 Jan;31(1):e16055. [doi: [10.1111/ene.16055](https://doi.org/10.1111/ene.16055)] [Medline: [37691341](https://pubmed.ncbi.nlm.nih.gov/37691341/)]
30. Klompstra L, Hägglund E, Jaarsma T, Kato NP, Strömberg A. Effects of exergaming and yoga on exercise capacity and physical and mental health in heart failure patients: a randomized sub-study. *Eur J Cardiovasc Nurs* 2025 Apr 11;24(3):389-398. [doi: [10.1093/eurjcn/zvae155](https://doi.org/10.1093/eurjcn/zvae155)]
31. Huang X, Xiang X, Liu Y, Wang Z, Jiang Z, Huang L. The use of gamification in the self-management of patients with chronic diseases: scoping review. *JMIR Serious Games* 2023 Dec 22;11:e39019. [doi: [10.2196/39019](https://doi.org/10.2196/39019)] [Medline: [38133907](https://pubmed.ncbi.nlm.nih.gov/38133907/)]

Edited by S Brini; submitted 09.Jun.2025; peer-reviewed by A Gazis, S Göbel; accepted 14.Feb.2026; published 27.Mar.2026.

Please cite as:

Antoniadou M, Orädd H, Berglund A, Berglund E, Strömberg A, Jaarsma T, Klompstra L

User Experiences of Behavioral and Psychological Change Techniques in a Walking-Based Mobile Exergame: Cross-Sectional Qualitative Study

JMIR Serious Games 2026;14:e78776

URL: <https://games.jmir.org/2026/1/e78776>

doi: [10.2196/78776](https://doi.org/10.2196/78776)

© Marianna Antoniadou, Helena Orädd, Aseel Berglund, Erik Berglund, Anna Strömberg, Tiny Jaarsma, Leonie Klompstra. Originally published in JMIR Serious Games (<https://games.jmir.org>), 27.Mar.2026. 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.

Original Paper

Gamified Physical Education and Cognitive Performance Among Chinese Secondary School Students: Cross-Sectional Moderation Mediation Study

Jingtong Zhang¹, PhD; Sae-Sook Oh¹, Prof Dr; Yaru Xu², MSc

¹Department of Physical Education, Kyonggi University, Suwon-si, Republic of Korea

²School of Health Sciences, Universiti Sains Malaysia, Kubang Kerian, Kelantan, Malaysia

Corresponding Author:

Yaru Xu, MSc

School of Health Sciences

Universiti Sains Malaysia

Exercise and Sports Science, School of Health Sciences

Universiti Sains Malaysia, Health Campus

Kubang Kerian, Kelantan, 16150

Malaysia

Phone: 86 17763575813

Email: xuyarustc@163.com

Abstract

Background: Educators are exploring new methods to educate beyond the classroom as global concerns about students' cognitive, emotional, and social well-being grow. Physical education (PE) has been demonstrated to boost cognitive and psychological outcomes in several studies. Most research has neglected the benefits of gamification and artificial intelligence (AI)-based feedback in PE, focusing instead on conventional PE formats. The impacts of technologically enhanced PE settings on students' cognitive performance through feedback and reward mechanisms remain understudied.

Objective: This study aimed to investigate how intrinsic motivation and AI-based feedback moderated the effects of gamified PE on students' cognitive performance.

Methods: The study used a cross-sectional design. In Beijing, Shanghai, Chengdu, and Guangzhou, a total of 1029 public high school students completed a standardized questionnaire. Students in secondary school (male: $n=490$, 47.6% and female: $n=539$, 52.4%) aged 10-18 years, were recruited from urban, suburban, and rural locales. Participants were sourced from public, private, and semigovernment schools, reflecting a range of academic achievement levels and access to technology. Students participating in standard PE sessions were included, whereas those with medical conditions that restricted physical exercise were excluded. Data were gathered via standardized questionnaires during designated PE sessions. Gamified PE, cognitive performance, intrinsic motivation, teacher support, collaboration, and AI feedback were examined using standardized instruments. Trained facilitators helped younger participants understand and follow ethical norms. The study used maximum likelihood estimation for structural equation modeling. Bootstrapping was used to analyze mediation and moderation effects at a 5% significance level ($\alpha=.05$).

Results: According to structural equation modeling, gamified PE highly predicts cognitive performance ($\beta=.34$; $P<.001$). Other significant factors were teacher support ($\beta=.31$; $P<.001$), physical exercise enjoyment ($\beta=.28$; $P<.001$), and teamwork ($\beta=.26$; $P<.001$). AI-based feedback strengthened the link between gamified PE and cognitive performance under moderation analysis ($\beta=.18$; $P<.001$). Mediation analysis indicated that intrinsic motivation partially mediated the relationship, resulting in a significant indirect effect ($\beta=.21$, 95% CI 0.12-0.31; SE=0.05).

Conclusions: This research integrates gamified PE with AI-based feedback mechanisms to evaluate students' cognitive outcomes, a domain that has been rarely investigated experimentally. This study highlights the combined effect of intrinsic motivation and AI-generated feedback in a technology-enhanced PE context, in contrast to previous research that primarily focuses on traditional PE methods or isolated gamification elements. The findings enhance the field by demonstrating that student-centered, feedback-rich PE environments may improve cognitive abilities through social interaction, enjoyment, and instructor support. AI-assisted, gamified PE programs may enhance learning outcomes and academic performance among secondary school students.

(JMIR Serious Games 2026;14:e81086) doi:[10.2196/81086](https://doi.org/10.2196/81086)

KEYWORDS

gamified physical education; cognitive performance; intrinsic motivation; AI-based feedback; teacher support; peer collaboration; China; artificial intelligence

Introduction

Background of Research

In the 21st century, multimodal techniques that integrate physical, cognitive, emotional, and technical components have become more important in education [1]. Contemporary research suggests that physical and intellectual education must be integrated, even though traditional schools generally divide them [2,3]. Physical education (PE) can change children's lives by enhancing their emotional, cognitive, and academic performance [4]. Concerns about student inactivity and cognitive overload have led to the development of gamified PE, an innovative teaching approach. This strategy uses game-design principles to make exercise more enjoyable, goal-oriented, and cognitively demanding [5]. Gaming methods combined with physical exercise have been shown to enhance motivation, teamwork, and self-control—essential for both mental and physical well-being [6].

Gamification incorporates features such as points, competition, prizes, badges, levels, and leaderboards into nongame contexts to motivate and engage users. Gamification, which originated in marketing and business, is now being used in classrooms to engage students [7]. This type of PE transforms static exercises into game-like cooperative learning settings where students work together to achieve goals. Research suggests that gamified methods improve students' engagement, enjoyment, and academic success [8]. Teenagers often find PE uninteresting or irrelevant; therefore, these strategies have been effective in capturing their attention. Gamified PE promotes active participation and cognitive engagement by allowing students to make choices, solve problems, and reflect on outcomes in real time [9].

Cognitive performance, encompassing attention, memory, reasoning, and problem-solving, has been extensively studied in educational psychology about physical well-being [10,11]. Regular exercise improves working memory, executive function, and information processing speed. Neurobiological mechanisms, including increased cerebral blood flow, reduced cortisol levels, and the release of brain-derived neurotrophic factor, may explain this connection [12]. Gamified PE may directly stimulate cognitive brain networks through rule-following, decision-making, instruction, memory, and strategy adjustment [13]. China's rapidly changing educational system is a fascinating environment for studying these dynamics. China has put student well-being alongside academic success in recent years to address excessive academic pressure, mental health issues, and physical inactivity in school-age children and adolescents [14]. National educational reforms aim to provide students with a "quality education" that nurtures their emotional, physical, and cognitive potential [15]. Chinese schools are experimenting with gamified teaching, notably in PE, to achieve this goal [16]. Despite well-defined policy aims, empirical research on the implementation processes and outcomes of these

innovations is lacking. Understanding the impact of gamified PE on cognitive development and academic achievement in Chinese children requires considering psychological factors, such as motivation, and external factors, such as technology use.

Intrinsic motivation links gamified PE to cognitive performance. This is the desire to do something because it is fun, fascinating, or valuable. Ryan and Deci's [17] self-determination theory states that intrinsic motivation is driven by autonomy, competence, and relatedness. Gamified situations provide autonomy, competence, and relatedness. Gamified aspects in PE may help students understand and connect with physical exercise by making repetitive tasks more interesting [18]. Intrinsic motivation enables students to self-regulate their participation, thereby improving cognitive functioning and task performance [19]. The mediating role of intrinsic motivation is crucial in understanding how gamified PE may impact cognitive outcomes. Gamified PE may be enhanced or mitigated by technology, particularly artificial intelligence (AI)-powered feedback systems. Many progressive schools worldwide use smart wearables, mobility trackers, and digital dashboards, with AI being increasingly integrated [20,21]. These technologies enhance training efficiency by offering adaptive learning environments, personalized feedback, and real-time performance monitoring. AI may transform PE by enabling real-time monitoring of students' and teachers' efficiency, speed, endurance, and progress [22]. By acknowledging effort, fostering reflection, and driving self-improvement, such technologies may increase student engagement when gamified. Student receptivity to technology, teacher implementation skills, and institutional resources can all impact their effectiveness. Thus, AI-based feedback may moderate the relationship between gamified PE and cognitive performance sufficiently to change its direction.

In addition to these core concepts, contextual considerations should also be addressed. In China, a collectivist culture that values group cohesiveness and cooperative learning, peer collaboration is important. Gamified PE involves collaborative, communicative, and goal-oriented teamwork [23]. This may improve relationships and higher-order thinking via collaborative problem-solving and strategy building. Teacher support also affects student engagement in gamified contexts. When teachers model healthy exercise habits, provide constructive feedback, and encourage autonomy, they may motivate and engage students. Instructor behavior and gamified content delivery are crucial factors in assessing the effectiveness of an intervention [24]. Another key component in China's educational background is the variety of school settings, from public to private and urban to rural. The accessibility of gamification and AI technologies may vary. In Shanghai and Beijing, schools have smart classrooms and instructors who use technology to make learning an engaging experience [25]. Rural or low-resource schools often struggle with outdated technology, a shortage of qualified personnel, and resistance to innovative

teaching methods. Demographics and institutional characteristics may help explain differences in student cognitive performance. Confucian values, including diligence, self-control, and academic achievement, also shape Chinese educational perspectives [26]. These principles influence students' perceptions of gamification and other innovative teaching methods. Gamified PE may be invigorating for some students but too limited for others. Thus, cultural perception is a subtle yet powerful factor that may influence new teaching approaches [27]. To achieve cognitive and motivational advantages, gamified PE may need culturally relevant values and communication strategies.

This study should be viewed from the perspective of modern educational research, which increasingly promotes interdisciplinary collaboration. Motivation theory, PE, educational technology, and cognitive science must collaborate to understand student learning. Gamified PE combines movement studies, behavioral psychology, and technology. These cross-domain interactions are seldom studied, particularly in China's sociocultural context. Gamification in academic fields such as mathematics or language learning has garnered the greatest attention [28], whereas most research has focused on the cognitive advantages of physical exercise [29]. Technical, motivational, and cognitive studies on gamified PE as a core intervention are lacking.

This study is theoretically and practically important for Chinese educational innovation and student growth. Gamifying PE programs can improve children's health and engagement, a trend that is becoming increasingly significant in schools worldwide [30]. Gamified PE may increase physical activity and cognitive function in China, where academic achievement often takes precedence over emotional and physical growth [31]. Studying the mediating role of intrinsic motivation reveals the psychological underpinnings behind student learning and engagement. While educational technology shapes learning settings, AI-based feedback serves as a moderating component that addresses this tendency. This study's results contribute to the global discussion on educational reform, technology-enhanced learning, and student motivation, as it is one of the first to systematically analyze these factors within the Chinese secondary school system. New technologies help school administrators, policymakers, and educators develop more engaging, productive, and intellectually stimulating PE programs.

Today's educational studies must grasp how PE, cognitive development, and motivational psychology interact, particularly in rapidly modernizing nations such as China. Given the growing use of gamified methods and AI in education, research on their influence on key learning outcomes is urgently needed. Over the past 20 years, research on the cognitive benefits of exercise has evolved. Still, the impact of 2 modern pedagogical innovations, namely gamification and AI-based feedback, on this area remains unknown. This study emphasizes gamified PE, intrinsic motivation, AI-based feedback, and cognitive performance. The 3 subsections of the review conclude with a hypothesis to guide empirical research.

Gamified PE and Physical Activity Enjoyment With Cognitive Performance

Several studies have demonstrated that gamified PE enhances children's learning and brain development [30,32]. Rule-based, competitive, and strategic games boost students' attention, engagement, and decision-making in gamified PE [33]. Gamified learning activities have been shown to improve task completion and working memory. Gamified physical activities compel students to solve problems, devise strategies, and assess their progress in real time, which promotes higher-order thinking. These benefits stem from PE games and competitions that keep children moving and engaged in critical-thinking activities [34]. Gamified PE lessons improved content retention and self-regulated learning. It is widely recognized that exercise enhances cognitive function. Students who like exercising are more likely to persist with it. This aids brain growth and executive function [35,36]. A meta-analysis found that children who were happier while exercising performed better on cognitive flexibility and memory tests. Engaging and pleasant activities promote academic learning, critical dual-task performance, and prefrontal brain activity. Despite a lack of scientific data, gamification is gaining popularity in China as a teaching method [37,38]. Gamified PE treatments in Shanghai, China, have improved students' memory and attention. Gamified PE increased mental rotation exercises and digit span recall. These results support the premise that engaging and interesting physical exercise promotes both physical and psychological health. Gamification's novelty must be managed carefully to keep users interested and minimize cognitive fatigue [39,40].

Peer Collaboration and Teacher Support to Cognitive Performance

Working together in class has long been shown to boost students' social and cognitive development. Collaborative learning plays a key role in intellectual growth. Teamwork and collaboration in PE foster effective communication, critical analysis, and perspective-taking [41,42]. Students who worked together on physical activities performed better on reflective thinking and adaptation exams. Peer contact in PE improves strategic thinking and decision-making, particularly when participants must collaborate to solve challenges or reach a consensus [43,44]. Coordinated peer collaboration in PE enhances cognitive engagement. This was especially evident when students discussed approaches and reflected on their outcomes. Cognitive benefits have been observed in teaching games for understanding models, which rely heavily on peer interactions [45,46]. These models improve foresight, planning, and problem-solving. Collectivist principles that promote group cohesiveness and shared accomplishment make peer collaboration meaningful in Chinese education. Chinese students who took PE programs with a partner performed better on executive functioning assessments. Peer collaboration and teacher support help influence students' cognitive involvement [47]. Teachers who foster autonomy and provide clear feedback help pupils build intrinsic motivation and learning perseverance. PE teachers' primary duties include engaging students and making the classroom exciting. Scaffolding, encouragement, and feedback improve cognitive performance [48,49]. Physical ability and academic success were higher in Chinese students

who were more encouraged by their PE teachers. Formative assessments and motivational signals reduced cognitive load and increased self-reflection in PE students [50]. Gamified PE highlighted the need for psychologically safe spaces where children can experiment, fail, and learn, which is essential for cognitive development [51]. Instructors remain crucial to PE courses that use AI to support students in understanding digital feedback and to foster metacognition [52]. These findings suggest that students need to collaborate and have their teachers' support to increase cognitive performance, particularly in dynamic and game-based learning environments.

Mediating Role of Intrinsic Motivation and Moderating Effect of AI-Based Feedback

Intrinsic motivation, which is the drive to do something for its own sake, has been linked to scholastic and cognitive success. Gamified learning settings make students feel more connected, competent, and independent [53]. Motivation to study enhances students' use of deeper cognitive techniques, sustained attention, and effort. Genuinely driven PE students performed better and persevered longer in motor learning tasks. Enjoyment and perceived competence during PE influenced the relationship between PE involvement and academic success [54,55]. Inner motivation mediated the relationship between physical activity and math performance in Chinese middle schools. A recent study suggests that gamified PE motivates students to enhance their executive functioning and facilitate learning transfer [56,57]. These findings support the cognitive benefits of gamified PE, including increased intrinsic motivation, perceived competence, and a sense of control. In contrast, AI has enhanced PE by providing students with real-time performance statistics. Smart wristbands equipped with AI-powered motion sensors facilitate reflective thinking and personalized learning. AI-based feedback enhanced students' metacognitive awareness and adaptive learning strategies [58,59]. This input may reinforce learning loops, give immediate incentives, and drive persistent participation in gamified contexts. Students who received AI-enhanced PE courses performed better on spatial thinking and problem-solving examinations than those who received conventional instructor feedback. AI's moderating influence is not always positive. Students may become excessively dependent on external validation or overwhelmed by AI technology, depending on deployment and preparation [60,61]. Thus, its effectiveness may vary depending on the situation. AI-enhanced PE pilot programs in China have shown promising outcomes, with improved engagement and physical literacy [62]; however, concerns remain regarding unequal access and teacher training. When paired with AI-based feedback, gamified PE may improve cognitive function, particularly depending on the quality of feedback, engagement, and customization.

Research Gaps and Contribution of the Study

There is growing evidence that PE is important for cognitive development; however, the impact of gamified PE interventions on cognitive outcomes remains unclear, particularly in the Chinese educational and cultural context. Exercise has been shown to improve attention, memory, and executive functioning in several studies [63,64]. These studies have primarily examined aerobic or endurance activities, rather than PE

programs that use games to enhance these areas. Gkintoni et al [65] found that motor skill memorization in conventional PE improves mental agility, while gamified activities that require flexibility, strategy, and decision-making in the present engage students' brains more effectively.

Another topic with limited research is the impact of PE on cognitive functioning and its underlying mechanisms. Self-determination theory has been applied in motivation research; however, few studies have examined intrinsic motivation as a moderator in gamified PE [66,67]. Sañudo et al [68] observed that gamification boosts physical activity, but their models do not account for the psychological relationship between intrinsic motivation, internal pleasure, and mental performance. To develop educational interventions that are both physically and psychologically beneficial, it is essential to understand the driving factors of gamified learning.

Research on the cognitive effects of gamified PE using AI-based feedback is seldom conducted. Although AI is becoming increasingly widespread in developing or transitional education systems, such as China's, academics have given less attention to its applications in PE. AI research in education has focused on academic learning platforms or disregarded pedagogy. Zha et al [69] study shows that AI-enhanced feedback tools may assist students in learning PE technical skills, but they do not link them to cognitive progress. In addition, there is limited knowledge about how technological tools affect motivational psychology. AI feedback in gamified PE has not been studied to determine whether it enhances self-awareness and metacognitive processing or merely motivates.

The literature on physically interactive, game-based learning environments does not address multilevel classroom dynamics, including student-teacher collaboration and instructor support. The social aspect of PE is often overlooked in empirical models that combine gamification and cognitive findings. Iglesias and Fernandez-Rio [70] conducted a comparative study. Students may be more engaged in collaborative PE. The research does not examine how these interactions increase cognition or how teacher scaffolding affects it. Comprehensive frameworks must incorporate instructional strategies, motivating factors, technological upgrades, and student interaction patterns, which affect cognitive performance.

This study examines how gamified PE, physical activity enjoyment, peer cooperation, and teacher support impact students' cognitive performance in China's unique educational system, aiming to fill these complex gaps. This study integrates pedagogical and psychological perspectives to better understand how innovative PE methods can enhance cognitive performance. The research incorporates both technology and physical exercise, unlike previous investigations. The study uses intrinsic motivation as a mediating component to enhance the theoretical understanding of internal motivational processes in gamified scenarios. It also examines how AI-based feedback moderates the effects of digital technology on learning in interactive, collaborative classrooms. The study suggests that secondary school PE programs facilitate students' cognitive and developmental progress by providing data to inform curriculum development, teacher professional development, and the

effective use of educational technology. It can help Chinese and other school administrators develop balanced, tech-integrated, and student-centered PE curricula to achieve 21st-century learning goals.

Research has linked gamified PE with an intrinsic desire to exercise and improved cognitive function to form the following research hypotheses:

- H1: Gamified physical education and physical activity enjoyment have a significant positive effect on students' cognitive performance.
- H2: Peer collaboration and teacher support have a significant positive effect on students' cognitive performance.
- H3: Intrinsic motivation mediates, and AI-based feedback moderates, the relationship between gamified physical education and cognitive performance.

Methods

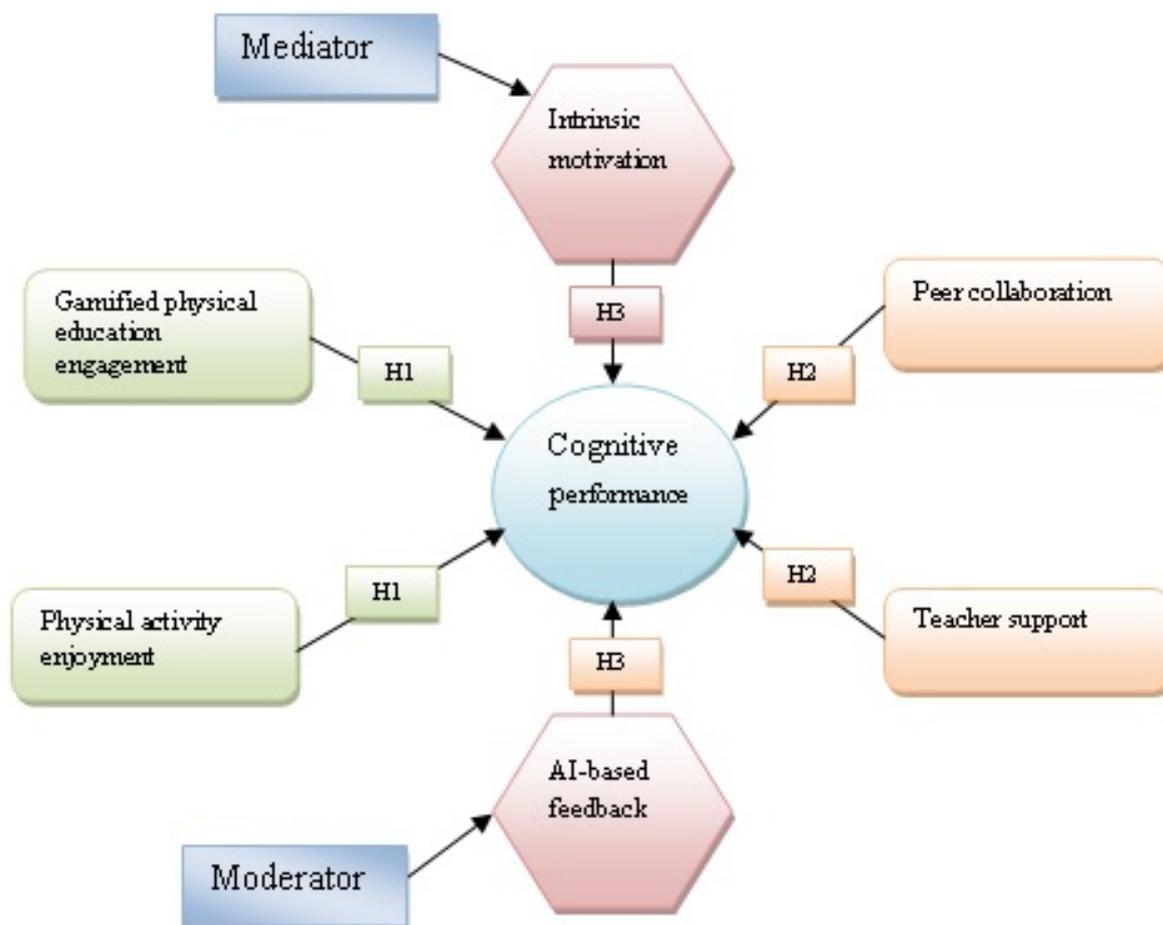
Theoretical Framework

This study is based on cognitive load theory (CLT) and self-determination theory (SDT), with contextual support from Constructivist Learning Theory. Deci and Ryan's [71]

self-determination theory explains how students' sense of enjoyment in connectivity, autonomy, and competence motivates them to pursue diverse activities. SDT is helpful for this study because it shows how intrinsic motivation moderates the relationship between gamified PE and cognitive function. Gamification elements, such as points, challenges, levels, and social engagement, help PE programs fulfill students' psychological needs and encourage enthusiastic participation. As gamified PE becomes more engaging and personalized, students' cognitive engagement, memory, and problem-solving abilities are expected to improve [72]. CLT suggests that gamification and other instructional methods may increase working memory and decrease cognitive load, particularly when combined with real-time AI-based feedback [73]. Such feedback may help students focus, repair mistakes quickly, and learn more via adaptive responses by strengthening the gamified PE-cognition link. According to Constructivist Learning Theory, teachers can direct students' exploration, social interaction, and feedback, while students can also help themselves through peer collaboration and instructor support [74]. Teamwork and teacher facilitation are highly effective in promoting engagement and the application of cognitive skills in PE.

These theoretical frameworks create the study's expected variable relationships (refer to Figure 1).

Figure 1. Conceptual framework showing hypothesized relationships among explanatory variables, mediator, moderator, and dependent variable (cognitive performance) with hypotheses H1–H3. Data from 1029 high school students in Beijing, Shanghai, Chengdu, and Guangzhou, 2025. Source: author's work. AI: artificial intelligence.



The first hypothesis (H1) states that gamified PE and physical activity enjoyment will enhance students' cognitive performance. Physically active classrooms boost students' physical health, mental clarity, memory, and problem-solving skills. H2, which builds on H1, suggests that teacher and peer interaction enhances students' cognitive performance by focusing on the social and pedagogical aspects of learning. These factors make the classroom a friendly yet challenging environment for students to learn through hands-on experiences, collaborate, and receive constructive feedback. H3, the technological and psychological components, reveals that intrinsic motivation mediates the relationship between gamified PE and cognitive performance and that AI-based feedback moderates this relationship. Intrinsically motivated students are more likely to focus, continue the course, and actively engage in learning using AI feedback systems, which improves outcomes. This paradigm encompasses motivation, instruction, technology, and cognition, and when combined, these elements demonstrate how gamified PE can support learning in today's classrooms.

Population, Sample, and Research Design

The study investigated the influence of intrinsic motivation and AI-based feedback on the relationship between gamified PE and students' cognitive performance in China's unique cultural and educational context. The research includes secondary school PE students from various regions in China. The sample included urban, suburban, and rural schools from Beijing, Shanghai, Chengdu, and Guangzhou municipalities. These regions were selected because they participate in national education innovation projects and have various degrees of classroom technology integration. The study used quantitative, cross-sectional surveys and structured questionnaires to obtain participant data. The 1029 completed and valid responses from 1175 surveys issued to 10- to 18-year-olds yielded an 87.6% response rate. Based on Cochran sample size calculation approach, which accounts for a 95% CI and a 3% margin of error, a sample of desired respondents was required. Multivariate statistical analysis, including structural equation modeling (SEM), can be reliably performed with the given sample.

Demographic Profile

The study collected demographic data to understand the participants' contextual profiles and ensure diversity across crucial parameters. High, moderate, and low achievers were categorized based on their academic achievement. The demographic component of the questionnaire inquired about school type (public, private, or semigovernment), location (urban, suburban, or rural), and access to technology (high, moderate, or low). The study cross-tabulated these factors to ensure consistency among student demographics and institutional settings.

Measurement Scales

This survey included key factors from theoretical and empirical investigations (refer to Table 1). Cognitive performance was assessed using a Cognitive Functioning Scale item derived from the Woodcock-Johnson Tests. The item includes: "My memory has improved due to gamified PE." A cognitive development study has validated this scale [75]. The first independent variable, gamified physical education engagement, was assessed using the game-based learning engagement scale. The item was "I feel more engaged in gamified PE," followed by Hamari et al [76]. The second independent variable, physical activity enjoyment, was assessed using the physical activity enjoyment scale with a representative item, "I feel energetic after physical activity," as suggested by Kendzierski and DeCarlo [77]. The third independent variable, examined by the collaborative learning scale, was "Team-based PE helps me learn more," as used by Laal and Ghodsi [78] for instrumentation. Belmont et al [79] used the scale and representative items, including "My teacher helps me reflect on progress," to score instructor support, the fourth independent variable, on the Teacher as Social Context Questionnaire–Autonomy Support subscale. The mediating variable, intrinsic motivation, was operationalized using the intrinsic motivation inventory-interest and enjoyment subscale. One key item is "I feel accomplished after PE." Deci and Ryan's [71] theoretical model and psychometric validation by McAuley et al [80] provided support for this instrument. AI-based feedback was measured using a sample item from the smart education technology-enhanced feedback scale, supporting Ifenthaler et al [81] and Caspari-Sadeghi et al [82]. The study used a 5-point Likert scale for all items, with 1 indicating "strongly disagree" and 5 indicating "strongly agree."

Table 1. Measurement scales and sample items for studied variables, based on responses from 1029 high school students in 4 Chinese cities, 2025.

Constructs	Measurement scale	Sample item from questionnaire	Literature support
Cognitive performance (DV ^a)	Cognitive Functioning Scale (adapted from Woodcock-Johnson Tests)	I can concentrate better during lessons after gamified PE ^b .	[75]
Gamified physical education engagement (IV1 ^c)	Game-Based Learning Engagement Scale	I enjoy gamified PE classes.	[76]
Physical activity enjoyment (IV2 ^d)	Physical Activity Enjoyment Scale (PACES)	Physical activities are fun for me.	[77]
Peer collaboration (IV3 ^e)	Collaborative Learning Scale	I work better in teams during PE.	[78]
Teacher support (IV4 ^f)	Teacher as Social Context Questionnaire–Autonomy Support subscale	My PE teacher gives helpful feedback.	[79]
Intrinsic motivation (mediator)	Intrinsic Motivation Inventory (IMI): interest and enjoyment subscale	I participate in PE because I enjoy it.	[71,80]
^g AI-based feedback (moderator)	Technology-Enhanced Feedback Scale	AI-based feedback helps me improve my physical skills.	[81,82]

^aDV: dependent variable.

^bPE: physical education.

^cIV1: independent variable 1.

^dIV2: independent variable 2.

^eIV3: independent variable 3.

^fIV4: independent variable 4.

^gAI: artificial intelligence.

Pilot Study and Diagnostic Tests

A pilot study evaluated the validity, reliability, and clarity of content and timing of the research instrument with 50 students from similar backgrounds in the different groups of research participants, aged 10-18 years. After conducting reliability analysis using Cronbach's alpha, all questionnaire components had coefficients above 0.70, indicating strong internal consistency. The dimensionality of each scale was assessed using exploratory factor analysis, and construct validity, including convergent and discriminant validity, was examined using confirmatory factor analysis (CFA) with using SPSS AMOS software (IBM Corp). The scale's average variance extracted (AVE) and composite reliability (CR) values were within the acceptable range for all structures, proving its validity. Variance inflation factor multicollinearity tests verified the dataset's suitability for regression-based modeling. These tests showed that the independent variables were not highly correlated.

Reporting Standards

This study followed the APA Journal Article Reporting Standards for Studies Using Structural Equation Modeling to provide comprehensive documentation of model construction, estimation methods, model fit indices, and mediation and moderation studies [83].

SEM was used to test the study's hypotheses (H1, H2, and H3) and to estimate the direct and indirect correlations among variables for further statistical analysis. Furthermore, a moderated-mediation analysis was also performed for robust inferences. The study applied bootstrapping with 5000 resamples to establish the relevance of the indirect effects; the study accordingly performed the analysis. An interaction term was

tested to assess whether AI-based feedback moderates the connection between gamified PE and cognitive performance. The relationship's intensity and direction were considered. Model fit indices, including comparative fit index (CFI), Tucker-Lewis Index (TLI), root-mean-square error of approximation (RMSEA), and standardized root-mean-square residual, were within acceptable bounds, demonstrating model robustness. The relationships were exhaustively analyzed using SPSS Statistics version 26 (IBM Corp) and AMOS version 24 (IBM Corp) for all statistical methods.

Finally, this research relies on reliable statistical methods, measurement instruments, and rigorous sampling. By incorporating both technical and psychological components into educational interventions, such as gamified PE, Chinese secondary school students can gain a multifaceted understanding of how modern pedagogical strategies impact cognitive development. The stratified sample, high response rate, and advanced SEM methodologies make the findings reliable and accessible to a large audience.

Ethical Considerations

This study was approved by the Research Ethics Committee of Kyonggi University, Suwon-si, Republic of Korea (Institutional Review Board no 12416). All procedures were conducted in accordance with institutional and national ethical guidelines. Written informed consent was obtained from all participants before data collection. For participants younger than 18 years, written and oral consent was obtained from their parents or legal guardians. Trained facilitators assisted younger participants to ensure they clearly understood the survey items and participated voluntarily. All data were anonymized at the time of collection; no personal identifiers were recorded, and participants were

assigned unique numeric codes for data handling and analysis. No images or other identifying materials of individual participants were collected or included in the paper. Therefore, no identifiable participant information is disclosed, and no additional consent was required for images.

Results

Stratified random sampling was used to recruit students from various schools across the 3 provinces, categorized by school type and geographic location. The study initially categorized schools by geography. Each stratum's students were randomly selected. The study carefully balanced gender and grade levels for equitable sampling. PE instructors distributed questionnaires during class time with the support of administrators. Participation was voluntary, and students were given anonymity and confidentiality, which enhanced the legitimacy of the responses. School coordinators and field supervisors collaborated to gather data using standardized instruments over 2 months. Given the wide age range of participants, methodological and ethical adjustments were implemented for children in the 10- to 12-year age group, who may experience developmental difficulties in recalling intricate survey questionnaires. For this age group, the questionnaire was read aloud by trained facilitators in a structured manner. The facilitators read questions to the students individually, provided simplified explanations, and ensured that the students responded voluntarily without asking leading questions. This approach ensured the validity, accuracy, and adherence to ethical guidelines when using young respondents in psychological and educational research. The 13- to 18-year-old respondents were able to complete the questionnaire independently.

Table 2 provides the demographics of respondents to illuminate the study's sample composition and contextualize the results. The sample comprised 1029 students from various educational institutions throughout China, ensuring a broad representation of student opinions and learning conditions. Male students comprised 47.60% (n=490) of the sample, while female students

comprised 52.40% (n=539). The findings are more generalizable due to a gender-diverse participant pool and relatively equal distribution. Three age groups of students were 10-12, 13-15, and 16-18 years. The students included 438 (42.60%) aged 13-15 years, 310 (30.10%) aged 10-12 years, and 281 (27.30%) aged 16-18 years. This age group encompasses a broad range of developmental periods in middle and upper elementary and secondary school, making it crucial for gamified learning and cognitive growth. The survey found that 479 (46.60%) students assessed their academic performance as moderate, 305 (29.67%) as excellent, and 245 (23.83%) as low achievers. These distributions show how students from diverse academic backgrounds use gamified PE and AI-based feedback. Public schools educated 428 (41.60%) students, private schools 311 (30.20%), and semigovernment schools 290 (28.20%). This indicates a balanced representation across institutional types, as educational systems and funding may affect the integration of gamified PE and AI. The respondents were drawn from a variety of areas, including urban (366/1029, 35.60%), suburban (360/1029, 35%), and rural (303/1029, 29.40%) locations. Geographic diversity is considered when assessing the impact of educational infrastructure and technological accessibility on cognitive performance outcomes. As a result, the study also reveals regional variations in teaching and educational resources. Finally, 400 (38.90%) students reported high access to digital tools and platforms, 451 (43.90%) reported moderate access, and 178 (17.30%) reported limited access, a significant variable given this study's focus on AI-based feedback. These results suggest that students' technological readiness may influence or moderate the effectiveness of gamified learning approaches. Given the relatively high percentages of students with moderate to high technology access, most students appeared to have participated meaningfully using digital and AI-driven teaching resources.

Table 3 provides descriptive statistics, validity indicators, and internal reliability and stability for each of the study's major latent variables, assessing the stability and internal consistency of the structural model's constructs.

Table 2. Demographic characteristics of 1029 high school student respondents across Beijing, Shanghai, Chengdu, and Guangzhou, 2025.

Demographic variable and category	Value, n (%)
Sex	
Male	490 (47.60)
Female	539 (52.40)
Age group	
10-12 years	310 (30.10)
13-15 years	438 (42.60)
16-18 years	281 (27.30)
Academic performance	
High achiever	305 (29.60)
Moderate achiever	479 (46.60)
Low achiever	245 (23.80)
School type	
Public	428 (41.60)
Private	311 (30.20)
Semigovernment	290 (28.20)
Geographic location	
Urban	366 (35.60)
Suburban	360 (35)
Rural	303 (29.40)
Technology access level	
High	400 (38.90)
Moderate	451 (43.80)
Low	178 (17.30)

Table 3. Descriptive statistics, reliability, composite reliability, and average variance extracted for all study constructs based on 1029 student responses, 2025.

Variables	Value, mean (SD)	Cronbach α	CR ^a	AVE ^b
Cognitive performance	4.18 (0.72)	0.89	0.91	0.66
Gamified PE ^c engagement	4.25 (0.68)	0.92	0.93	0.70
Physical activity enjoyment	4.12 (0.75)	0.87	0.89	0.64
Peer collaboration	4.05 (0.81)	0.86	0.88	0.61
Teacher support	4.1 (0.70)	0.91	0.92	0.69
Intrinsic motivation	4.21 (0.65)	0.90	0.91	0.68
AI ^d -based feedback	3.89 (0.79)	0.88	0.90	0.65

^aCR: composite reliability.

^bAVE: average variance extracted.

^cPE: physical education.

^dAI: artificial intelligence.

Cognitive performance, the study's primary dependent measure, had a mean score of 4.18 (SD 0.72). With modest fluctuation, participants reported strong cognitive engagement and outcomes. The scale's Cronbach α of 0.89 indicated high internal consistency. The CR was 0.91, and the AVE was 0.66, both

exceeding the convergent validity criteria of 0.50. Scale elements adequately explained the hidden component. Gamified PE Engagement, a major independent variable, had the highest mean score (mean 4.25, SD 0.68), reflecting positive participant views of PE gamification. Cronbach α of 0.92, CR of 0.93, and

AVE of 0.70 indicate the reliability of this scale. These statistics demonstrate that the gamified engagement measurement model is statistically valid and substantively relevant due to its strong internal cohesion. Another independent variable with a high mean score of 4.12 (SD 0.75) supported favorable participant attitudes toward PE. AVE 0.64, Cronbach α 0.87, and CR 0.89 all exceeded acceptable norms, indicating the construct's internal reliability. Gamified activities motivated and engaged participants in PE. Peer collaboration also had strong psychometric features, with a mean of 4.05 (SD 0.81), CR=0.88, AVE=0.61, and Cronbach α =0.86. These findings support the favorable benefits of collaborative learning on academic success and demonstrate that the instrument effectively captures the social dynamics and cooperative learning aspects of PE. Teacher support was another major independent variable with high reliability and validity. A mean of 4.10 (SD 0.70), Cronbach α of 0.91, CR of 0.92, and AVE of 0.69 indicated a robust construct. These statistics highlight the role of instructors in

designing, guiding, and supporting gamified education and physical learning, as well as cognitive outcomes and participant motivation. Most participants were motivated and self-driven in gamified learning contexts, with an average Intrinsic Motivation score of 4.21 (SD 0.65). This construct had an AVE of 0.68, a CR of 0.91, and a Cronbach α of 0.90. Self-determination theory suggests that intrinsic motivation is crucial for academic achievement, and these values support the statistical validity of the motivation construct. Finally, AI-based feedback had a lower mean of 3.89 (SD 0.79), suggesting a wider range of responses. However, Cronbach α scores of 0.88, CR 0.90, and AVE 0.65 indicate strong construct validity and internal consistency. Although experiences with AI feedback varied (possibly due to familiarity or availability), participants generally reported that it improved engagement and performance. [Table 4](#) shows the item-level standardized loadings and metric invariance estimates.

Table 4. Item-level standardized loadings and metric invariance results for all constructs measured in 1029 high school students across 4 Chinese cities, 2025.

Variable and item	λ (standardized loading)	Cross-loading (max)	Mandarin translation ΔCFI^a
Cognitive performance (CP)			
CP1	0.78	0.12	0.003
CP2	0.81	0.10	0.003
CP3	0.83	0.09	0.003
CP4	0.79	0.11	0.003
CP5	0.82	0.10	0.003
Gamified PE^b engagement (GPE)			
GPE1	0.85	0.11	0.004
GPE2	0.88	0.10	0.004
GPE3	0.86	0.12	0.004
GPE4	0.84	0.09	0.004
GPE5	0.87	0.10	0.004
Physical activity enjoyment (PAE)			
PAE1	0.79	0.08	0.002
PAE2	0.81	0.07	0.002
PAE3	0.80	0.09	0.002
PAE4	0.82	0.08	0.002
PAE5	0.79	0.07	0.002
Peer collaboration (PC)			
PC1	0.75	0.10	0.003
PC2	0.78	0.11	0.003
PC3	0.77	0.09	0.003
PC4	0.76	0.08	0.003
PC5	0.79	0.10	0.003
Teacher support (TS)			
TS1	0.84	0.09	0.003
TS2	0.87	0.08	0.003
TS3	0.85	0.10	0.003
TS4	0.86	0.09	0.003
TS5	0.88	0.08	0.003
Intrinsic motivation (IM)			
IM1	0.82	0.10	0.002
IM2	0.85	0.11	0.002
IM3	0.83	0.09	0.002
IM4	0.81	0.08	0.002
IM5	0.84	0.10	0.002
AI^c-based feedback (AIF)			
AIF1	0.80	0.09	0.003
AIF2	0.83	0.10	0.003
AIF3	0.81	0.08	0.003
AIF4	0.82	0.09	0.003

Variable and item	λ (standardized loading)	Cross-loading (max)	Mandarin translation ΔCFI^a
AIF5	0.84	0.10	0.003

^aCFI: comparative fit index.

^bPE: physical education.

^cAI: artificial intelligence.

The study looked at item-level standardized loadings (λ) and cross-loadings for all latent constructs to ensure the study instrument was reliable and valid (refer to Table 5). The standardized loadings for each item were 0.75–0.88, indicating robust construct representation, while the cross-loadings were less than 0.12, indicating discriminant validity. The translation

into Mandarin was tested for metric invariance, and all constructs had ΔCFI values <0.01 , which indicates that the translated items perform the same as the original instrument. Table 5 provides a comparison of original CF and higher-order CFA models, which helps determine whether the measurement model in this research is legitimate.

Table 5. Comparison of original confirmatory factor analysis (CFA) and higher-order CFA models, including model fit indices, for constructs measured in 1029 high school students in Beijing, Shanghai, Chengdu, and Guangzhou, 2025.

Model	CFI ^a	TLI ^b	RMSEA ^c	SRMR ^d	χ^2 (<i>df</i>)	AIC ^e	BIC ^f
Original CFA ^g	0.961	0.953	0.042	0.038	2.85 (532)	2180.45	2245.67
Higher-order CFA (reduced model)	0.954	0.946	0.043	0.040	2.87 (534)	2145.32	2210.58

^aCFI: comparative fit index.

^bTLI: Tucker-Lewis Index.

^cRMSEA: root-mean-square error of approximation.

^dSRMR: standardized root-mean-square residual.

^eAIC: Akaike information criterion

^fBIC: Bayesian information criterion

^gCFA: confirmatory factor analysis.

In a higher-order structural factor analysis (CFA), the study combined 2 latent factors: intrinsic motivation and enjoyment of physical exercise, and cognitive performance and teamwork. Having a χ^2/df ratio of 2.87 and good fit indices (CFI=0.954, TLI=0.946, RMSEA=0.043, and standardized root-mean-square residual=0.040), the higher-order model effectively fit the measurement model. The higher-order CFA has lower AIC (2145.32) and BIC (2210.58) than the original CFA, indicating

that the reduced model is more parsimonious without losing theoretical or empirical validity. These results suggest that the dual-framework model (SDT + CLT) may retain construct validity and conceptual coherence with fewer latent variables.

The Fornell-Larcker criterion and the heterotrait-monotrait (HTMT) ratio of correlations were used to assess the discriminant validity of SEM. Table 6 provides their results.

Table 6. Fornell-Larcker criteria and the heterotrait-monotrait (HTMT) ratio assessing the discriminant validity of all constructs in 1029 high school students, 2025.

Constructs	Fornell-Larcker diagonal	HTMT ^a ratios (max)
Cognitive performance	0.81	0.82
Gamified PE ^b	0.84	0.79
PE ^b enjoyment	0.80	0.76
Peer collaboration	0.78	0.74
Teacher support	0.83	0.77
Motivation	0.82	0.81
AI ^c feedback	0.81	0.80

^aHTMT: heterotrait-monotrait.

^bPE: physical education.

^cAI: artificial intelligence.

The diagonal line shows the square root of each construct's AVE compared to the interconstruct correlations, using Fornell-Larcker criteria. This method demonstrates discriminant

validity when the square root of a concept's AVE is greater than its highest correlation with any other construct. Cognitive performance's Fornell-Larcker diagonal value of 0.81 is greater

than its connections with gamified PE, motivation, and AI feedback. AI feedback (0.81), gamified PE (0.84), PE enjoyment (0.80), peer collaboration (0.78), teacher support (0.83), motivation (0.82), and AI Support (0.83) all have top diagonal values, indicating that each construct's items are more strongly related to their latent variable than to any other. The HTMT ratio was used to enhance this investigation. PLS-SEM applies HTMT for a more sensitive and reliable discriminant validity test. Discriminant validity is good when HTMT values are

<0.85. All HTMT scores in the results were between 0.74 and 0.82. For cognitive performance, none of the constructs has an HTMT ratio >0.82, which is sufficient. AI feedback (0.80), gamified PE (0.79), PE enjoyment (0.76), peer collaboration (0.74), teacher support (0.77), motivation (0.81), and AI support (0.79) all fall below the threshold, supporting the premise that each latent concept is empirically distinct. Table 7 shows the estimates of Harman single-factor and latent common method variance.

Table 7. Harman single-factor and latent common method variance (CMV) tests to evaluate potential survey bias in 1029 high school student responses, 2025.

Test type	No of factors extracted	Variance explained by first factor (%)	Total variance explained (%)	ΔCFI^a (with CMV ^b factor)	$\Delta\chi^2^c$ (df)	P value	Conclusion
Harman single-factor test	7	28.6	72.4	—	—	—	No CMV concern (first factor<40%)
Latent common method factor test	7 + CMV	26.9	74.1	0.006	21.38 (1)	.09	No significant CMV effect ($\Delta CFI < 0.01$)

^aCFI: common method variance.

^bCMV: common method variance.

^c $\Delta\chi^2$: change in chi-square value.

Harman single-factor test and the latent CMV factor test were used to assess for self-reported data-related CMV. The unrotated factor analysis found 7 variables, although the first accounted for 28.6% of the variance (vs 40%). This indicates that no factor dominated the item-level covariance structure. This finding was confirmed by adding a latent CMV component to the measurement model. The model fit comparison showed no significant reduction in fit, even after controlling for method effects ($\Delta CFI=0.006$ and $\Delta\chi^2_1=21.38$; $P=.09$). These findings strongly suggest that common method bias does not threaten the study's validity and that discriminant integrity of latent constructs is preserved.

Table 8 provides the SEM findings, which strongly support the research's theoretical approach. There is a positive relationship between gamified PE and students' cognitive performance. Gamified PE may enhance students' concentration, memory, and information-processing skills [84]. According to the engagement-learning paradigm, students learn meaningfully when they are emotionally, behaviorally, and cognitively involved [85]. Game elements, such as goal-setting, fast feedback, and reward systems, in PE sessions, may help children learn more effectively and improve cognitively [86]. Chaiyarat [87] and Aibar-Almazán et al [88] reported that gamification can make the classroom more dynamic and engaging, thereby enhancing students' problem-solving and critical-thinking skills.

Table 8. Structural equation modeling (SEM) path estimates showing relationships among studied factors with standardized beta coefficients and significance levels, based on 1029 students, 2025.

Path	Estimates (β)	SE	t statistic	P value
Gamified PE ^a → cognitive performance	.34 ^b	0.05	6.80	<.001
Enjoyment → cognitive performance	.28 ^b	0.06	4.67	<.001
Collaboration → cognitive performance	.26 ^b	0.07	3.71	<.001
Teacher support → cognitive performance	.31 ^b	0.06	5.17	<.001

^aPE: physical education.

^b $P < .001$ (1% significance level).

Teacher support is the second strongest predictor of students' cognitive performance. This suggests that instructors must understand how to positively impact students' cognitive growth through effective lesson preparation, positive reinforcement, and constructive feedback. Gamified and AI-feedback classrooms require teacher facilitation [89]. Teachers should motivate students by establishing a psychologically safe classroom, helping them grasp and implement challenging feedback, and scaffolding their learning to support their growth

and development [90]. Cha et al [91] found that students feel more confident, engaged, and cognitively involved in classroom tasks, especially in active and digitally mediated settings, when they view their teachers as accessible, attentive, and helpful.

There is a positive relationship between physical activity enjoyment and students' cognitive performance, which suggests that students who enjoy PE may focus better, experience lower cognitive stress, and feel emotionally good, which helps them

learn. According to SDT, individuals are more engaged and learn more when they have a personal interest in the result [92]. Chen [93] found that children who enjoy physical activities are more willing to participate and better able to harness the psychological benefits of exercise, leading to improved classroom concentration, memory, and performance. Enjoyment enhances cognition in gamified environments by reducing performance anxiety and boosting self-confidence.

Peer collaboration and cognitive performance were significantly positively associated. Collaborating with peers during gamified PE classes boosts cognitive development. The focus on idea-sharing, collaborative problem-solving, and social and emotional support in peer learning enhances understanding and mental agility. This result is supported by Vygotsky's sociocultural theory, which emphasizes the role of social interaction in internalizing information [94]. Qi and Derakhshan [95] found that physically active educational activities help students develop social and cognitive skills for academic

success. Collaboration, negotiation, and peer feedback improve metacognition and learning.

The study analyzed cluster effects in Beijing, Shanghai, Chengdu, and Guangzhou using intraclass correlation coefficients (ρ) and design effects for each latent construct. The intraclass correlation coefficients ranged from 0.02 to 0.06, resulting in design effects of 7.18 to 19.90 with an average cluster size of 310 respondents per city (Table 9). After achieving the threshold ($\rho > 0.05$), a multilevel structural equation model with random intercepts was used to review cognitive performance and AI-based feedback. There were minor differences in explained variance ($\Delta R^2 = +0.02$ to $+0.05$) and standardized path coefficients ($\Delta \beta < .02$) between single-level and multilevel estimations. The findings reveal that provincial clustering did not significantly affect gamified PE engagement, intrinsic motivation, and cognitive performance. Therefore, the estimation technique does not influence structural model robustness.

Table 9. Cluster diagnostics and multilevel structural equation modeling (SEM) robustness checks verifying the stability of results across school clusters in 1029 high school students, 2025.

Construct	Number of clusters (cities)	Average cluster size (n)	ICC ^a (ρ)	Design effect= $1+(n-1)\rho$	Single-level β (gamified PE \rightarrow cognitive performance)	Multilevel β (random intercept model)	$\Delta \beta^b$ (bias)	ΔR^2 (change in explained variance)
Cognitive performance (CP)	4 (Beijing, Shanghai, Chengdu, and Guangzhou)	310	0.05	16.45	.45	.43	-.02	0.04
Gamified PE ^c engagement	4	310	0.04	13.36	.34	.33	-.01	0.03
Intrinsic motivation	4	310	0.03	10.27	.21 (indirect $\kappa^2=0.16$)	.20	-.01	0.02
AI ^d -based feedback	4	310	0.06	19.9	.18 (interaction β)	.17	-.01	0.05
Teacher support	4	310	0.02	7.18	.31	.30	-.01	0.02
Peer collaboration	4	310	0.03	10.27	.26	.25	-.01	0.02

^aICC: intraclass correlation coefficient.

^b $\Delta \beta$: change in bias.

^cPE: physical education.

^dAI: artificial intelligence.

Table 10 assessed the impact of AI-based feedback on the model using effect-size indices (κ^2 and PM), conditional effects at ± 1 SD, and ΔR^2 . The mediation study revealed a medium-to-large effect size ($\kappa^2=0.19$) and PM=0.36, suggesting that intrinsic motivation indirectly accounts for 36% of the connection. Intrinsic motivation substantially influenced the link between gamified PE and cognitive performance ($\beta=.23$; $P<.001$). A moderated analysis indicated that AI-based feedback increased the positive correlation between gamified PE and cognitive

performance ($\beta=.17$; $P<.001$). The correlation between AI feedback and gamified learning gains increased with complexity, with $\beta=.13$ at low levels (-1 SD), $\beta=.20$ at medium, and $\beta=.28$ at high levels ($+1$ SD). The inclusion of the interaction term raised the model's explanatory power from $R^2=0.47$ to $R^2=0.55$, resulting in a ΔR^2 of 0.08, indicating that AI-based feedback explained an additional 8% of the variance in students' cognitive performance.

Table 10. Moderation-mediation analysis showing effect sizes, conditional effects (± 1 SD of artificial intelligence [AI]-based feedback), and ΔR^2 for relationships between gamified physical education (PE), intrinsic motivation, and cognitive performance in 1029 students, 2025.

Path	Estimate (β)	SE	<i>t</i> value	<i>P</i> value	LLCI ^a	ULCI ^b	κ^2 ^c	PM ^d	Condi- tional effect (-1 SD) ^e	Condition- al effect (mean) ^e	Condi- tional effect (+1 SD) ^e	<i>R</i> ²	ΔR^2 ^f
Mediation effect: gamified PE → intrinsic motivation → cognitive performance	.23 ^g	0.04	5.75	<.001	0.15	0.32	0.19	.36	—	—	—	0.47	—
Moderation effect: gamified PE × AI feedback → cognitive performance	.17 ^g	0.03	5.67	<.001	0.10	0.24	—	—	0.13	0.20	0.28	0.55	+0.08
Total effect: gamified PE → cognitive performance (including mediation and moderation)	.45 ^g	0.05	9.00	<.001	0.35	0.55	—	—	—	—	—	—	—

^aLLCI: lower limit CI.^bULCI: upper limit CI.^c κ^2 : standardized indirect effect size.^dPM: proportion mediated.^eConditional effects: changes in the slope of gamified PE → cognitive performance at low (-1 SD), mean, and high (+1 SD) levels of artificial intelligence (AI)-based feedback.^f ΔR^2 : 0.08, indicating an 8% increase in explained variance after counting the moderation term.^g*: indicates a 1% significance level.

The quantity of AI feedback in gamified PE sessions affects student engagement and cognitive performance. Gamified PE improves cognitive function, and AI-based feedback enhances this impact. Children who participate in gamified PE and receive continuous, real-time AI-powered feedback are more likely to experience enhanced cognitive outcomes, including improved memory recall, attention, problem-solving, and mental engagement [96]. Feedback intervention theory suggests that timely and personalized feedback helps learners concentrate on task-related goals, self-regulate, and engage cognitively [97]. It may be challenging to obtain adaptive, data-driven insights in traditional PE settings, but AI-powered feedback is immediate, objective, and tailored. These systems, integrated into a gamified framework, make learning more engaging and challenging for students, providing explicit guidance on how to improve, which in turn fosters deeper thinking. According to Suresh Babu and Dhakshina Moorthy [98], gamification alone can boost student engagement. However, intelligent feedback systems can amplify these effects on cognition by influencing learner behavior and performance in real-time. The study's findings support the use of technology-enhanced, customized learning environments in modern school design. Gamified education uses AI feedback as a cognitive support system to help students recall and apply what they have learned via hands-on, interactive activities [99]. CLT suggests that real-time

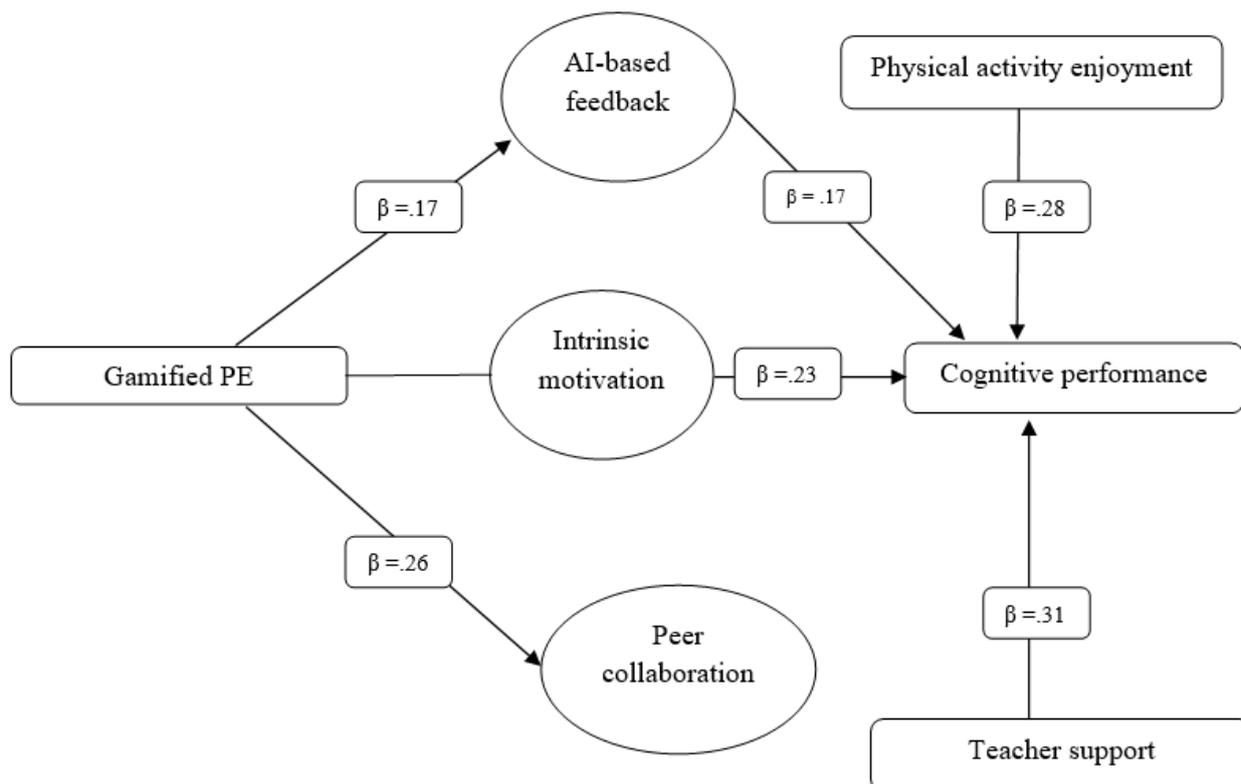
AI feedback, which optimizes task difficulty and decreases ambiguity, helps learners focus on important cognitive activities. The moderating impact also affects curriculum and educational policy, suggesting that complex feedback systems are needed for gamified physical instruction to maximize cognitive outcomes. For gamified learning to be most effective, schools and instructors should use or invest in AI-powered solutions that tailor insights and feedback to each student's unique profile and cognitive capabilities.

Gamified PE improves cognitive function, and intrinsic motivation is a crucial psychological factor. Gamified PE enhances intrinsic motivation, leading to improved cognitive performance. The indirect effect explains most of the variance in cognitive performance, confirming that gamification's structure and qualities are significant, but what counts most are learners' psychological moods. Self-determination theory emphasizes relatedness, competence, and autonomy as components of intrinsic motivation, which is reflected in this mediation effect. Gamified learning environments enhance intrinsic interest and satisfaction by fostering autonomy through choice, competence through manageable tasks, and relatedness through peer collaboration [53]. Students who are genuinely motivated to study are more likely to use metacognitive skills, pay attention, and learn more deeply. Shalгимbekova et al [100] suggest that motivation moderates the link between instructional

design and learning outcomes. This result highlights the role of motivation in mediating the pedagogical efficacy of innovative teaching methods. Even if gamified PE is fun and structured, internalizing values and desires drives cognitive growth. This study supports the idea that the motivating processes of instructional inputs are as essential as the inputs themselves in cognitive performance [101]. Gamification sets the scene, but intrinsic motivation propels cognitive functioning. This mediation strengthens Hypothesis H3, which states that intrinsic motivation mediates the relationship between gamified PE and

cognitive performance. Gamification alone is ineffective; however, students' natural incentive to participate and accomplish tasks substantially enhances outcomes, as evidenced by the considerable indirect path. This means that instructional designers, policymakers, and instructors should support children's intrinsic drive to learn and incorporate engaging aspects into PE interventions to enhance academic cognition. Figure 2 shows the structural model with standardized path coefficients.

Figure 2. Structural model displaying standardized path coefficients for all hypothesized relationships among studied factors in 1029 high school students, 2025. Source: Author's estimate. AI: artificial intelligence; PE: physical education.



The significant first-order factor loadings ($\lambda=0.65-0.88$; $P<.001$) indicate high correlations between observable indicators and their latent components (refer to Table 11). The substantial second-order factor loadings ($\lambda=0.71-0.84$; $P<.001$) suggest

that SDT and CLT collaborate to create a more sophisticated integrative model. The complete model suited the specified conceptual framework well, with CFI=0.957, TLI=0.949, and RMSEA=0.041.

Table 11. Structural equation modeling (SEM) results for the dual-framework model with standardized path coefficients in 1029 students, 2025.

Analysis	Statistic	P value
Second-order CFA ^a	Factor loadings: $\lambda=0.65-0.88$ (1st-order) and $\lambda=0.71-0.84$ (2nd-order)	<.001
Model fit indices	CFI ^b =0.957, TLI ^c =0.949, and RMSEA ^d =0.041	— ^e
Covariance	Covariance=0.42	<.001
Cross-loading check	Max cross-loading=0.25	— ^e
Latent interaction test	$\Delta\chi^2=4.12^f$.04

^aCFA: confirmatory factor analysis.

^bCFI: comparative fit index.

^cTLI: Tucker-Lewis Index.

^dRMSEA: root-mean-square error of approximation.

^eNot available.

^fIndicates a 5% significance level.

Further, a significant association (covariance=0.42; $P<.001$) was found between intrinsic motivation and AI-based feedback, using a covariance approach. While preserving discriminant validity, this study supports the significant link between these factors. The measurement model was tested, and cross-loadings indicated that cognitive performance, intrinsic motivation, and

AI-based feedback are empirically independent, supporting the structural coherence of the dual-framework SEM.

To determine whether AI-based feedback has context-specific effects, the study conducted multigroup moderation tests by level of technology access (high, moderate, and low) and location (urban vs rural; refer to [Table 12](#)).

Table 12. Multigroup moderation analysis examining differences in study relationships by urban vs rural location and technology access in 1029 high school students across 4 cities, 2025.

Path	Group	Sample size (n)	β estimate	SE	t value	P value	ΔR^2^a	f^2^b
Gamified PE ^c → cognitive performance	Urban	366	.35 ^d	0.06	5.83	<.001	0.05	0.16
Gamified PE → cognitive performance	Rural	303	.31 ^d	0.07	4.43	<.001	0.05	0.15
Gamified PE → cognitive performance	High-tech access	400	.37 ^d	0.05	7	<.001	0.07	0.18
Gamified PE → Cognitive Performance	Moderate tech access	451	.34 ^d	.05	6.8	<.001	0.05	0.16
Gamified PE → cognitive performance	Low-tech access	178	.28 ^d	0.06	4.67	<.001	0.04	0.14

^a ΔR^2 : $P<.05$.

^b f^2 : $P<.05$.

^cPE: $P<.05$.

^dIndicates a 1% significance level.

[Table 12](#) shows a substantial influence of gamified PE on cognitive performance across all groups ($\beta=.28-.37$; $P<.001$). Impacts were higher for urban students and those with high technological access ($\beta=.35$ and $.37$, respectively) and lower for rural students and those with low access ($\beta=.31$ and $.28$, respectively). Moderation had medium to high impacts, with ΔR^2 values ranging from 0.04 to 0.07 and f^2 effect sizes from 0.14 to 0.18. These findings suggest that gamified PE interventions based on AI are generally effective, with urban students and those with greater technological access benefiting more cognitively.

Discussion

Principal Findings

This study examined how intrinsic motivation and AI-based feedback affect high school students' cognitive performance after gamified PE. The first objective was to explore the effects of gamified PE and teacher support on students' cognitive performance; statistically significant benefits were shown, confirming Hypothesis 1. The second goal was to assess how physical exercise enjoyment and peer collaboration affect cognitive performance, which was confirmed by the strong positive connections between these factors and cognitive performance, supporting Hypothesis 2. The findings supported the third aim, that is, determining whether gamified PE

influences cognitive performance and whether intrinsic motivation mediated this impact. The positive effect of gamified PE on cognitive performance was partially mediated by intrinsic motivation, but AI-based feedback enhanced it, supporting Hypothesis 3. Gamified PE improves cognitive performance through motivational and technical processes, in line with the research's goals and theoretical framework.

Interpretation and Implications

Gamified PE improves cognitive function, suggesting that schools can create engaging, dynamic, and technologically advanced learning environments. Gamification boosts cognitive functioning by focusing students, simplifying problem-solving, and stimulating active involvement. Well-designed, gamified PE interventions may achieve these aims, supporting the view of Barz et al [102] that game-based learning may boost motivation and cognition. The SDT concept was reinforced by intrinsic motivation, which asserts that students who feel autonomous, competent, and socially engaged are more likely to participate in meaningful activities and achieve higher-order cognitive achievements. This suggests that gamified PE improves cognition and intrinsic motivation, which sustains engagement and learning [103]. Student involvement and academic success are highest in PE programs that incorporate enjoyable and independent activities. Social and instructional factors, including a friendly teacher, exercise, and teamwork, promote learning. Game-based and AI-supported learning benefit from human facilitation, although instructors' guidance, encouragement, and scaffolding are still needed. Teamwork promotes collaboration, communication, and executive functioning, which may explain its cognitive advantages. These findings support more comprehensive theoretical frameworks that promote student-centered, socially engaged, and holistic learning environments [104]. AI-based feedback moderates the effect of adaptive, real-time recommendations on learning. AI-assisted strategy correction, refinement, and progressive advancement, delivered through rapid, tailored feedback, boosts cognitive performance beyond the reach of gamification. Recent AI-assisted learning experiments have shown potential for cognitive optimization and personalized skill improvement [105].

Comparison With Existing Literature

This study builds on previous classroom gamification research. Gamification and AI feedback in PE can enhance cognitive performance, contrary to previous research on academic participants [106,107]. While gamification has been shown to improve motivation, little research has linked it to cognitive outcomes in PE [108,109]. These findings suggest that technology-driven feedback and compelling design may boost students' cognitive performance. The study contributes to AI in education research by confirming previous classroom-based studies [110,111] that real-time adaptive feedback increases learning. This study combines SDT and CLT to advance theory. Gamified experiences increase learning via intrinsic motivation and cognitive load. These results provide a foundation for future research on integrated, technologically advanced, motivation-based educational interventions.

Practical Implications

The findings suggest that AI-based feedback mechanisms could be beneficial to educational institutions, particularly in technologically advanced cities. To achieve this, AI-powered educational systems must be funded. These systems should have adaptive capabilities and age-appropriate user interfaces to provide children with timely, individualized feedback on their PE, behavioral engagement, and cognitive attention. Policy cooperation among Chinese education ministries, IT businesses, and AI research institutes may accelerate the development of educational systems tailored to the nation. To ensure that technology enhances human teaching, a national AI-in-education framework must guide deployment, data ethics, privacy protection, and teacher capacity-building. Pilot studies in schools with suitable digital infrastructure may examine the effects of AI-enhanced, gamified learning on academic achievement. Since public and semigovernment institutions had larger performance disparities, the findings imply a shift toward socially and motivationally enhanced learning. Policies should prioritize training PE instructors to create welcoming, inclusive, and inspirational learning environments. To stimulate cognitive development through intrinsic motivation and peer collaboration, schools must prioritize students' mental health, incorporate team-building activities, and use inclusive instructional techniques that cater to their unique social identities and needs. Teacher performance measurements should incorporate student-centered learning, emotional support, and inclusive engagement as part of education reforms. Physical exercise, reflection, feedback, and cognitive challenge must be balanced in gamified learning settings that promote collaboration and active learning.

Limitations

This study has made many essential contributions. However, there are some limitations. Although the research used a large and diverse sample from 4 metropolitan cities, its cross-sectional design limits causal inferences. SEM and bootstrapping provide more compelling findings; however, experimental or longitudinal designs are necessary to determine the long-term impact of gamified PE and AI feedback on cognitive function. Second, self-report assessments may be biased by social desirability or cognitive misunderstanding, even when administered under supervision to children aged 10-12 years. Future studies could triangulate findings and improve measurement validity by using multi-informant data, such as behavioral observations or teacher ratings. The generalizability of the results is another issue. The sample included students from urban and suburban public, private, and semigovernment schools, but not rural or low-tech institutions. This raises the question of how AI-based gamified systems can effectively serve diverse socioeconomic and geographical conditions. Future research should examine how infrastructural variations affect such efforts and the digital divide. Although this study focused on cognitive function, it did not examine emotional regulation, academic resilience, or physical health. Future multidimensional research should examine how gamified and AI-supported instructional techniques affect broader student capacities. This study focused on AI feedback as a moderator and intrinsic motivation as a mediator. Other factors, such as self-efficacy,

goal orientation, and support from family and friends, influence students' attitudes toward gamified learning. Future studies should incorporate additional mediators and moderators to gain a deeper understanding of the complex interplay among human, systemic, and technological factors. In conclusion, qualitative research methods, such as student interviews or classroom ethnographies, can complement quantitative approaches by providing contextual insights into learners' experiences and preferences. By addressing these limitations, future research may enhance the theoretical and practical understanding of how gamified, technology-enhanced education can improve student learning and growth.

Conclusions

The study concludes that Gamified PE with intrinsic motivation and AI-based feedback improves high school students' cognitive performance. These findings demonstrate that technology-based student-centered PE programs improve engagement, motivation, and cognition. Gamification and AI in education may improve students' overall development and academic and social results outside of the classroom. Future research should examine the long-term impacts, demographic factors, and technological developments to enhance the cognitive and motivational benefits of school-based treatments. To capitalize on the full potential of gamified PE for cognitive development, policymakers in education must institutionalize gamified pedagogy into the

national PE curriculum. Schools should update their PE curriculum to highlight the connection between physical health and emotional and intellectual development. The curriculum should incorporate more structured, game-based learning methods that enhance cognitive abilities. In PE, goal-setting, point-scoring, and challenge-based learning help students concentrate, recall, and solve issues. The Ministry of Education, along with state and regional education bureaus, should educate PE teachers in gamified teaching methods to meet pedagogical criteria, engage students, and measure academic achievement.

This study advances the literature by introducing a technology-driven model of PE that combines gamification with AI-based feedback to enhance students' cognitive abilities. This study demonstrates the synergistic effects of gamification, intrinsic motivation, and AI-supported feedback on cognitive development, in contrast to previous research that has primarily focused on conventional PE or singular teaching methodologies. This study contributes to the existing literature on digital and AI-enhanced education by using a substantial, multicity sample and SEM to demonstrate the validity of these connections. The findings have substantial implications for educational institutions, educators, and policymakers, suggesting that interactive, AI-driven PE programs can effectively enhance student engagement, learning processes, and cognitive performance in genuine educational environments.

Acknowledgments

No generative artificial intelligence tools were used in the writing, editing, data analysis, or preparation of this paper.

Data Availability

The dataset generated and analyzed during this study are not publicly available due to privacy protections for minors and institutional restrictions. Deidentified data may be made available upon reasonable request to the corresponding author, subject to ethical approval.

Funding

No external financial support or grants were received for this work.

Authors' Contributions

JZ contributed to the study's conceptualization, methodology development, data collection, data curation, formal analysis, and drafting of the original paper. SSO supported the conceptualization of the research, assisted with data collection, and contributed to critical review and paper editing. YX supervised the overall study, provided guidance throughout the research process, contributed to study design and interpretation of findings, participated in paper revision, and managed project administration as the corresponding author.

Conflicts of Interest

None declared.

References

1. Guo X, Chen S, Guo Y. Advancing multimodal teaching: a bibliometric and content analysis of trends, influences, and future directions. *Humanit Soc Sci Commun* 2024;11(1):1-22. [doi: [10.1057/s41599-024-04254-0](https://doi.org/10.1057/s41599-024-04254-0)]
2. Triola MM, Rodman A. Integrating generative artificial intelligence into medical education: curriculum, policy, and governance strategies. *Academic Medicine* 2025;100(4):413-418. [doi: [10.1097/ACM.0000000000005963](https://doi.org/10.1097/ACM.0000000000005963)]
3. Mallek F, Mazhar T, Faisal Abbas Shah S, Ghadi YY, Hamam H. A review on cultivating effective learning: synthesizing educational theories and virtual reality for enhanced educational experiences. *PeerJ Comput Sci* 2024;10:e2000 [FREE Full text] [doi: [10.7717/peerj-cs.2000](https://doi.org/10.7717/peerj-cs.2000)] [Medline: [38855256](https://pubmed.ncbi.nlm.nih.gov/38855256/)]

4. Latino F, Romano G, Tafuri F. Physical education teacher's continuing professional development affects the physiological and cognitive well-being of school-age children. *Education Sciences* 2024;14(11):1199. [doi: [10.3390/educsci14111199](https://doi.org/10.3390/educsci14111199)]
5. Christopoulos A, Mystakidis S, Stylios C, Tsoulos IG. Unplugged gamification in education: developing computational thinking skills through embodied gameplay. *J Educ Res* 2025;119(1):77-90 [FREE Full text] [doi: [10.1080/00220671.2025.2517266](https://doi.org/10.1080/00220671.2025.2517266)]
6. Martín-Rodríguez A, Gostian-Ropotin LA, Beltrán-Velasco AI, Belando-Pedreño N, Simón JA, López-Mora C, et al. Sporting mind: the interplay of physical activity and psychological health. *Sports (Basel)* 2024;12(1):37 [FREE Full text] [doi: [10.3390/sports12010037](https://doi.org/10.3390/sports12010037)] [Medline: [38275986](https://pubmed.ncbi.nlm.nih.gov/38275986/)]
7. Shah P, Gadekar NA, Taylor SS. Escape from passive learning: harnessing gamified digital educational escape rooms for immersive marketing education. *J Mark Educ* 2024;47(3):211-231 [FREE Full text] [doi: [10.1177/02734753241306093](https://doi.org/10.1177/02734753241306093)]
8. Lampropoulos G, Sidiropoulos A. Impact of gamification on students' learning outcomes and academic performance: a longitudinal study comparing online, traditional, and gamified learning. *Education Sciences* 2024;14(4):367. [doi: [10.3390/educsci14040367](https://doi.org/10.3390/educsci14040367)]
9. Kim J, Lindqvist AK, Castelli DM. Feasibility of utilizing gamified learning as a motivational strategy for promoting physical activity participation and healthy eating among college students. *J Am Coll Health* 2025;73(6):2651-2659. [doi: [10.1080/07448481.2024.2325929](https://doi.org/10.1080/07448481.2024.2325929)] [Medline: [38546660](https://pubmed.ncbi.nlm.nih.gov/38546660/)]
10. Guamanga MH, Saiz C, Rivas SF, Almeida LS. Analysis of the contribution of critical thinking and psychological well-being to academic performance. *Front. Educ* 2024;9:1423441. [doi: [10.3389/educ.2024.1423441](https://doi.org/10.3389/educ.2024.1423441)]
11. Zhang Y. Impact of arts activities on psychological well-being: emotional intelligence as mediator and perceived stress as moderator. *Acta Psychol (Amst)* 2025;254:104865 [FREE Full text] [doi: [10.1016/j.actpsy.2025.104865](https://doi.org/10.1016/j.actpsy.2025.104865)] [Medline: [40049081](https://pubmed.ncbi.nlm.nih.gov/40049081/)]
12. Sanaeifar F, Pourranjbar S, Pourranjbar M, Ramezani S, Mehr SR, Wadan AS, et al. Beneficial effects of physical exercise on cognitive-behavioral impairments and brain-derived neurotrophic factor alteration in the limbic system induced by neurodegeneration. *Exp Gerontol* 2024;195:112539 [FREE Full text] [doi: [10.1016/j.exger.2024.112539](https://doi.org/10.1016/j.exger.2024.112539)] [Medline: [39116955](https://pubmed.ncbi.nlm.nih.gov/39116955/)]
13. Martín-Rodríguez A, Madrigal-Cerezo R. Technology-enhanced pedagogy in physical education: bridging engagement, learning, and lifelong activity. *Education Sciences* 2025;15(4):409. [doi: [10.3390/educsci15040409](https://doi.org/10.3390/educsci15040409)]
14. Li Q, Li L, He X, Wang H. Exploring adolescent moderate-to-vigorous physical activity in China: mediating roles of school climate, perceived barriers, and physical education satisfaction. *Risk Manag Healthc Policy* 2024;17:3125-3136 [FREE Full text] [doi: [10.2147/RMHP.S497472](https://doi.org/10.2147/RMHP.S497472)] [Medline: [39687748](https://pubmed.ncbi.nlm.nih.gov/39687748/)]
15. Zhang J. Shaping the future: analysing early childhood education policy reforms and development in China. *Euro J of Education* 2025 Feb 21;60(1):e70055. [doi: [10.1111/ejed.70055](https://doi.org/10.1111/ejed.70055)]
16. Zhou S. Gamifying language education: the impact of digital game-based learning on Chinese EFL learners. *Humanit Soc Sci Commun* 2024;11(1):1-14. [doi: [10.1057/s41599-024-04073-3](https://doi.org/10.1057/s41599-024-04073-3)]
17. Ryan RM, Deci EL. Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. *Contemp Educ Psychol* 2020;61:101860. [doi: [10.1016/j.cedpsych.2020.101860](https://doi.org/10.1016/j.cedpsych.2020.101860)]
18. Sotos-Martínez VJ, Tortosa-Martínez J, Baena-Morales S, Ferriz-Valero A. It's game time: improving basic psychological needs and promoting positive behaviours through gamification in physical education. *Eur Phys Educ Rev* 2023;30(3):435-457. [doi: [10.1177/1356336x231217404](https://doi.org/10.1177/1356336x231217404)]
19. Xiao Y, Hew KFT. Intangible rewards versus tangible rewards in gamified online learning: which promotes student intrinsic motivation, behavioural engagement, cognitive engagement and learning performance? *Brit J Educational Tech* 2023;55(1):297-317. [doi: [10.1111/bjet.13361](https://doi.org/10.1111/bjet.13361)]
20. Hong H, Dai L, Zheng X. Advances in wearable sensors for learning analytics: trends, challenges, and prospects. *Sensors (Basel)* 2025;25(9):2714. [doi: [10.3390/s25092714](https://doi.org/10.3390/s25092714)] [Medline: [40363153](https://pubmed.ncbi.nlm.nih.gov/40363153/)]
21. Wang Y. Report on smart education in China. In: Zhuang R, editor. *Smart Education in China and Central & Eastern European Countries*. Singapore: Springer; 2023.
22. Cao F, Xiang M, Chen K, Lei M. Intelligent physical education teaching tracking system based on multimedia data analysis and artificial intelligence. *Mob Info Sys* 2022;2022:7666615. [doi: [10.1155/2022/7666615](https://doi.org/10.1155/2022/7666615)]
23. Li N. Fostering creative thinking through gamified intercultural communication: innovative pedagogical strategies in higher education English programs. *Cogent Education* 2025;12(1):2489271. [doi: [10.1080/2331186x.2025.2489271](https://doi.org/10.1080/2331186x.2025.2489271)]
24. Gómez-Carrasco CJ, Monteagudo-Fernández J, Moreno-Vera JR, Sainz-Gómez M. Evaluation of a gamification and flipped-classroom program used in teacher training: perception of learning and outcome. *PLoS One* 2020;15(7):e0236083 [FREE Full text] [doi: [10.1371/journal.pone.0236083](https://doi.org/10.1371/journal.pone.0236083)] [Medline: [32673373](https://pubmed.ncbi.nlm.nih.gov/32673373/)]
25. Zhu T, Gu M. The application and implications of science and technology innovation in the management of education for Chinese students studying abroad in China in the epidemic era. *PLoS One* 2024;19(8):e0306785. [doi: [10.1371/journal.pone.0306785](https://doi.org/10.1371/journal.pone.0306785)] [Medline: [39150972](https://pubmed.ncbi.nlm.nih.gov/39150972/)]
26. Zhao B, Zhang Y. The impact of a personality trait rooted in Chinese confucian culture-Junzi personality-on loneliness and subjective well-being among Chinese individuals: the mediating role of self-reflection and self-control. *BMC Psychol* 2025;13(1):137. [doi: [10.1186/s40359-025-02488-4](https://doi.org/10.1186/s40359-025-02488-4)] [Medline: [39972519](https://pubmed.ncbi.nlm.nih.gov/39972519/)]

27. Wong D, Liu H, Meng-Lewis Y, Sun Y, Zhang Y. Gamified money: exploring the effectiveness of gamification in mobile payment adoption among the silver generation in China. *ITP* 2021;35(1):281-315. [doi: [10.1108/itp-09-2019-0456](https://doi.org/10.1108/itp-09-2019-0456)]
28. Ishaq K, Mat Zin NA, Rosdi F, Jehanghir M, Ishaq S, Abid A. Mobile-assisted and gamification-based language learning: a systematic literature review. *PeerJ Comput Sci* 2021;7:e496 [FREE Full text] [doi: [10.7717/peerj-cs.496](https://doi.org/10.7717/peerj-cs.496)] [Medline: [34084920](https://pubmed.ncbi.nlm.nih.gov/34084920/)]
29. Taşkın N, Kılıç Çakmak E. Effects of gamification on behavioral and cognitive engagement of students in the online learning environment. *Int J Hum-Comput Interact* 2022;39(17):3334-3345. [doi: [10.1080/10447318.2022.2096190](https://doi.org/10.1080/10447318.2022.2096190)]
30. Arufe-Giráldez V, Sanmiguel-Rodríguez A, Ramos-Álvarez O, Navarro-Patón R. Gamification in physical education: a systematic review. *Educ Sci* 2022;12(8):540. [doi: [10.3390/educsci12080540](https://doi.org/10.3390/educsci12080540)]
31. Wei Y, Wang L, Tang Y, Deng H, Su J, Li Q. Enhancing young children's executive function through physical activities: a three-level meta-analysis. *Ment Health Phys Act* 2024;26:100592. [doi: [10.1016/j.mhpa.2024.100592](https://doi.org/10.1016/j.mhpa.2024.100592)]
32. Camacho-Sánchez R, Manzano-León A, Rodríguez-Ferrer JM, Serna J, Lavega-Burgués P. Game-based learning and gamification in physical education: a systematic review. *Educ Sci* 2023;13(2):183. [doi: [10.3390/educsci13020183](https://doi.org/10.3390/educsci13020183)]
33. Christopoulos A, Mystakidis S. Gamification in education. *Encycl* 2023;3(4):1223-1243. [doi: [10.3390/encyclopedia3040089](https://doi.org/10.3390/encyclopedia3040089)]
34. Hsu CY, Wu TT. Application of business simulation games in flipped classrooms to facilitate student engagement and higher-order thinking skills for sustainable learning practices. *Sustainability* 2023;15(24):16867. [doi: [10.3390/su152416867](https://doi.org/10.3390/su152416867)]
35. Du Q. How artificially intelligent conversational agents influence EFL learners' self-regulated learning and retention. *Educ Inf Technol* 2025;30(15):21635-21701 [FREE Full text] [doi: [10.1007/s10639-025-13602-9](https://doi.org/10.1007/s10639-025-13602-9)]
36. Eng CM, Tsegai-Moore A, Fisher AV. Incorporating evidence-based gamification and machine learning to assess preschool executive function: a feasibility study. *Brain Sci* 2024;14(5):451 [FREE Full text] [doi: [10.3390/brainsci14050451](https://doi.org/10.3390/brainsci14050451)] [Medline: [38790430](https://pubmed.ncbi.nlm.nih.gov/38790430/)]
37. Zang W, Zhu J, Xiao N, Fang M, Li D, Li H, et al. Effects of aerobic exercise on children's executive function and academic performance: a systematic review and meta-analysis. *Heliyon* 2024;10(7):e28633 [FREE Full text] [doi: [10.1016/j.heliyon.2024.e28633](https://doi.org/10.1016/j.heliyon.2024.e28633)] [Medline: [38586404](https://pubmed.ncbi.nlm.nih.gov/38586404/)]
38. Ahmad F, Ahmed Z, Obaid I, Shaheen M, Rahman H, Singh G, et al. Motivation, engagement, and performance in two distinct modes of brain games: a mixed methods study in children. *International Journal of Human-Computer Interaction* 2024;41(15):9652-9666 [FREE Full text] [doi: [10.1080/10447318.2024.2426915](https://doi.org/10.1080/10447318.2024.2426915)]
39. Mo D, Xiang M, Luo M, Dong Y, Fang Y, Zhang S, et al. Using gamification and social incentives to increase physical activity and related social cognition among undergraduate students in Shanghai, China. *Int J Environ Res Public Health* 2019;16(5):858. [doi: [10.3390/ijerph16050858](https://doi.org/10.3390/ijerph16050858)]
40. Tsay CH, Kofinas AK, Trivedi SK, Yang Y. Overcoming the novelty effect in online gamified learning systems: an empirical evaluation of student engagement and performance. *Comput Assist Lang Learn* 2019;36(2):128-146. [doi: [10.1111/jcal.12385](https://doi.org/10.1111/jcal.12385)]
41. Sharma K, Nguyen A, Hong Y. Self - regulation and shared regulation in collaborative learning in adaptive digital learning environments: a systematic review of empirical studies. *Brit J Educational Tech* 2024;55(4):1398-1436. [doi: [10.1111/bjet.13459](https://doi.org/10.1111/bjet.13459)]
42. Ribeiro E, Farias C, Mesquita I. "The Game Changers": How equity-driven pedagogical scaffolding reduces participation disparities in physical education. *Educ Sci* 2024;14(10):1077. [doi: [10.3390/educsci14101077](https://doi.org/10.3390/educsci14101077)]
43. Rincon-Flores EG, Castano L, Guerrero Solis SL, Olmos Lopez O, Rodríguez Hernández CF, Castillo Lara LA, et al. Improving the learning-teaching process through adaptive learning strategy. *Smart Learn Environ* 2024;11(1):27. [doi: [10.1186/s40561-024-00314-9](https://doi.org/10.1186/s40561-024-00314-9)]
44. Boke H, Aygun Y, Tufekci S, Yagin FH, Canpolat B, Norman G, et al. Effects of cooperative learning on students' learning outcomes in physical education: a meta-analysis. *Front Psychol* 2025;16:1508808. [doi: [10.3389/fpsyg.2025.1508808](https://doi.org/10.3389/fpsyg.2025.1508808)] [Medline: [40432799](https://pubmed.ncbi.nlm.nih.gov/40432799/)]
45. Tariq MU, Sergio RP. Innovative assessment techniques in physical education: exploring technology-enhanced and student-centered models for holistic student development. In: *Global Innovations in Physical Education and Health*. Hershey, PA: IGI Global; 2025:85-112.
46. Fernandez-Rio J, Iglesias D. What do we know about pedagogical models in physical education so far? An umbrella review. *Phys Educ Sport Pedagogy* 2022;29(2):190-205. [doi: [10.1080/17408989.2022.2039615](https://doi.org/10.1080/17408989.2022.2039615)]
47. Sun Z, Yuan Y, Xiong X, Meng S, Shi Y, Chen A. Predicting academic achievement from the collaborative influences of executive function, physical fitness, and demographic factors among primary school students in China: ensemble learning methods. *BMC Public Health* 2024;24(1):274 [FREE Full text] [doi: [10.1186/s12889-024-17769-7](https://doi.org/10.1186/s12889-024-17769-7)] [Medline: [38263081](https://pubmed.ncbi.nlm.nih.gov/38263081/)]
48. Wang X, Yang L, Chen K, Zheng Y. Understanding teacher emotional exhaustion: exploring the role of teaching motivation, perceived autonomy, and teacher-student relationships. *Front Psychol* 2023;14:1342598 [FREE Full text] [doi: [10.3389/fpsyg.2023.1342598](https://doi.org/10.3389/fpsyg.2023.1342598)] [Medline: [38259554](https://pubmed.ncbi.nlm.nih.gov/38259554/)]
49. Faber TJE, Dankbaar MEW, van den Broek WW, Bruinink LJ, Hogeveen M, van Merriënboer JGG. Effects of adaptive scaffolding on performance, cognitive load and engagement in game-based learning: a randomized controlled trial. *BMC Med Educ* 2024;24(1):943 [FREE Full text] [doi: [10.1186/s12909-024-05698-3](https://doi.org/10.1186/s12909-024-05698-3)] [Medline: [39210381](https://pubmed.ncbi.nlm.nih.gov/39210381/)]
50. Xu L, Zou X, Hou Y. Effects of feedback visualisation of peer - assessment on pre - service teachers' data literacy, learning motivation, and cognitive load. *J Comput Assist Learn* 2024;40(4):1447-1462. [doi: [10.1111/jcal.12955](https://doi.org/10.1111/jcal.12955)]

51. Carcelén-Fraile MDC. Active gamification in the emotional well-being and social skills of primary education students. *Educ Sci* 2025;15(2):212. [doi: [10.3390/educsci15020212](https://doi.org/10.3390/educsci15020212)]
52. Hsia LH, Hwang GJ, Hwang JP. AI-facilitated reflective practice in physical education: an auto-assessment and feedback approach. *Interact Learn Environ* 2023;32(9):5267-5286. [doi: [10.1080/10494820.2023.2212712](https://doi.org/10.1080/10494820.2023.2212712)]
53. Kam AHT, Umar IN. Fostering autonomous motivation: a deeper evaluation of gamified learning. *J Comput High Educ* 2023;36(2):368-388. [doi: [10.1007/s12528-023-09358-1](https://doi.org/10.1007/s12528-023-09358-1)]
54. De Meester A, Galle J, Soenens B, Haerens L. Perseverance in motor tasks: the impact of different types of positive feedback. *Phys Educ Sport Pedagogy* 2022;29(2):221-234. [doi: [10.1080/17408989.2022.2054969](https://doi.org/10.1080/17408989.2022.2054969)]
55. Castle N, Howells K. A model for effective learning in competition: a pedagogical tool to enhance enjoyment and perceptions of competency in physical education lessons for young children. *Children (Basel)* 2024;11(1):111 [FREE Full text] [doi: [10.3390/children11010111](https://doi.org/10.3390/children11010111)] [Medline: [38255424](https://pubmed.ncbi.nlm.nih.gov/38255424/)]
56. Gan Y, Yang Y, Wu Y, Li H, Zhang Y. The effects of teacher persuasion on Chinese junior high school students' math self-efficacy: the multiple mediation of intrinsic value and math anxiety. *Curr Psychol* 2025;44(5):3956-3968. [doi: [10.1007/s12144-025-07381-2](https://doi.org/10.1007/s12144-025-07381-2)]
57. Sal-de-Rellán A, Hernández-Suárez Á, Hernaiz-Sánchez A. Gamification and motivation in adolescents. Systematic review from physical education. *Front Psychol* 2025;16:1575104 [FREE Full text] [doi: [10.3389/fpsyg.2025.1575104](https://doi.org/10.3389/fpsyg.2025.1575104)] [Medline: [40196203](https://pubmed.ncbi.nlm.nih.gov/40196203/)]
58. Spaho E, Çiço B, Shabani I. IoT integration approaches into personalized online learning: systematic review. *Computers* 2025;14(2):63. [doi: [10.3390/computers14020063](https://doi.org/10.3390/computers14020063)]
59. Strielkowski W, Grebennikova V, Lisovskiy A, Rakhimova G, Vasileva T. AI-driven adaptive learning for sustainable educational transformation. *Sustain Dev* 2024;33(2):1921-1947. [doi: [10.1002/sd.3221](https://doi.org/10.1002/sd.3221)]
60. Saini M, Sengupta E, Sharma N. AI-assisted hand gestures based smart feedback system for educators. *Educ Inf Technol* 2025;30(10):13281-13308 [FREE Full text] [doi: [10.1007/s10639-025-13358-2](https://doi.org/10.1007/s10639-025-13358-2)]
61. Lin H, Chen Q. Artificial intelligence (AI) -integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes. *BMC Psychol* 2024;12(1):487 [FREE Full text] [doi: [10.1186/s40359-024-01979-0](https://doi.org/10.1186/s40359-024-01979-0)] [Medline: [39285268](https://pubmed.ncbi.nlm.nih.gov/39285268/)]
62. Fang W, Fan S, Zheng H, Fang Z, You Y, Yin B, et al. How to improve mental health in the older adults through AI-enhanced physical activity: an emerging research topic. *Humanit Soc Sci Commun* 2025;12(1):1-15. [doi: [10.20944/preprints202502.2037.v1](https://doi.org/10.20944/preprints202502.2037.v1)]
63. Singh B, Bennett H, Miatke A, Dumuid D, Curtis R, Ferguson T, et al. Effectiveness of exercise for improving cognition, memory and executive function: a systematic umbrella review and meta-meta-analysis. *Br J Sports Med* 2025;59(12):866-876 [FREE Full text] [doi: [10.1136/bjsports-2024-108589](https://doi.org/10.1136/bjsports-2024-108589)] [Medline: [40049759](https://pubmed.ncbi.nlm.nih.gov/40049759/)]
64. Furley P, Schütz LM, Wood G. A critical review of research on executive functions in sport and exercise. *Int Rev Sport Exerc Psychol* 2023;18(1):316-344. [doi: [10.1080/1750984x.2023.2217437](https://doi.org/10.1080/1750984x.2023.2217437)]
65. Gkintoni E, Vantaraki F, Skoulidi C, Anastassopoulos P, Vantarakis A. Gamified health promotion in schools: the integration of neuropsychological aspects and CBT-a systematic review. *Medicina (Kaunas)* 2024;60(12):2085 [FREE Full text] [doi: [10.3390/medicina60122085](https://doi.org/10.3390/medicina60122085)] [Medline: [39768964](https://pubmed.ncbi.nlm.nih.gov/39768964/)]
66. Nguyen-Viet B, Nguyen-Duy C, Nguyen-Viet B. How does gamification affect learning effectiveness? The mediating roles of engagement, satisfaction, and intrinsic motivation. *Interact Learn Environ* 2024;33(3):2635-2653. [doi: [10.1080/10494820.2024.2414356](https://doi.org/10.1080/10494820.2024.2414356)]
67. Mercan G, Selçuk ZV. Investigating the impact of game-based learning and gamification strategies in physical education: a comprehensive systematic review. *JIETP* 2024;6(1):1-14. [doi: [10.47157/jietp.1389843](https://doi.org/10.47157/jietp.1389843)]
68. Sañudo B, Sanchez-Trigo H, Domínguez R, Flores-Aguilar G, Sánchez-Oliver A, Moral JE, et al. A randomized controlled mHealth trial that evaluates social comparison-oriented gamification to improve physical activity, sleep quantity, and quality of life in young adults. *Psychol Sport Exerc* 2024;72:102590. [doi: [10.1016/j.psychsport.2024.102590](https://doi.org/10.1016/j.psychsport.2024.102590)] [Medline: [38218327](https://pubmed.ncbi.nlm.nih.gov/38218327/)]
69. Zha H, Li W, Wang W, Xiao J. The paradox of AI empowerment in primary school physical education: why technology may hinder, not help, teaching efficiency. *Behav Sci (Basel)* 2025;15(2):240 [FREE Full text] [doi: [10.3390/bs15020240](https://doi.org/10.3390/bs15020240)] [Medline: [40001871](https://pubmed.ncbi.nlm.nih.gov/40001871/)]
70. Iglesias D, Fernandez-Rio J. A model fidelity check in cooperative learning research in physical education. *Phys Educ Sport Pedagogy* 2024:1-16 [FREE Full text] [doi: [10.1080/17408989.2024.2342831](https://doi.org/10.1080/17408989.2024.2342831)]
71. Deci EL, Ryan RM. *Intrinsic Motivation and Self-Determination in Human Behavior*. Berlin, Germany: Springer Science & Business Media; 2013.
72. Kassenkhan AM, Moldagulova AN, Serbin VV. Gamification and artificial intelligence in education: a review of innovative approaches to fostering critical thinking. *IEEE Access* 2025;13:98699-98728. [doi: [10.1109/access.2025.3576147](https://doi.org/10.1109/access.2025.3576147)]
73. Chen CH, Chang CL. Effectiveness of AI-assisted game-based learning on science learning outcomes, intrinsic motivation, cognitive load, and learning behavior. *Educ Inf Technol* 2024;29(14):18621-18642. [doi: [10.1007/s10639-024-12553-x](https://doi.org/10.1007/s10639-024-12553-x)]
74. Taggart J, Wheeler LB. Collaborative learning as constructivist practice: an exploratory qualitative descriptive study of faculty approaches to student group work. *Act Learn High Educ* 2023;26(1):59-76. [doi: [10.1177/14697874231193938](https://doi.org/10.1177/14697874231193938)]

75. Schrank FA, Wendling, BJ. The woodcock-johnson III normative update: tests of cognitive abilities and tests of achievement. In: Flanagan DP, Harrison PL, editors. *Contemporary Intellectual Assessment: Theories, Tests, and Issues*. New York, NY: The Guilford Press; 2012:297-335.
76. 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 6-9, 2014; Waikoloa, HI p. 3025-3034. [doi: [10.1109/hicss.2014.377](https://doi.org/10.1109/hicss.2014.377)]
77. Kendzierski D, DeCarlo KJ. Physical activity enjoyment scale: two validation studies. *J Sport Exerc Psychol* 1991;13(1):50-64. [doi: [10.1123/jsep.13.1.50](https://doi.org/10.1123/jsep.13.1.50)]
78. Laal M, Ghodsi SM. Benefits of collaborative learning. *Procedia Soc Behav Sci* 2012;31:486-490. [doi: [10.1016/j.sbspro.2011.12.091](https://doi.org/10.1016/j.sbspro.2011.12.091)]
79. Belmont M, Skinner EA, Wellborn J, Connell J. Teacher as Social Context Questionnaire. *APA PsycNet*. 1988. URL: <https://psycnet.apa.org/doiLanding?doi=10.1037%2F10488-000> [accessed 2026-01-24]
80. 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;60(1):48-58. [doi: [10.1080/02701367.1989.10607413](https://doi.org/10.1080/02701367.1989.10607413)] [Medline: [2489825](https://pubmed.ncbi.nlm.nih.gov/2489825/)]
81. Ifenthaler D, Eseryel D, Ge X. *Assessment in Game-Based Learning*. New York, NY: Springer; 2012:1-8.
82. Caspari-Sadeghi S. Artificial intelligence in technology-enhanced assessment: a survey of machine learning. *J Educ Technol Syst* 2022;51(3):372-386. [doi: [10.1177/00472395221138791](https://doi.org/10.1177/00472395221138791)]
83. *Publication Manual of the American Psychological Association, Seventh Edition*. Washington, DC: American Psychological Association; 2020.
84. Criollo-C S, Guerrero-Arias A, Guña-Moya J, Samala A, Luján-Mora S. Towards sustainable education with the use of mobile augmented reality in early childhood and primary education: a systematic mapping. *Sustainability* 2024;16(3):1192. [doi: [10.3390/su16031192](https://doi.org/10.3390/su16031192)]
85. Falola H, Joel OO, Akinbode J, Ojebola O. Emotional intelligence and academic engagement of graduate students: the moderating role of digital learning support. *Afr J Econ Manag Stud* 2025:1-20 [FREE Full text] [doi: [10.1108/ajems-03-2024-0203](https://doi.org/10.1108/ajems-03-2024-0203)]
86. Gkintoni E, Vantaraki F, Skoulidi C, Anastassopoulos P, Vantarakis A. Promoting physical and mental health among children and adolescents via gamification-a conceptual systematic review. *Behav Sci (Basel)* 2024;14(2):102 [FREE Full text] [doi: [10.3390/bs14020102](https://doi.org/10.3390/bs14020102)] [Medline: [38392455](https://pubmed.ncbi.nlm.nih.gov/38392455/)]
87. Chaiyarat K. Enhancing creative problem solving and learning motivation in social studies classrooms with gamified cooperative learning. *Think Ski Creat* 2024;54:101616. [doi: [10.1016/j.tsc.2024.101616](https://doi.org/10.1016/j.tsc.2024.101616)]
88. Aibar-Almazán A, Castellote-Caballero Y, Carcelén-Fraile MDC, Rivas-Campo Y, González-Martín AM. Gamification in the classroom: Kahoot! As a tool for university teaching innovation. *Front Psychol* 2024;15:1370084 [FREE Full text] [doi: [10.3389/fpsyg.2024.1370084](https://doi.org/10.3389/fpsyg.2024.1370084)] [Medline: [38646120](https://pubmed.ncbi.nlm.nih.gov/38646120/)]
89. Trajkovski G, Hayes H. AI-assisted formative assessment and feedback. In: *AI-Assisted Assessment in Education*. Cham: Palgrave Macmillan; 2025:283-312.
90. To J. Feedback partnerships: strengthening students' proactive reciprocity through co-creating dialogic feedback. *High Educ Res Dev* 2025;44(4) [FREE Full text] [doi: [10.1080/07294360.2024.2445570](https://doi.org/10.1080/07294360.2024.2445570)]
91. Cha W, Hong M, Glassman M, Anderman EM, Lin T. The roles of technology efficacy and networking agency in elementary students' engagement in online and face-to-face technology-mediated learning. *Brit J Educational Tech* 2025;56(6):2623-2646 [FREE Full text] [doi: [10.1111/bjet.13597](https://doi.org/10.1111/bjet.13597)]
92. Ryan RM, Deci EL. Self-determination theory. In: Maggino F, editor. *Encyclopedia of Quality of Life and Well-Being Research*. Cham: Springer; 2023.
93. Chen C. Exploring the impact of acute physical activity on creative thinking: a comprehensive narrative review with a focus on activity type and intensity. *Discov Psychol* 2024;4(1):3. [doi: [10.1007/s44202-024-00114-9](https://doi.org/10.1007/s44202-024-00114-9)]
94. Chen BC. Using Vygotsky's sociocultural theory to explore ethnic cultural representation in Taiwanese preschool children's play. *Front Educ* 2025;10:1569322. [doi: [10.3389/educ.2025.1569322](https://doi.org/10.3389/educ.2025.1569322)]
95. Qi S, Derakhshan A. Technology-based collaborative learning: EFL learners' social regulation and modifications in their academic emotions and academic performance. *Educ Inf Technol* 2024;30(7):8611-8636. [doi: [10.1007/s10639-024-13167-z](https://doi.org/10.1007/s10639-024-13167-z)]
96. Banik BG, Gullapelly A. AI-powered gamification and interactive learning tools for enhancing student engagement. In: *InDriving Quality Education Through AI and Data Science*. Hershey, PA: IGI Global Scientific Publishing; 2025:283-310.
97. Dinh CT, Phuong HY. MOOC learners' perspectives of the effects of self-regulated learning strategy intervention on their self-regulation and speaking performance. *Cogent Education* 2024;11(1):2378497. [doi: [10.1080/2331186x.2024.2378497](https://doi.org/10.1080/2331186x.2024.2378497)]
98. Suresh Babu S, Dhakshina Moorthy A. Application of artificial intelligence in adaptation of gamification in education: a literature review. *Comp Applic In Engineering* 2023;32(1):e22683. [doi: [10.1002/cae.22683](https://doi.org/10.1002/cae.22683)]
99. Ng DTK, Xinyu C, Leung JKL, Chu SKW. Fostering students' AI literacy development through educational games: AI knowledge, affective and cognitive engagement. *Computer Assisted Learning* 2024;40(5):2049-2064. [doi: [10.1111/jcal.13009](https://doi.org/10.1111/jcal.13009)]

100. Shalgimbekova K, Smagliy T, Kalimzhanova R, Suleimenova Z. Innovative teaching technologies in higher education: efficiency and student motivation. *Cogent Education* 2024;11(1):2425205. [doi: [10.1080/2331186x.2024.2425205](https://doi.org/10.1080/2331186x.2024.2425205)]
101. Shekh-Abed A. Metacognitive self-knowledge and cognitive skills in project-based learning of high school electronics students. *Eur J Eng Educ* 2024;50(1):214-229. [doi: [10.1080/03043797.2024.2374479](https://doi.org/10.1080/03043797.2024.2374479)]
102. Barz N, Benick M, Dörrenbächer-Ulrich L, Perels F. The effect of digital game-based learning interventions on cognitive, metacognitive, and affective-motivational learning outcomes in school: a meta-analysis. *Rev Educ Res* 2023;94(2):193-227. [doi: [10.3102/00346543231167795](https://doi.org/10.3102/00346543231167795)]
103. Hsia LH, Lin YN, Lin CH, Hwang GJ. Effectiveness of gamified intelligent tutoring in physical education through the lens of self-determination theory. *Comput Educ* 2025;227:105212. [doi: [10.1016/j.compedu.2024.105212](https://doi.org/10.1016/j.compedu.2024.105212)]
104. Bhardwaj V, Zhang S, Tan YQ, Pandey V. Redefining learning: student-centered strategies for academic and personal growth. *Front. Educ* 2025;10:1518602. [doi: [10.3389/educ.2025.1518602](https://doi.org/10.3389/educ.2025.1518602)]
105. Plass JL, Froehlich F. The future of personalized learning with AI. *Learn Individ Differ* 2025;124:102813. [doi: [10.1016/j.lindif.2025.102813](https://doi.org/10.1016/j.lindif.2025.102813)]
106. Huseinović L. The effects of gamification on student motivation and achievement in learning english as a foreign language in higher education. *MAPEH* 2023;4(1):10-36. [doi: [10.53880/2744-2373.2023.4.10](https://doi.org/10.53880/2744-2373.2023.4.10)]
107. Kalogiannakis M, Papadakis S, Zourmpakis A. Gamification in science education. A systematic review of the literature. *Educ Sci* 2021;11(1):22. [doi: [10.3390/educsci11010022](https://doi.org/10.3390/educsci11010022)]
108. Ferriz-Valero A, Østerlie O, García Martínez S, García-Jaén M. Gamification in physical education: evaluation of impact on motivation and academic performance within higher education. *Int J Environ Res Public Health* 2020;17(12):4465 [FREE Full text] [doi: [10.3390/ijerph17124465](https://doi.org/10.3390/ijerph17124465)] [Medline: [32575919](https://pubmed.ncbi.nlm.nih.gov/32575919/)]
109. Sotos-Martínez VJ, Ferriz-Valero A, García-Martínez S, Tortosa-Martínez J. The effects of gamification on the motivation and basic psychological needs of secondary school physical education students. *Phys Educ Sport Pedagogy* 2022;29(2):160-176. [doi: [10.1080/17408989.2022.2039611](https://doi.org/10.1080/17408989.2022.2039611)]
110. Biswas U, Bhattacharya S. ML-based intelligent real-time feedback system for blended classroom. *Educ Inf Technol* 2023;29(4):3923-3951. [doi: [10.1007/s10639-023-11949-5](https://doi.org/10.1007/s10639-023-11949-5)]
111. Qadir HM, Khan RA, Rasool M, Sohaib M, Shah MA, Hasan MJ. An adaptive feedback system for the improvement of learners. *Sci Rep* 2025;15(1):17242 [FREE Full text] [doi: [10.1038/s41598-025-01429-w](https://doi.org/10.1038/s41598-025-01429-w)] [Medline: [40383780](https://pubmed.ncbi.nlm.nih.gov/40383780/)]

Abbreviations

- AI:** artificial intelligence
AVE: average variance extracted
CFA: confirmatory factor analysis
CFI: comparative fit index
CLT: cognitive load theory
CR: composite reliability
HTMT: heterotrait-monotrait
PE: physical education
RMSEA: root-mean-square error of approximation
SDT: self-determination theory
SEM: structural equation modeling
TLI: Tucker-Lewis Index

Edited by S Brini; submitted 22.Jul.2025; peer-reviewed by YC Loh, FMJM Shamrat; comments to author 26.Oct.2025; accepted 06.Jan.2026; published 02.Feb.2026.

Please cite as:

Zhang J, Oh SS, Xu Y

Gamified Physical Education and Cognitive Performance Among Chinese Secondary School Students: Cross-Sectional Moderation Mediation Study

JMIR Serious Games 2026;14:e81086

URL: <https://games.jmir.org/2026/1/e81086>

doi: [10.2196/81086](https://doi.org/10.2196/81086)

PMID:

©Jingtong Zhang, Sae-Sook Oh, Yaru Xu. Originally published in JMIR Serious Games (<https://games.jmir.org>), 02.Feb.2026. 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.

Original Paper

Effects of Exergaming Tennis on Players' Tennis Skills and Mental State Compared to Regular Tennis in Adult Players: Quasi-Experimental Study

Jia Sheng Ngo¹, MMedSc; See Ziau Hoe¹, PhD; Maziah Mat Rosly¹, PhD

Department of Physiology, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia

Corresponding Author:

Maziah Mat Rosly, PhD

Department of Physiology

Faculty of Medicine

Universiti Malaya

Kuala Lumpur, 50603

Malaysia

Phone: 60 379674924

Email: mazahmr@um.edu.my

Abstract

Background: Exergaming, which combines physical activity with interactive gaming, has been shown to improve motor skills and fitness. However, exergaming's potential in complex, open-skill sports such as tennis, which require real-time coordination, decision-making, and technical precision, remains underexplored. Furthermore, only a few studies have evaluated the impact of exergaming on both technical skill development and psychological outcomes such as motivation and confidence, especially among novice players. This study addresses these gaps by comparing the combination of exergame-based tennis training and on-court tennis training (EBTT+OCTT) with on-court tennis training alone (OCTT×2) in improving technical skills, grip strength, confidence, and motivation.

Objective: This study aims to assess the effect of exergaming tennis in developing tennis technical skills, grip strength, confidence level, and motivation level.

Methods: In a 12-week quasi-experimental trial, 66 novices were randomized to either the EBTT+OCTT or the OCTT×2 group. Participants consisted of 22 males and 44 females, with the mean BMI and age of approximately 22 (SD 2.9) kg/m² and 26 (SD 7) years, respectively. The EBTT+OCTT group had 1 weekly session of OCTT and 1 session of exergaming training using Virtual Tennis (PlayStation 3 Move), while the OCTT×2 group completed 2 weekly on-court sessions. Outcomes included tennis technical skills such as groundstroke technique, service speed, and service accuracy (assessed using the Hewitt tennis test), grip strength (using a handheld dynamometer), confidence (using the Sports Confidence Inventory), and motivation (using the Sport Motivation Scale). Mean differences (MDs) between pre- and postintervention were collected and analyzed using repeated-measures ANOVA ($\alpha=.05$) and Pearson correlation analysis. Effect sizes were calculated using partial eta squared (η^2), where values ≥ 0.14 indicated large effects.

Results: After 12 weeks of training, both EBTT+OCTT and OCTT×2 groups showed significant improvements ($P<.001$) in tennis technical skills (MD=17.06-22.62), grip strength (MD=9.59-11.04 kg), and confidence levels (MD=23.29-26.28). These outcome measures have large effect sizes ($\eta^2=0.84-0.92$); however, they did not significantly differ ($P>.05$) across the groups when compared, with $P_{grip\ strength}=.24$, $P_{hit}=.97$, $P_{accuracy}=.86$, $P_{speed}=.72$, and $P_{confidence}=.31$. In terms of motivation, EBTT+OCTT retained intrinsic motivation (IM) better than OCTT×2, with significant reductions in IM, mainly IM-to-know (MD=7, SD 2.95) and IM-to-accomplish (MD=5, SD 3.77) observed in the OCTT×2 group ($P<.001$). Grip strength, confidence, and motivation levels (except amotivation) showed positive correlations with tennis technical skills ($r=0.39-0.80$).

Conclusions: EBTT+OCTT and OCTT×2 significantly improve tennis skills and confidence levels in novice players, although no significant differences were found between the two. However, EBTT+OCTT appeared to better sustain IM. Thus, EBTT+OCTT may serve as a supplementary tool for novice players to better learn tennis.

(JMIR Serious Games 2026;14:e73732) doi:[10.2196/73732](https://doi.org/10.2196/73732)

KEYWORDS

exergaming; tennis; sports skills; psychology; sport training

Introduction

Technological innovations have facilitated the integration of physical activity through electronic platforms, particularly via exergaming—interactive video games that require physical movement for gameplay [1]. Substantial evidence demonstrates the efficacy of exergaming in enhancing motor coordination, reaction time, and physical fitness across various populations, including pediatric, geriatric, rehabilitation, and special needs cohorts [2,3]. These systems present novel opportunities for incorporating exercise into daily life, prompting investigation of their applications in sports training. According to the existing literature, most research has focused primarily on endurance and closed-skill sports (eg, cycling, running, and rowing), as their predictable, self-paced nature translates effectively to virtual environments [4-6]. In contrast, open-skill sports such as tennis have received limited attention due to the technical challenges of accurately replicating their complex biomechanics and decision-making demands in digital platforms [4]. Tennis demands precise mind-body coordination, decision-making skills, and real-time reactions [7]. While exergaming shows potential for early skills acquisition, the current literature lacks studies examining its effects in skill-based sports.

Tennis is a complex, open-skill sport that demands not only physical skill and execution of techniques such as serving and groundstrokes but also strong psychological attributes, including confidence and motivation [7,8]. Technical proficiency, such as accurate and powerful strokes, is fundamental to performance, yet these motor skills do not operate in isolation, and physical capability, such as grip strength, may also play a role in ensuring these skills are executed appropriately [4,9,10]. In addition, psychological readiness can influence how consistently and effectively players perform under pressure, make decisions, and maintain effort during training or competition [7,11].

While physical outcomes have been moderately investigated, there is a lack of research examining psychological outcomes, particularly confidence and motivation, which are known to influence motor learning and sports performance [4,7,11]. Confidence, for example, plays a key role in motor learning and performance; players with greater sport confidence are more likely to take risks, recover from mistakes, and execute techniques successfully under stress [7,12]. Similarly, motivation, especially intrinsic motivation (IM), affects the degree of engagement, persistence, and enjoyment during practice, which can directly enhance motor learning and skill retention [7,13]. These psychological factors are not only relevant to sport performance but may also be interrelated with technical skill development, making it important to study both domains simultaneously [7,11-13]. For instance, a motivated and confident player may be more willing to engage in deliberate practice, which leads to better technical improvement. Only a limited number of studies have concurrently examined the interrelationship between technical and psychological skills in an exergaming context, leaving a gap in understanding how these domains interact to support performance and engagement.

Exergaming, by blending physical movement with gaming elements, has the potential to impact both domains [1]. It can foster motor skill repetition in a stimulating environment while simultaneously supporting motivation and self-confidence through goal-setting and feedback [3]. However, most previous exergaming research has focused either on general fitness or cognitive outcomes, with limited exploration of its impact on sport-specific technical skills and psychological outcomes in a combined manner.

While both exergaming and traditional on-field exercises aim to promote physical fitness, they differ considerably in context, delivery, and user engagement. Research has consistently demonstrated that on-field exercises such as tennis, football, and badminton contribute to improved cardiovascular fitness, enhanced muscular strength, better body composition, and overall physical and mental well-being [8,14]. However, despite their benefits, on-field exercises can present certain challenges, such as time constraints, limited access to appropriate facilities, weather conditions, and personal motivation [14-16]. Exergaming, in contrast, combines physical movement with interactive digital gameplay, making exercise more accessible, enjoyable, and gamified [3]. Research has shown that adding exergaming to exercise regimes such as badminton, tennis, and golf can help encourage social interaction, allowing users to participate in cooperative or competitive activities with family, friends, or online communities, which may improve the appeal and sustainability of physical activity [17,18].

Tennis was chosen in this study because there is limited research examining this sport. There was also a need to evaluate whether skills learned through exergaming can transfer to real tennis performance. Although understanding the rules and strategies of tennis is important, mastery of the sport requires repeated physical practice to develop technical proficiency, coordination, and real-time decision-making skills. Exergaming may help players develop motor skills to respond to service and shots through repetitive practice, combining mental, visual, and motor learning.

Thus, the main objectives of this research were to determine whether integrating exergame-based tennis training and on-court tennis training (EBTT+OCTT) can help improve tennis technical skills, grip strength, confidence, and motivation levels compared with only the traditional on-court tennis training (OCTT×2). In addition, this research aimed to determine whether there is any correlation between grip strength, confidence, and motivation, to better understand their interrelationship and impact on tennis performance.

Methods**Study and Trial Designs**

This quasi-experimental study was conducted in Malaysia, where participants were initially randomly allocated to either the EBTT+OCTT group or the OCTT×2 group as a parallel-group experimental study (aged 18 to 40 years, with

balanced randomization in a ratio of 1:1). However, about 9%-10% (OCTT×2 [n=3] and EBTT+OCTT [n=4]) of participants in each group requested to change their assigned allocation, and these requests were accommodated to minimize participant dropout. These requests were made before commencement of the experimental procedure, and no participants were switched during the experimental period.

Participants

Social media postings and word-of-mouth marketing were used to recruit research participants. Participants were predominantly urban residents, consisting of university students and young professionals from diverse socioeconomic strata. The sample size calculator program G*Power (version 3.1.9.4; Heinrich Heine University) was used to determine the sample size. A repeated measures ANOVA (within- and between-groups' interactions) was selected as the statistical test. The calculation

was based on a large effect size ($f=0.4$), a significance level of 0.05, and a power ($1-\beta$) of 0.95 [19]. The large effect size was chosen based on previous literature in sports and exergaming interventions reporting medium to large effect sizes for motor and psychological outcomes [20,21]. In addition, as this study was exploratory and involved multiple dependent variables with expected training-related gains, a conservative large effect size was used to ensure adequate statistical power. Two groups (EBTT+OCTT and OCTT×2) and 2 measurement time points (pre- and posttraining) were included, with an estimated correlation of 0.5 between repeated measures. The analysis determined that a minimum of 24 participants (12 per group) were required. To account for a potential dropout rate of 20%, 29 participants were required in total. This study recruited 66 participants in total, with 33 participants assigned to each group (Textbox 1).

Textbox 1. The inclusion and exclusion criteria of the study.

Inclusion criteria

- Individuals with less than 1 year of tennis experience.
- Aged between 18 and 40 years.
- No medical history or surgical history.
- Physical activity level below moderate as assessed by the International Physical Activity Questionnaire (IPAQ).

Exclusion criteria

- Individuals who are obese ($BMI > 30 \text{ kg/m}^2$).
- Individuals who do not understand English or Malay.

In the recruitment process, 74 individuals were screened, and 66 met the inclusion criteria. The others were excluded for reasons such as having over a year of tennis experience (3 individuals), a high International Physical Activity Questionnaire (IPAQ) score (2 individuals), a BMI exceeding 30 kg/m^2 (2 individuals), or cardiorespiratory conditions such as asthma (1 individual). The final sample (22 males and 44 females with a mean age of 26.3 years) represented the university-affiliated and local recreational tennis community, which is roughly 60% female and predominantly aged 18 to 35 years. The participants' weekly physical activity levels, averaging about 760 metabolic equivalents of task minutes on the IPAQ scale, aligned with those of a low-to-moderately active urban adult population. The low to moderate IPAQ score was selected because it can enhance the responsiveness to training interventions, as individuals with lower activity levels are more likely to demonstrate measurable improvements in physical and psychological outcomes [4,7]. Furthermore, exergaming interventions have been primarily designed for and shown to be particularly beneficial in low-active or novice populations, supporting the appropriateness of this inclusion criterion [7,22]. A standardized IPAQ score was also required to reduce variability in fitness and motor performance that might influence the responsiveness to training, particularly since the study targeted novice players [22].

Experimental Procedure

Sixty-six individuals were divided into two groups, with a total of 33 individuals in each of the OCTT x 2 and EBTT+OCTT groups. The participants in the EBTT+OCTT group practiced tennis on both the exergaming platform and the traditional tennis court, whereas the OCTT×2 group practiced only on the traditional tennis court. Before the training started, the participants' consent, demographic data, physical activity level via IPAQ, and baseline data such as sports motivation test, sports confidence inventory, grip strength via handheld dynamometer, and technical skills via the Hewitt tennis test were collected. All questionnaires and outcome measures were valid and reliable [22-26].

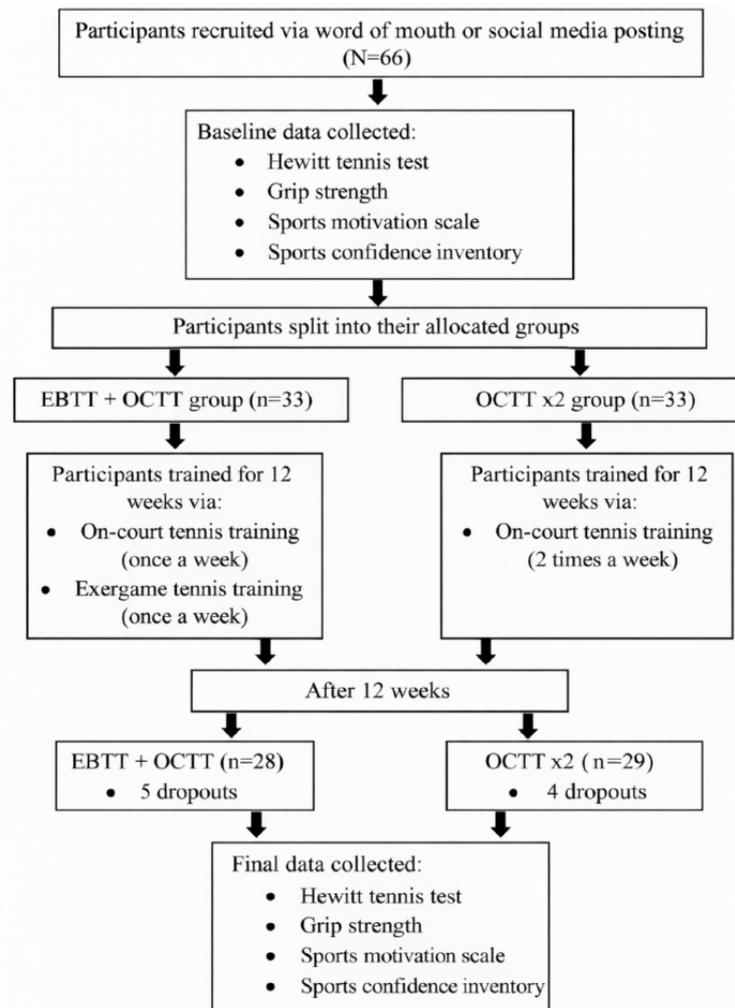
Both groups followed identical protocols for their on-court training sessions, which were conducted on tennis courts and with equipment meeting International Tennis Federation regulatory standards [27]. The sessions were done for 1.5 hours, comprising a warm-up period of 15 minutes, a 1-hour main practice, and a 15-minute cool-down. The exercise session duration and frequency were set following the American College of Sports Medicine guidelines [28,29]. The OCTT×2 group practiced on-court tennis twice a week, whereas the EBTT+OCTT group had one on-court and one exergaming session per week. In addition to the on-court tennis practice, participants in the EBTT+OCTT group also practiced tennis using a video game version of Virtual Tennis (Sports Champion

2, Zindagi Games, San Diego Studio) on a gaming device, the PlayStation 3 Move, for 1.5 hours.

Training included swing practices such as service, forehand, and backhand, and matches (either with a virtual avatar or among themselves). Matches were conducted between the 10th and 12th weeks. Both groups performed the task for 12 weeks

(24 practice sessions). There was no group switching among participants during the experimental period. Four individuals in the OCTT×2 group and 5 in the EBTT+OCTT group dropped out during the second and third weeks, respectively, during the 12-week training period. Figure 1 provides the flowchart of the experimental design.

Figure 1. Flowchart of the quasi-experimental study. EBTT: exergame-based tennis training; OCTT: on-court tennis training.



Outcome Measures

Measurement of Tennis Technical Skills

In this study, the Hewitt tennis test was used to assess the player's technical skills, such as service speed, service accuracy, and tennis strokes (forehand and backhand). This test was done pre- and post-12 weeks of practice. It has been demonstrated that this test is reliable and valid to conduct [25]. Participants in the tennis hit test aimed to score points by hitting balls into specific zones over the net, using forehand or backhand strokes during 10 trials. In the service accuracy and speed test, they served 10 balls into a target area, with accuracy scores based on placement and speed determined by the ball's bounce distance after hitting the service court. Each subtest score is:

- Groundstroke (hit) score: participants earned 1-5 points per stroke based on where the ball landed on a court target zone. The maximum score for this test was 50.

- Service accuracy score: participants were instructed to serve into marked target zones on the court. Points were awarded based on the accuracy of placement, with a maximum of 50 points.
- Service speed score: speed was estimated based on the bounce distance of the ball in the service court, categorized into point zones, with a maximum possible score of 50.

All outcomes were treated as continuous variables (range: 0-50) and used in the statistical analysis. The normal value given by Hewitt's Tennis test scores (score of 1-50) for beginners was graded from A (>18 points) to F (<3 points), where "A" represents an excellent score and "F" represents a very poor score [25]. While scoring the participant, the assessor was aware of the participant's group allocation.

Measurement of Grip Strength

Measuring grip strength with handheld dynamometers was used due to their high validity and reliability [30,31]. Standardized

protocols were followed to ensure accurate and comparable data. This typically involves a seated position with elbows flexed at 90 degrees, neutral wrists, and the handgrip aligned with the forearm [30,31]. Following the standardized protocol, 3 attempts were made for the dominant hand with rest intervals in between attempts. The highest value in kilograms (kg) was recorded and used for analysis. Grip strength was treated as a continuous variable in the statistical analysis. Measurements were taken pre- and post-12 weeks of training. The normal value for grip strength in Malaysia for males aged 18 to 54 ranged from 37 kg to 42 kg, whereas for females of a similar age, it ranged from 22 kg to 25 kg [32,33].

Measurement of Motivation and Confidence Level

The Sports Motivation Test (SMT) and the Sports Confidence Inventory (SCI) were used to measure motivation and confidence levels of participants when playing sports pre- and post-12 weeks of training. The SMT, developed by Pelletier et al [34], is a self-report questionnaire that assessed 7 types or subscales of motivation: IM-to-know, IM-to-accomplish, IM-to-experience stimulation, extrinsic motivation (EM)-identified, EM-introjected, EM-external regulation, and amotivation. Participants indicated the extent to which each statement reflected their motivation for engaging in sports, ranging from “not at all true” to “very true.” Each subscale score was calculated as the sum of all 4 items belonging to that subscale, resulting in 7 separate subscale scores ranging from 4 to 28. These subscale scores were used as separate continuous outcome variables in the analysis of motivation change over time and in correlation with tennis performance. The SMT is widely used in research settings to examine athletes’ motivation and understand the factors that drive their engagement in sports. High internal consistency is indicated by Cronbach alpha coefficients ranging from 0.70 to 0.90 [24,34].

The SCI, developed by Vealey [35], is a widely used questionnaire designed to assess athletes’ self-confidence in sports situations. Participants rate their level of confidence in each statement on a scale ranging from “strongly disagree” to “strongly agree.” The SCI measures both general and sport-specific self-confidence across various domains, such as skill execution, physical conditioning, and resiliency [23]. The total confidence score ranges from 8 to 72, with higher scores indicating greater sport-specific confidence. All 8 items were summed to calculate a total confidence score, which was used in pre-post comparisons and correlation analyses as a continuous variable. According to reliability assessments, the SCI has a high degree of internal consistency, with Cronbach alpha values between 0.70 and 0.90 [35]. Test-retest reliability has also been established, demonstrating stability over time.

Data Analysis

SPSS software, created by International Business Machines, was used to gather and analyze the data. Normality was evaluated using the Shapiro-Wilk test, which is highly effective at detecting nonnormality [36]. For normally distributed data, repeated measures ANOVA was performed. Final analysis was conducted per protocol outlined in the “Data Analysis” section, where participants were analyzed according to the group they actually participated in, instead of their original randomized

groups. In addition, a modified intention-to-treat (ITT) analysis was also performed, in which all participants who were randomized initially and completed the study were included in the analysis according to their originally assigned groups. This analytical approach excludes missing data from participants who dropped out of their respective interventions. To evaluate whether the treatment effects differed significantly between modified ITT analyses and per-protocol analyses, treatment effects from both approaches were compared. The primary comparison was based on the magnitude and 95% CIs of the treatment effects. Consistency between modified ITT and per-protocol analyses effects was interpreted as evidence of robustness. Effect sizes were evaluated using partial eta squared (η^2), with results classified as small (0.01-0.06), medium (0.06-0.14), or large (≥ 0.14) [37]. Partial eta squared (η^2) was used as the measure of effect size for the repeated-measures ANOVA results, as it is appropriate for estimating effect sizes in factorial and within-subject designs. Unlike Cohen *d*, which is suited for simple 2-group comparisons, η^2 provides a proportion of variance explained by each effect (main and interaction), making it more suitable for analyzing group \times time interactions in repeated-measures designs [37]. If the time (pre-post training) and group interaction are significant, posthoc pairwise comparisons were used; however, if only the time effect is significant, but there is no interaction between time and group, paired *t* tests were used to confirm improvement or mean difference (MD) in each group. A 95% CI of the MD was also calculated to determine the range within which the true difference between population means is likely to fall. The correlation between the changes (differences between pre- and postintervention) in players’ grip strength, confidence, and motivation levels with their tennis technical skills was analyzed using Pearson’s product-moment correlation coefficient (*r*). According to Cohen [38], the strength of the relationship is weak ($r=0.10$), moderate ($r=0.30$), or strong ($r=0.50$). Pearson *r* value was squared and multiplied by 100 (r^2) to determine the percentage of variation explained by 1 variable for another. In this study, r^2 was used to determine the percentage of variability in tennis technical skills accounted for by grip strength, confidence, and motivation levels. Findings with a $P < .05$ were considered statistically significant.

Ethical Considerations

This study involving human participants was reviewed and approved by the Universiti Malaya Research Ethics Committee, Malaysia. The ethics approval was granted on January 26, 2023 (UM.TNC2/UMREC_2286). Written informed consent was obtained from all participants before enrollment in the study. Participants were informed about the study objectives, procedures, potential risks and benefits, and their right to withdraw at any time without penalty. The study did not involve any secondary use of existing data; all data were collected prospectively as part of this trial. To ensure participant privacy and data confidentiality, all data collected remained anonymous by assigning unique participant codes. Personal identifiers were not linked to the study data. Electronic records were stored securely on password-protected devices. Participants did not receive any monetary compensation for their participation. However, they were provided with gift vouchers from local

grocery stores and restaurants and were allowed to retain access to the tennis training facilities during the study period at no cost, as a token of appreciation for their involvement. Written informed consents were also obtained from participants when their images were used in the papers. This trial was not preregistered, as it did not measure any clinical outcomes.

Results

Overview

The results are provided in 2 parts: the first part is the comparative analysis of the physical variables (tennis technical

skills and grip strength) and psychological variables (confidence and motivation levels) between EBTT+OCTT and OCTT×2, and the second part is the correlation analysis of the relationships between grip strength, confidence, and motivation with the tennis technical skills. All variables are presented as mean and SD for parametric data. Demographic data, including BMI, age, sex, and physical activity level, were collected. Independent *t* tests revealed no significant differences between groups in age, training experience, sex (male and female), BMI, and physical activity levels (Table 1).

Table 1. Summary of demographic data for the quasi-experimental study.

Parameter	EBTT+OCTT ^a (n=33)	OCTT×2 ^b (n=33)	<i>P</i> value
Age (y), mean (SD)	26.82 (7.08)	25.85 (6.15)	.64
Sex, n (%)			
Male	10 (30)	12 (36)	.68
Female	23 (70)	21 (64)	.79
Training experience (mo), mean (SD)	3.3 (2.44)	2.9 (3.12)	.86
Height (m), mean (SD)	1.64 (0.66)	1.65 (0.58)	.34
Weight (kg), mean (SD)	60.1 (7.8)	59.6 (7.7)	.54
BMI (kg/m ²), mean (SD)	22.35 (2.94)	21.89 (2.90)	.57
International Physical Activity Questionnaire (METs ^c), mean (SD)	808.94 (300.97)	712.35 (311.81)	.43
Dropouts, n (%)	5 (15)	4 (12)	— ^d

^aEBTT+OCTT: exergame-based tennis training with on-court tennis training.

^bOCTT: on-court tennis training.

^cMET: metabolic equivalent of task.

^dNot applicable.

Data collections were conducted from January to April 2024. During the 2nd and 3rd weeks of the study, 4 participants from the OCTT×2 group and 5 from the EBTT+OCTT group dropped out due to work schedules, health issues, transportation difficulties, or family emergencies (Table 1). The data of participants who dropped out were not used in the analysis. Shapiro-Wilk normality tests indicated that the differences between paired variables in the EBTT+OCTT and OCTT×2 groups were not statistically significant ($P>.05$), confirming that all paired variables were normally distributed.

In general, the treatment effect estimates from the modified ITT and per-protocol analyses were generally consistent for all outcome measures. The outcome measures showed overlapping 95% CIs and similar significance levels (*P*) between the 2 analyses, indicating no meaningful difference between the 2

analytical approaches. The table on the modified ITT analysis is provided in Multimedia Appendix 1.

Effects of EBTT+OCTT and OCTT×2 on Physical and Psychological Variables

Effect of EBTT+OCTT and OCTT×2 on Grip Strength

Both groups showed a significant improvement ($P<.001$) with a large effect size ($\eta^2=0.84$) from pre- to posttraining (Table 2). A total of 28 of 33 participants in the EBTT+OCTT group have an increase in grip strength by 1.5 times or 11.04 kg on average, while 29 of 33 participants in the OCTT×2 group improved their grip strength by 1.6 times or an average of 9.59 kg. However, grip strength showed no significant difference between groups ($P=.24$) with a small effect size ($\eta^2=0.03$).

Table 2. Comparison of grip strength, hit scores, accuracy scores, speed scores, and confidence level within and between groups.

Outcome measure and group	Pretest, mean (SD)	Posttest, mean (SD)	Value, MD ^a (95% CI; % changed)	Within-group <i>P</i> value	Between groups <i>P</i> value
Grip strength				<.001	.24
EBTT+OCTT ^b	21.68 (14.96)	32.71 (0.82)	11.04 (9.24-12.83; 51.10)		
OCTT ^c ×2	16.10 (13.97)	25.69 (0.68)	9.59 (7.86-11.31; 59.57)		
Hit scores				<.001	.97
EBTT+OCTT	4.32 (4.68)	21.46 (7.21)	17.14 (14.95-19.34; 400)		
OCTT×2	5.31 (5.41)	22.38 (9.23)	17.06 (14.27-19.86; 321.28)		
Accuracy scores				<.001	.86
EBTT+OCTT	4.21 (3.25)	26.54 (8.33)	22.32 (19.55-25.09; 530.17)		
OCTT×2	4.66 (3.30)	27.28 (7.51)	22.62 (20.38-24.86; 485.41)		
Speed scores				<.001	.72
EBTT+OCTT	3.50 (2.92)	21.25 (6.03)	17.75 (15.62-19.88; 507.14)		
OCTT×2	2.97 (2.16)	20.17 (6.34)	17.21 (14.97-19.44; 579.46)		
Confidence level				<.001	.31
EBTT+OCTT	37.04 (9.21)	60.32 (2.90)	23.29 (18.93-27.64; 62.88)		
OCTT×2	34.17 (7.31)	60.45 (3.80)	26.28 (22.19-30.36; 76.91)		

^aMD: mean difference.

^bEBTT+OCTT: exergame-based tennis training with on-court tennis training.

^cOCTT: on-court tennis training.

Effect of EBTT+OCTT and OCTT×2 on Technical Skills

Technical skills were categorized into 3 measures: forehand and backhand (hit) score, service accuracy, and service speed. Significant differences ($P<.001$) were found in all 3 measures, pre- and posttraining, for both groups, with large effect sizes: $\eta^2_{\text{hit}}=0.88$, $\eta^2_{\text{accuracy}}=0.92$, and $\eta^2_{\text{speed}}=0.91$. In the EBTT+OCTT group, 28 of 33 participants' hit score increased by 5 times (MD 17.14), service accuracy by 6.3 times (MD 22.32), and service speed by 6.1 times (MD 17.75). In the OCTT×2 group, 29 of 33 participants' hit score increased by 4.2 times (MD 17.06), service accuracy by 5.9 times (MD 22.62), and service speed by 6.8 times (MD 17.21). However, for all technical skill outcome measures, there were no significant differences between groups ($P_{\text{hit}}=.97$, $P_{\text{accuracy}}=.86$, and $P_{\text{speed}}=.72$), with small effect sizes ($\eta^2_{\text{hit}}=0.01$, $\eta^2_{\text{accuracy}}=0.03$, and $\eta^2_{\text{speed}}=0.01$) (Table 2).

Effect of EBTT+OCTT and OCTT×2 on Confidence Level

The main effect of time was significant ($P<.001$) with a large effect size ($\eta^2=0.84$), indicating that participants in both groups showed significant improvement from pre- to posttraining. In the EBTT+OCTT group, 28 of 33 participants improved their confidence level by 23.29 points (~1.6 times on average), whereas in the OCTT×2 group, 29 of 33 participants increased by 26.28 points (~1.8 times) compared with baseline. However, the interaction between pre-post training and the group was not significant ($P=.31$) with a small effect size of 0.02, suggesting

no overall difference between the EBTT+OCTT and OCTT×2 groups (Table 2).

Effect of EBTT+OCTT and OCTT×2 on Motivation Level

Overall, 28 of 33 participants in the EBTT+OCTT group showed a significant decrease in IM-to-accomplish, EM-identified, and EM-introjected. Whereas 29 of 33 participants in the OCTT×2 group, other than a significant increase in amotivation level, there were significant decreases in IM-to-know, IM-to-accomplish, EM-identified, EM-introjected, and EM-external regulation. In terms of the comparison between groups, the reduction of IM-to-know score and IM-to-accomplish score was greater in the OCTT×2 group than the EBTT+OCTT group, leading to a statistical significance, $P<.001$, when compared. A large effect size was also found, η^2 (IM-to-know)=0.25, and η^2 (IM-to-accomplish)=0.12. The only finding in EM that was statistically significant ($P<.001$) with a large effect size of $\eta^2=0.30$ was the decrease in the EM-external regulation score of OCTT×2 in comparison to EBTT+OCTT (Table 3).

Relationship Between Grip Strength, Confidence, and Motivation With Tennis Technical Skills

Relationship Between Grip Strength and Tennis Technical Skill

The bivariate correlation between grip strength and tennis hit score was significantly positive ($r=0.70$; $P<.001$). Grip strength was also significantly positively correlated with service accuracy ($r=0.45$; $P<.001$) and service speed ($r=0.42$; $P=.001$). The

proportion of variance (r^2) explained by grip strength for tennis hit score, service accuracy, and service speed was 0.50, 0.20, and 0.17, respectively.

Table 3. Comparison of different motivation levels within and between groups.

Outcome measures	Pretest, mean (SD)	Posttest, mean (SD)	Value, MD ^a (95% CI; % changed)	Within-group <i>P</i> value	Between groups <i>P</i> value
IM^b-to-know					<.001
EBTT+OCTT ^c	24.18 (2.92)	22.39 (4.12)	-1.79 (-3.85 to 0.28; 7.40)	.04	
OCTT ^d ×2	24.21 (2.53)	17.21 (2.88)	-7 (-8.43 to 5.57; 28.91)	<.001	
IM-to-accomplish					<.001
EBTT+OCTT	23.29 (2.57)	21.82 (2.21)	-1.46 (-2.90 to -0.03; 6.27)	.04	
OCTT×2	24.38 (2.21)	18.41 (2.21)	-5.97 (-7.09 to 4.85; 24.49)	<.001	
IM-to-experience stimulation					.66
EBTT+OCTT	24.50 (2.66)	24.18 (2.71)	-0.32 (-1.78 to 1.14; 1.31)	.66	
OCTT×2	24.21 (2.37)	24.31 (2.71)	0.10 (-1.26 to 1.47; 0.41)	.88	
EM^e-identified					.42
EBTT+OCTT	24.54 (2.62)	19.11 (2.50)	-5.43 (-6.72 to -4.13; 22.13)	<.001	
OCTT×2	24.14 (2.84)	19.41 (2.69)	-4.72 (-5.95 to -3.50; 19.55)	<.001	
EM-introjected					.93
EBTT+OCTT	17.25 (3.40)	14.64 (2.83)	-2.61 (-4.35 to -0.86; 15.13)	.01	
OCTT×2	17.17 (4.48)	14.69 (3.39)	-2.48 (-4.53 to -0.43; 14.44)	.02	
EM-external regulation					<.001
EBTT+OCTT	24.54 (2.83)	23.11 (2.60)	-1.43 (-2.99 to 0.13; 5.83)	.07	
OCTT×2	24.34 (2.79)	18.24 (2.40)	-6.10 (-7.29 to -4.92; 25.06)	<.001	
Amotivation					.91
EBTT+OCTT	13.68 (2.98)	20.76 (2.06)	7.08 (6.31 to 9.04; 51.75)	.06	
OCTT×2	13.86 (3.16)	21.66 (2.44)	7.79 (6.20 to 9.38; 56.20)	<.001	

^aMD: mean difference.

^bIM: intrinsic motivation.

^cEBTT+OCTT: exergame-based tennis training with on-court tennis training.

^dOCTT: on-court tennis training.

^eEM: extrinsic motivation.

Relationship Between Confidence Level and Tennis Technical Skill

For confidence, the bivariate correlation with tennis hit score ($r=0.70$; $P<.001$) and service speed ($r=0.33$; $P=.01$) was positive and significant. However, the correlation between confidence and service accuracy was not significant ($r=0.21$; $P=.12$). The coefficient of determination (r^2) for confidence with tennis hit score and service speed were 0.50 and 0.11, respectively.

Relationship Between Motivation Level and Tennis Technical Skill

IM had significant positive correlations with tennis hit score and service accuracy. The correlations with IM-to-know, IM-to-accomplish, and IM-to-experience stimulation were $r=0.60$, 0.45 , and 0.80 for tennis hit score, and $r=0.35$, 0.40 , and 0.40 for service accuracy. However, there was no significant

correlation between tennis service speed score and IM-to-know, IM-to-accomplish, and IM-to-experience, where $r=0.21$ ($P=.11$), 0.18 ($P=.18$), and 0.18 ($P=.18$), respectively. The r^2 values for tennis hit score and IM-to-know, IM-to-accomplish, and IM-to-experience stimulation were 0.36, 0.21, and 0.62, respectively. For service accuracy, the r^2 values were 0.12, 0.16, and 0.16.

EM showed significant correlations ($P<.001$) only with tennis hit score, with $r=0.60$ (EM-identified), $r=0.67$ (EM-introjected), and $r=0.64$ (EM-external regulation). There were no significant correlations with service speed ($P_{EM-identified}=.21$, $P_{EM-introjected}=.72$, and $P_{EM-external regulation}=.65$) or accuracy ($P_{EM-identified}=.86$, $P_{EM-introjected}=.26$, and $P_{EM-external regulation}=.45$). The r^2 values for tennis hit score and EM-identified, EM-introjected, and EM-external regulation were 0.36, 0.45, and 0.41, respectively.

Finally, amotivation was negatively correlated with tennis hit score ($r=-0.39$; $P<.001$), but there were no significant correlations with service speed ($P=.46$) or accuracy ($P=.60$). The r^2 between tennis hit score and amotivation was 0.15,

meaning that 15% of the variability in tennis hit scores could be predicted by amotivation. Table 4 provides the correlation between grip strength, confidence, and motivation levels with tennis technical skills.

Table 4. Correlation analysis (Pearson r and 2-tailed P value) between grip strength, confidence, and motivation levels with tennis technical skills.

Outcome measures	Hit scores	Accuracy scores	Speed scores
Grip strength			
r	0.70 ^a	0.45 ^a	0.42 ^a
P value	<.001	<.001	.001
Confidence level			
r	0.70 ^a	0.21	0.33 ^a
P value	<.001	.12	.01
IM^b-to-know			
r	0.60 ^a	0.35 ^a	0.21
P value	<.001	.01	.11
IM-to-accomplish			
r	0.45 ^a	0.40 ^a	0.18
P value	<.001	<.001	.18
IM-to-experience stimulation			
r	0.80 ^a	0.40 ^a	0.18
P value	<.001	<.001	.18
EM^c-identified			
r	0.60 ^a	0.17	-0.02
P value	<.001	.21	.86
EM-introjected			
r	0.67 ^a	0.05	-0.15
P value	<.001	.72	.26
EM-external regulation			
r	0.64 ^a	0.06	-0.10
P value	<.001	.65	.45
Amotivation			
r	-0.39 ^a	-0.07	-0.10
P value	<.001	.60	.46

^aThe correlation is significant at a significant level of .05 (2-tailed).

^bIM: intrinsic motivation.

^cEM: extrinsic motivation.

Discussion

Principal Findings

In general, both EBTT+OCTT and OCTT×2 groups demonstrated significant improvements ($P<.001$) in tennis technical skills, all with large effect sizes, though no significant between-group differences were found across hit scores ($P=.97$), service accuracy ($P=.86$), and service speed ($P=.72$). The

absence of significant between-group differences should not be interpreted as evidence of equivalence between the interventions, as this study was not designed or powered to assess equivalence or noninferiority. Grip strength increased significantly ($P<.001$) in both groups, with no significant interaction effects ($P=.24$). Confidence levels also improved significantly ($P<.001$) in both groups, but without significant between-group differences ($P=.31$). In terms of motivation, the EBTT+OCTT group showed significant decreases ($P<.001$) in EM-identified and

EM-introjected, while the OCTT×2 group showed significant reductions ($P<.001$) in IM-to-know, IM-to-accomplish, EM-identified, and EM-external regulation. Notably, the OCTT×2 group experienced significantly greater decreases in IM-to-know, IM-to-accomplish, and EM-external regulation compared to EBTT+OCTT, all with large effect sizes. Correlational analysis revealed strong positive relationships ($r>.50$) between hit skill and grip strength, confidence level, IM-to-know, IM-to-experience stimulation, and all EM subscales. However, a negative correlation was observed between hit skill and amotivation ($r=-0.39$).

Comparison With Previous Work

Effect of Exergaming on Grip Strength

This study found significant improvements in grip strength in both EBTT+OCTT and OCTT×2 groups. Grip strength increased by 1.5 times in the EBTT+OCTT group and by 1.6 times in the OCTT×2 group. These results align with previous research on sports training and muscular strength. The likely mechanism is the repeated high-intensity activation of the forearm and hand muscles required for racket control and stroke execution. A previous study and a review have corroborated that regular engagement in sports requiring sustained gripping actions and upper limb coordination, such as tennis, induces significant neuromuscular adaptations that manifest as measurable strength gains in distal upper extremity musculature [9,39]. The observed grip strength improvements in the EBTT+OCTT group (1.5 times) demonstrated comparability with findings from an electromyography biofeedback exergaming study, which reported a 0.17 times enhancement in grip strength following training [40]. The substantially greater improvements observed in this quasi-experimental study may be attributable to the longer intervention duration (12 weeks vs 2 weeks), allowing for more pronounced neuromuscular adaptation. These results support the theoretical framework suggesting that exergaming systems designed to replicate sport-specific movements can elicit muscular engagement patterns and strength development comparable to conventional training modalities [40]. While the between-group comparison revealed no statistically significant difference in grip strength improvement, the marginally lower gain in the EBTT+OCTT condition may reflect biomechanical differences in equipment characteristics. Specifically, the substantially lighter mass of the exergaming controller (90 g) compared to a standard tennis racket (280 g-300 g) potentially reduced the mechanical loading and strength demands during gameplay. Nevertheless, the significant improvements observed in the EBTT+OCTT group suggest that exergaming tennis, while not specifically designed for grip strengthening, can effectively enhance grip strength through repeated sport-specific movement patterns. These findings expand current understanding of EBTT+OCTT's therapeutic potential and its capacity to produce physiological adaptations similar to traditional training approaches.

Effect of Exergaming on Technical Skills

This study showed significant improvements in technical skill scores, with large effect sizes in both EBTT+OCTT and OCTT×2 groups. These results suggest that exergaming and traditional training are both effective for enhancing tennis skills.

Improvements were likely driven by repetitive practice, whether in virtual or real settings, which aligns with previous research on motor learning in neurological rehabilitation [41]. This reflects the motor learning principle of practice, where skills must be rehearsed correctly to be mastered [42].

In the EBTT+OCTT group, hit scores increased by 5.5 times, service accuracy by 6.3 times, and service speed by 6.1 times. In the OCTT×2 group, hit scores rose by 4.2 times, service accuracy by 5.9 times, and service speed by 6.8 times. These findings are consistent with earlier studies, showing that exergaming can enhance motor skills and physical performance comparably to on-court training [42-45]. For example, one study reported that both exergaming and on-court training improved tennis players' reaction time by 1.03 times, indicating that exergaming may help players prepare for incoming shots [43]. Similar results have been observed in children, where exergaming improved forehand and backhand skills by 1.5 times compared to 1.33 times in traditional training [44]. The greater improvement observed in this study (average skill increase of 5.5 times) compared with previous work (around 1.3 times) may be explained by the novice status of participants. Before the training, their baseline skills were poor, scoring below 6 of 50 on the Hewitt tennis test. However, with structured guidance and consistent practice, their skills improved dramatically over 12 weeks. This reflects the principle of specificity, where training effects are specific to the practiced activity [46]. As EBTT+OCTT replicates tennis movements, skill gains in the virtual setting were transferred to real-world performance.

Supporting evidence comes from studies in other domains. Exergaming-based balance training improved balance by 1.67 times, similar to conventional balance training [47]. Another study on adolescent tennis players reported a 0.38-times increase in Universal Tennis Rating when training combined with virtual reality and tablet-based cognitive drills was compared to regular training [45]. Collectively, these findings highlight that exergaming can be as effective as traditional training for enhancing technical performance. The shared success likely stems from common motor learning principles and practice specificity [46,47]. Finally, this study extends the literature by focusing on novice players. Exergaming provided an engaging, space-efficient alternative for skill development, requiring only about 4×3 meters of training space.

Effect of Exergaming on Confidence Level

The study found that both EBTT+OCTT and OCTT×2 significantly improved participants' confidence in tennis performance. Confidence increased by 1.6 times in the EBTT+OCTT group and by 1.8 times in the OCTT×2 group. These results align with the self-efficacy theory, which highlights mastery experiences as a core source of confidence [48]. In this trial, structured and repetitive practice enabled participants to progressively master tennis skills, enhancing their belief in their abilities. The large effect sizes further indicate that both training types were highly effective in fostering a sense of control over performance, consistent with self-efficacy principles [49].

Previous research supports this view, showing that repeated practice in real sports settings increased athletes' confidence

by 1.21 times, thus preparing them for competition [50]. Similarly, exergaming interventions have been shown to enhance confidence in other contexts, such as fall-prevention training for older adults, which raised confidence in avoiding falls by 1.05 times [51]. However, only a few studies have directly assessed exergaming's impact on sports confidence. This trial helps address that gap, showing that tennis-specific exergaming can enhance confidence comparably to on-court training. Notably, there was a study that found exergaming to be less effective than traditional training for psychological outcomes. The trial reported that while exergaming boosted enjoyment and engagement, it did not significantly improve confidence in novice athletes [52]. The difference may reflect variations in exergame type, intervention length, or participant skill level. This study used tennis-specific exergaming directly relevant to participants' skill goals, while the comparison study [52] used non-sport-specific games.

Overall, the comparable benefits of EBTT+OCTT and OCTT×2 suggest that exergaming can be a valuable addition to sports training, particularly where access to courts or equipment is limited. It may also provide a psychologically supportive environment for beginners or individuals with low self-confidence by reducing the pressures of real-world competition.

Effect of Exergaming on Motivation Level

In the EBTT+OCTT group, apart from an increase in amotivation levels, there was a decline in all other motivation categories. Significant decreases were observed only in IM-to-accomplish (by 0.09 times), EM-identified (by 0.22 times), and EM-introjected (by 0.15 times), with medium and small effect sizes, respectively. In the OCTT×2 group, aside from increases in amotivation and IM-to-experience, other motivation levels also declined. Significant decreases were noted in IM-to-know (by 0.28 times), IM-to-accomplish (by 0.24 times), EM-identified (by 0.20 times), and EM-external regulation (by 0.25 times), all with large effect sizes ($\eta^2=0.12-0.30$). These results suggest that both EBTT+OCTT and OCTT×2 may influence IM in complex ways. According to self-determination theory, IM is driven by the inherent satisfaction of performing an activity [53]. The observed decreases in IM components, such as IM-to-accomplish, could suggest that the initial novelty and enjoyment of mastering new skills may diminish over time as tasks become repetitive or as participants reach a plateau in skill development [53]. This aligns with past research showing that continuous, repetitive training can sometimes reduce IM, especially when the perceived challenge or novelty decreases [54]. This pattern is also supported by another study, where motivation level was reduced by 1.08 times in general, and amotivation increased by 3.69 times when the participants were provided with a wearable healthy lifestyle technology (Fitbit by Fitbit Inc) for 8 weeks [55]. These results may indicate that, over time, both intrinsic and EM decline. However, it should be noted that maintaining motivation is essential for sustaining motivation without the need to alter or add new training equipment. Thus, the results from this quasi-experimental study suggest that EBTT+OCTT was able to sustain IM more than OCTT×2.

The Interrelationship Between Grip Strength, Confidence Level, and Motivation Level Toward an Individual's Tennis Skills

The interaction between motivation, confidence, and grip strength appears central to enhancing tennis performance. In this study, grip strength showed strong positive correlations with tennis hit score ($r=0.70$), service accuracy ($r=0.45$), and service speed ($r=0.42$). This suggested that stronger grip strength contributed to more powerful and accurate strokes. Similar findings have been reported in previous studies, where grip-strengthening exercises improved forearm speed and service performance [10,20,56]. Stronger grip strength enhances the transfer of force from the lower body and core to the racket, leading to greater ball velocity and control. In this study, participants initially had low grip strength (mean=18.74 kg) and poor Hewitt's tennis test scores, often falling in the "D" or "F" range. Their early struggles in returning balls accurately highlighted the importance of grip strength for technical execution.

Exergaming, such as traditional training, improved grip strength in this study. The EBTT+OCTT group increased grip strength by 1.5 times, while the OCTT×2 group improved by 1.6 times. Although no significant difference was found between the groups, the repetitive upper-limb motions in exergaming likely contributed to strength gains. Past research supports this, showing that interactive games involving hand and arm actions can build muscular endurance and functional fitness [21,57,58]. The moderate correlation between grip strength and service speed ($r=0.42$) further supports its role in stroke performance. While this study did not establish equivalence between EBTT+OCTT and OCTT×2, the findings suggested that exergaming offered a meaningful, engaging way to strengthen tennis-related muscles.

Confidence also emerged as a critical factor. Posttraining, confidence levels rose significantly in both groups and were positively correlated with tennis hit scores ($r=0.70$) and service speed ($r=0.33$). Confidence helped participants take risks, handle pressure, and execute skills more effectively. Early in training, participants' low confidence coincided with poor accuracy and hesitancy. By the end of the training, confidence scores improved from 35 of 72 to 61 of 72, accompanied by fourfold increases in hit scores. Previous studies confirm that self-efficacy and confidence are linked to better performance, reduced anxiety, and resilience under stress [48,59-62]. Exergaming may support confidence development by providing progressive challenges and instant feedback, enabling athletes to see measurable progress. This study found confidence increased by 1.6 times in EBTT+OCTT and by 1.8 times in OCTT×2, showing that both methods are effective.

Motivation, particularly IM, also played a key role. Players with high IM were more committed to practice, strengthening both skills and confidence. In this study, IM correlated strongly with hit scores ($r=0.60$, 0.45, and 0.80 for IM-to-know, IM-to-accomplish, and IM-to-experience, respectively), while EM showed weaker links to performance. Although overall motivation decreased slightly over the 12 weeks, the decline was smaller in EBTT+OCTT, suggesting exergaming helped

sustain interest. This aligns with previous research showing that task-oriented environments emphasizing skill mastery foster stronger IM than competition-focused settings [63-66]. Exergaming creates such an environment by offering clear goals, instant feedback, and playful competition, all of which help maintain engagement.

Taken together, the results suggest that grip strength, confidence, and motivation interact in a reinforcing cycle. Stronger grip strength enhances stroke performance, which boosts confidence. Increased confidence encourages effort and persistence, feeding back into skill gains. Intrinsic motivation sustains this process by making practice enjoyable and rewarding. Exergaming supports this loop by providing variety, feedback, and achievable challenges, which help maintain both confidence and motivation. While not a full replacement for on-court training, EBTT+OCTT appears to be a valuable complement, especially for novices or those with limited access to facilities. Exergaming requires minimal space, engages players psychologically, and offers cost-effective benefits in both sporting and rehabilitation contexts [49]. Coaches and practitioners could integrate exergaming to enhance early skill acquisition, sustain motivation, and reduce barriers to participation.

Exergaming Studies in Other Sports and Psychological Domains

The results of this study are consistent with previous exergaming interventions across a range of sports and populations. For example, in a study involving junior athletes practicing basketball, handball, and volleyball, exergaming significantly improved reaction time over 3 months, indicating its utility in sport-specific skill enhancement [67]. Similarly, in a soccer-based exergaming study, players demonstrated improved reaction time and passing accuracy, with larger effects observed in novices compared to experienced athletes [68]. These findings support the results of this quasi-experimental study, whereby novice players may be particularly responsive to exergaming interventions, potentially due to a greater scope for technical and psychological adaptation.

From a psychological perspective, the results align with previous studies showing that exergaming enhances motivation, confidence, and enjoyment. A study found that adolescents who engaged in exergaming reported significantly greater IM and self-efficacy compared to traditional physical education groups [69,70]. Likewise, another study demonstrated that virtual cycling games increased motivation and adherence among older adults [71]. While the effect sizes in this quasi-experimental study were large in terms of motivational outcomes, this may reflect the unique engagement and gamified feedback inherent in exergaming platforms, which are particularly effective in novice or recreational populations. In contrast, elite athletes may show more modest gains due to ceiling effects and previous high-level training.

Limitations

This study has several limitations that should be acknowledged. First, the findings may only be generalizable to novice players, as the participants recruited had less than 1 year of tennis experience. This may limit the applicability of the results to

more experienced or elite players, who may respond differently to exergaming-based interventions. To address this, this quasi-experimental study standardized the training program and ensured participants had similar baseline experience levels. Future studies should explore the effects of exergaming among intermediate or advanced players to assess broader applicability.

Second, the dropout rate was approximately 12% in the OCTT×2 group and 15% in the EBTT+OCTT group. Although this attrition was within acceptable limits for a quasi-experimental design, it could have influenced group comparisons and reduced statistical power. Another key limitation was the group allocation switching that occurred before the start of the experiment, where 7 participants requested to join the group of their preference (OCTT×2 [n=3] and EBTT+OCTT [n=4]). This introduced self-selection bias and could potentially affect the internal validity of the study's results, especially for outcomes related to motivation. Furthermore, the per-protocol analysis of data reflected the partial allocation of participants' chosen group, as opposed to randomized groups. This limits the statistical strength and validity of the results reporting motivational outcomes typically associated with randomized designs. As such, findings about motivation in this study should be interpreted with caution, and future studies should consider more rigorous strategies to ensure the robustness of randomization. To mitigate this, the study recruited more participants than the minimum sample size required by power analysis to maintain sufficient statistical validity. Although participants who had requested the group change may have had pre-existing differences in motivation, their baseline levels remained similar compared to those who remained in their allocated groups. In addition, the consistency in results between the modified ITT and per-protocol analyses strengthens the validity of the findings, suggesting that the results are robust to missing data and protocol deviations. However, future studies should incorporate strategies such as participant incentives, participant blinding, or shorter program durations to enhance retention.

Third, the exergaming environment may not fully replicate the physical, cognitive, and environmental demands of real-life tennis. Factors such as court size, surface type, weather conditions, and opponent variability are absent in virtual simulations, potentially limiting the ecological validity of the findings. To address this limitation, EBTT+OCTT participants also received one on-court tennis session per week to complement the virtual training. However, further research should explore advanced platforms (for example, augmented or mixed reality) to better simulate real-world playing conditions.

Fourth, the sex distribution in the study sample was skewed, with approximately 63% female participants. This imbalance may introduce sex-related bias and limit the generalizability of the results across sexes. While the study analyzed group-level performance rather than sex-specific differences, future studies should aim for a more balanced male-to-female ratio or consider conducting sex-stratified analyses.

Fifth, although the study used validated tools to assess motivation, confidence, grip strength, and tennis skills, it relied

primarily on quantitative methods. This limit understanding of participant experiences and perspectives regarding the training. Incorporating qualitative methods, such as interviews or focus groups, in future studies could provide richer insight into how and why exergaming may influence motivation and performance. At the same time, the assessor was not blinded to group allocation due to practical constraints inherent in sport-specific field testing, which could introduce potential bias. The tennis skill assessments required direct interaction and real-time scoring, making blinding operationally difficult and largely unfeasible. To minimize bias, a single trained assessor was used to ensure consistent scoring procedures across all participants.

Finally, while no statistically significant differences were found between EBTT+OCTT and OCTT×2 for most outcome measures, this should not be interpreted as evidence of equivalence between the 2 training modalities. The study used standard hypothesis testing and was not designed or powered to formally assess equivalence or noninferiority. Therefore, the finding of nonsignificant *P* values reflects a failure to reject the null hypothesis, not confirmation that the interventions are equally effective. Future studies using equivalence or noninferiority trial designs may be warranted to further evaluate whether EBTT+OCTT can produce outcomes comparable to traditional training.

Future Directions

Future research should consider several aspects to build upon the findings of this study. First, as mentioned previously, studies should explore the effectiveness of EBTT+OCTT in populations with varying levels of tennis proficiency, including intermediate and elite players, to determine whether skill transfer and psychological benefits persist across experience levels. Second, longitudinal studies with extended follow-up periods are required to assess the sustainability of improvements in technical

skills, grip strength, motivation, and confidence beyond the 12-week intervention period. Third, more immersive technologies such as virtual reality or augmented reality could be integrated into exergaming systems to better replicate real-world conditions and improve ecological validity. Fourth, future trials should investigate how exergaming can be personalized based on individual learning styles, physical abilities, or motivational profiles, which may enhance engagement and outcomes. Fifth, incorporating qualitative methods could provide deeper insights into users' experiences and perceived barriers or facilitators to exergaming adoption. Finally, exergaming could be examined as a rehabilitation tool for individuals with musculoskeletal or neurological impairments, to determine its utility in clinical sports medicine or physiotherapy settings.

Conclusion

In conclusion, this quasi-experimental study demonstrated that both EBTT+OCTT and OCTT×2 significantly improved tennis technical skills, grip strength, and confidence level among novice players over 12 weeks. EBTT+OCTT also helped sustain participants' IM level. Although between-group differences were not significant for most outcomes, EBTT+OCTT was more effective in sustaining IM. The observed positive associations between tennis skills, grip strength, confidence, and motivation suggest a correlated process where improvements in one domain support gains in others. Overall, evidence points toward EBTT+OCTT as a feasible and engaging supplementary approach for developing both physical and psychological aspects of tennis performance in novice players. However, given the study's design and population, conclusions should be interpreted with caution and not generalized beyond novice recreational players. Further controlled studies are needed to confirm the broader applicability and long-term effects of exergaming in sport-specific training.

Acknowledgments

The content of the paper does not represent the opinion of the funder or the university itself. The funding sources did not influence any aspects of this study.

Funding

This research was funded by the Universiti Malaya Sustainable Development Centre grant LL2023ECO015 and the Universiti Malaya Faculty of Medicine Postgraduate Scholarship Scheme.

Data Availability

In accordance with the stipulations provided by the ethics governing committee (Universiti Malaya Ethics Committee), all data belongs to Universiti Malaya, and distribution of the data require previous permission from the committee. Study data can be provided upon reasonable request to the corresponding author.

Authors' Contributions

Conceptualization was led by JSN, with MMR and SZH contributing equally. Data curation and formal analysis were led by JSN, with MMR and SZH providing supporting contributions. Funding acquisition was led by MMR, with SZH contributing equally and JSN in a supporting role. Investigation, methodology, and project administration were led by JSN, with MMR and SZH providing support. Resources were managed by JSN (lead), with MMR and SZH contributing equally. Supervision was led by MMR, with SZH contributing equally and JSN supporting. Validation and visualization were led by JSN, with MMR and SZH

contributing equally. Writing of the original draft and review, and editing were led by JSN, with MMR and SZH providing supporting contributions.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Results before group reassignment for modified intention-to-treat analysis.

[\[DOCX File, 19 KB - games_v14i1e73732_app1.docx\]](#)

References

1. Sween J, Wallington S, Sheppard V, Taylor T, Llanos A, Adams-Campbell L. The role of exergaming in improving physical activity: a review. *J Phys Act Health* 2014;11(4):864-870 [FREE Full text] [doi: [10.1123/jpah.2011-0425](https://doi.org/10.1123/jpah.2011-0425)] [Medline: [25078529](https://pubmed.ncbi.nlm.nih.gov/25078529/)]
2. Mat Rosly M, Mat Rosly H, Davis Oam GM, Husain R, Hasnan N. Exergaming for individuals with neurological disability: a systematic review. *Disabil Rehabil* 2017;39(8):727-735. [doi: [10.3109/09638288.2016.1161086](https://doi.org/10.3109/09638288.2016.1161086)] [Medline: [27108475](https://pubmed.ncbi.nlm.nih.gov/27108475/)]
3. Benzing V, Schmidt M. The effect of exergaming on executive functions in children with ADHD: a randomized clinical trial. *Scand J Med Sci Sports* 2019;29(8):1243-1253 [FREE Full text] [doi: [10.1111/sms.13446](https://doi.org/10.1111/sms.13446)] [Medline: [31050851](https://pubmed.ncbi.nlm.nih.gov/31050851/)]
4. Guo W, Wang B, Smoter M, Yan J. Effects of open-skill exercises on cognition on community dwelling older adults: protocol of a randomized controlled trial. *Brain Sci* 2021;11(5):609 [FREE Full text] [doi: [10.3390/brainsci11050609](https://doi.org/10.3390/brainsci11050609)] [Medline: [34068726](https://pubmed.ncbi.nlm.nih.gov/34068726/)]
5. Martinez-Millana A, Michalsen H, Berg V, Anke A, Gil Martinez S, Muzny M, et al. Motivating physical activity for individuals with intellectual disability through indoor bike cycling and exergaming. *Int J Environ Res Public Health* 2022;19(5):2914 [FREE Full text] [doi: [10.3390/ijerph19052914](https://doi.org/10.3390/ijerph19052914)] [Medline: [35270607](https://pubmed.ncbi.nlm.nih.gov/35270607/)]
6. Khundam C, Noël F. A study of physical fitness and enjoyment on virtual running for exergames. *Int J Comput Games Technol* 2021;2021:1-16. [doi: [10.1155/2021/6668280](https://doi.org/10.1155/2021/6668280)]
7. Levac DE, Huber ME, Sternad D. Learning and transfer of complex motor skills in virtual reality: a perspective review. *J Neuroeng Rehabil* 2019;16(1):121 [FREE Full text] [doi: [10.1186/s12984-019-0587-8](https://doi.org/10.1186/s12984-019-0587-8)] [Medline: [31627755](https://pubmed.ncbi.nlm.nih.gov/31627755/)]
8. Biddle SJH, Gorely T, Faulkner G, Mutrie N. Psychology of physical activity: a 30-year reflection on correlates, barriers, and theory. *Int J Sport Exerc Psychol* 2022;21(1):1-14. [doi: [10.1080/1612197x.2022.2147261](https://doi.org/10.1080/1612197x.2022.2147261)]
9. Cronin J, Lawton T, Harris N, Kilding A, McMaster D. A brief review of handgrip strength and sport performance. *J Strength Cond Res* 2017;31(11):3187-3217. [doi: [10.1519/JSC.0000000000002149](https://doi.org/10.1519/JSC.0000000000002149)] [Medline: [28820854](https://pubmed.ncbi.nlm.nih.gov/28820854/)]
10. Kovač T, Marček T, Šarkanj B, Borišev I, Ižaković M, Jukić K, et al. Fullerol C60(OH)24 nanoparticles and drought impact on wheat (*Triticum aestivum* L.) during growth and infection with *Aspergillus flavus*. *J Fungi (Basel)* 2021;7(3):236 [FREE Full text] [doi: [10.3390/jof7030236](https://doi.org/10.3390/jof7030236)] [Medline: [33809896](https://pubmed.ncbi.nlm.nih.gov/33809896/)]
11. Song Y, Lin H, Wang B, Yang M. Factors and measures affecting players' psychology in tennis. *JSSHL* 2024;7(3):106-111. [doi: [10.53469/jsshl.2024.07\(03\).14](https://doi.org/10.53469/jsshl.2024.07(03).14)]
12. Hidayat Y, Yudianta Y, Hambali B, Sultoni K, Ustun UD, Singnoy C. The effect of the combined self-talk and mental imagery program on the badminton motor skills and self-confidence of youth beginner student-athletes. *BMC Psychol* 2023;11(1):35 [FREE Full text] [doi: [10.1186/s40359-023-01073-x](https://doi.org/10.1186/s40359-023-01073-x)] [Medline: [36737818](https://pubmed.ncbi.nlm.nih.gov/36737818/)]
13. Mercader-Rubio I, Ángel NG, Silva S, Furtado G, Brito-Costa S. Intrinsic motivation: knowledge, achievement, and experimentation in sports science students-relations with emotional intelligence. *Behav Sci (Basel)* 2023;13(7):589 [FREE Full text] [doi: [10.3390/bs13070589](https://doi.org/10.3390/bs13070589)] [Medline: [37504036](https://pubmed.ncbi.nlm.nih.gov/37504036/)]
14. Bangari D, Choudhary PK, Choudhary S, Kandpal A, Singh H, Mohit. Effects of a 12-week integrated core and plyometric training program on tennis skills, agility, strength, and balance in adolescent tennis players. *Pedagogy Phys Cult Sports* 2025;29(4):308-319. [doi: [10.15561/26649837.2025.0408](https://doi.org/10.15561/26649837.2025.0408)]
15. Buyle M, Jung Y, Pavlou M, Gonzalez S, Bamiou DE. The role of motivation factors in exergame interventions for fall prevention in older adults: a systematic review and meta-analysis. *Front Neurol* 2022;13:903673 [FREE Full text] [doi: [10.3389/fneur.2022.903673](https://doi.org/10.3389/fneur.2022.903673)] [Medline: [35989930](https://pubmed.ncbi.nlm.nih.gov/35989930/)]
16. Birnstiel S, Morschheuser B. Exploring gamification in team sports: a scoping review and research agenda. 2024 Presented at: 8th International GamiFIN Conference 2024 (GamiFIN 2024); April 2-5, 2024; Ruka, Finland p. 33-46 URL: https://www.researchgate.net/publication/379873210_Exploring_gamification_in_team_sports_A_scoping_review_and_research_agenda
17. Biddiss E, Irwin J. Active video games to promote physical activity in children and youth: a systematic review. *Arch Pediatr Adolesc Med* 2010;164(7):664-672. [doi: [10.1001/archpediatrics.2010.104](https://doi.org/10.1001/archpediatrics.2010.104)] [Medline: [20603468](https://pubmed.ncbi.nlm.nih.gov/20603468/)]
18. Qiu S, Kaisar E, Ding R, Han T, Hu J, Rauterberg M. Social balance ball: designing and evaluating an exergame that promotes social interaction between older and younger players. *Int J Hum-Comput Interact* 2023;40(11):2838-2861. [doi: [10.1080/10447318.2023.2175157](https://doi.org/10.1080/10447318.2023.2175157)]

19. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen* 2012;141(1):2-18. [doi: [10.1037/a0024338](https://doi.org/10.1037/a0024338)] [Medline: [21823805](https://pubmed.ncbi.nlm.nih.gov/21823805/)]
20. Genevois C, Frican B, Creveaux T, Hautier C, Rogowski I. Effects of two training protocols on the forehand drive performance in tennis. *J Strength Cond Res* 2013;27(3):677-682. [doi: [10.1519/JSC.0b013e31825c3290](https://doi.org/10.1519/JSC.0b013e31825c3290)] [Medline: [22592176](https://pubmed.ncbi.nlm.nih.gov/22592176/)]
21. Sun H. Impact of exergames on physical activity and motivation in elementary school students: a follow-up study. *J Sport Health Sci* 2013;2(3):138-145. [doi: [10.1016/j.jshs.2013.02.003](https://doi.org/10.1016/j.jshs.2013.02.003)]
22. Sember V, Meh K, Sorić M, Starc G, Rocha P, Jurak G. Validity and reliability of international physical activity questionnaires for adults across EU Countries: systematic review and meta analysis. *Int J Environ Res Public Health* 2020;17(19):7161 [FREE Full text] [doi: [10.3390/ijerph17197161](https://doi.org/10.3390/ijerph17197161)] [Medline: [33007880](https://pubmed.ncbi.nlm.nih.gov/33007880/)]
23. Barbosa DG, Vealey R, Pelegrini A, Farias G, Corazza S, Felden É. Validation and cultural adaptation of the sport confidence inventory (SCI) for Brazilian adolescent athletes. *Rev Bras Psic Esporte* 2022;11(2). [doi: [10.31501/rbpe.v11i2.12498](https://doi.org/10.31501/rbpe.v11i2.12498)]
24. Barreira J, Amaro AS, Borges K, Luque LL, Fernandes PT. Internal consistency of the sport motivation scale-II questionnaire in the Brazilian context: potentialities and limitations. *Motriz: rev educ fis* 2022;28:e10220010721. [doi: [10.1590/s1980-657420220010721](https://doi.org/10.1590/s1980-657420220010721)]
25. Hewitt JE. Hewitt's tennis achievement test. *Res Q Am Assoc Health Phys Educ* 2013;37(2):231-240. [doi: [10.1080/10671188.1966.10613366](https://doi.org/10.1080/10671188.1966.10613366)]
26. Reuter SE, Massy-Westropp N, Evans AM. Reliability and validity of indices of hand-grip strength and endurance. *Aust Occup Ther J* 2011;58(2):82-87. [doi: [10.1111/j.1440-1630.2010.00888.x](https://doi.org/10.1111/j.1440-1630.2010.00888.x)] [Medline: [21418230](https://pubmed.ncbi.nlm.nih.gov/21418230/)]
27. Eugster N, Ewerhart C. Rules in international tennis tournaments: a game theoretic perspective. *J Sport Game* 2024;6(1):33-42. [doi: [10.22259/2642-8466.0601004](https://doi.org/10.22259/2642-8466.0601004)]
28. Pescatello L. *ACSM's Guidelines for Exercise Testing and Prescription*. Philadelphia, PA: Lippincott Williams & Wilkins; 2014.
29. Sawali L, Refiater UH. Effectiveness of training programs to improve tennis service skills. *HON* 2022;5(2):609. [doi: [10.31851/hon.v5i2.8279](https://doi.org/10.31851/hon.v5i2.8279)]
30. González-Rosalén J, Benítez-Martínez JC, Medina-Mirapeix F, Cuerda-Del Pino A, Cervelló A, Martín-San Agustín R. Intra- and inter-rater reliability of strength measurements using a pull hand-held dynamometer fixed to the examiner's body and comparison with push dynamometry. *Diagnostics (Basel)* 2021;11(7):1230 [FREE Full text] [doi: [10.3390/diagnostics11071230](https://doi.org/10.3390/diagnostics11071230)] [Medline: [34359313](https://pubmed.ncbi.nlm.nih.gov/34359313/)]
31. Mathiowetz V. Comparison of Rolyan and Jamar dynamometers for measuring grip strength. *Occup Ther Int* 2002;9(3):201-209 [FREE Full text] [doi: [10.1002/oti.165](https://doi.org/10.1002/oti.165)] [Medline: [12374997](https://pubmed.ncbi.nlm.nih.gov/12374997/)]
32. Jaafar MH, Ismail R, Ismail NH, Md Isa Z, Mohd Tamil A, Mat Nasir N, et al. Normative reference values and predicting factors of handgrip strength for dominant and non-dominant hands among healthy Malay adults in Malaysia. *BMC Musculoskelet Disord* 2023;24(1):74 [FREE Full text] [doi: [10.1186/s12891-023-06181-8](https://doi.org/10.1186/s12891-023-06181-8)] [Medline: [36709276](https://pubmed.ncbi.nlm.nih.gov/36709276/)]
33. Kamarul T, Ahmad T, Loh W. Hand grip strength in the adult Malaysian population. *J Orthop Surg (Hong Kong)* 2006;14(2):172-177 [FREE Full text] [doi: [10.1177/230949900601400213](https://doi.org/10.1177/230949900601400213)] [Medline: [16914783](https://pubmed.ncbi.nlm.nih.gov/16914783/)]
34. Pelletier LG, Tuson KM, Fortier MS, Vallerand RJ, Brière NM, Blais MR. Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: the sport motivation scale (SMS). *J Sport Exerc Psychol* 1995;17(1):53. [doi: [10.1123/jsep.17.1.35](https://doi.org/10.1123/jsep.17.1.35)]
35. Vealey RS. Conceptualization of sport-confidence and competitive orientation: preliminary investigation and instrument development. *J Sport Psychol* 1986;8(3):221-246 [FREE Full text] [doi: [10.1123/jsp.8.3.221](https://doi.org/10.1123/jsp.8.3.221)]
36. Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. *Ann Card Anaesth* 2019;22(1):67-72 [FREE Full text] [doi: [10.4103/aca.ACA_157_18](https://doi.org/10.4103/aca.ACA_157_18)] [Medline: [30648682](https://pubmed.ncbi.nlm.nih.gov/30648682/)]
37. Richardson JTE. Eta squared and partial eta squared as measures of effect size in educational research. *Educ Res Rev* 2011;6(2):135-147. [doi: [10.1016/j.edurev.2010.12.001](https://doi.org/10.1016/j.edurev.2010.12.001)]
38. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge; Dec 1988.
39. Meigh NJ, Keogh JW, Schram B, Hing W, Rathbone EN. Effects of supervised high-intensity hardstyle kettlebell training on grip strength and health-related physical fitness in insufficiently active older adults: the BELL pragmatic controlled trial. *BMC Geriatr* 2022;22(1):354 [FREE Full text] [doi: [10.1186/s12877-022-02958-z](https://doi.org/10.1186/s12877-022-02958-z)] [Medline: [35459114](https://pubmed.ncbi.nlm.nih.gov/35459114/)]
40. Garcia-Hernandez N, Garza-Martinez K, Parra-Vega V. Electromyography biofeedback exergames to enhance grip strength and motivation. *Games Health J* 2018;7(1):75-82. [doi: [10.1089/g4h.2017.0054](https://doi.org/10.1089/g4h.2017.0054)] [Medline: [29227162](https://pubmed.ncbi.nlm.nih.gov/29227162/)]
41. Levin MF, Demers M. Motor learning in neurological rehabilitation. *Disabil Rehabil* 2021;43(24):3445-3453. [doi: [10.1080/09638288.2020.1752317](https://doi.org/10.1080/09638288.2020.1752317)] [Medline: [32320305](https://pubmed.ncbi.nlm.nih.gov/32320305/)]
42. Kitago T, Krakauer JW. Chapter 8 - motor learning principles for neurorehabilitation. In: *Handbook of Clinical Neurology*. Amsterdam, NL: Elsevier; 2013:93-103.
43. Flôres F, Silva A, Willig R, Reyes A, Serpa J, Marconcin P, et al. Exergames as an effective alternative to real environmental tennis training for adolescents. *Sports (Basel)* 2024;12(6):168 [FREE Full text] [doi: [10.3390/sports12060168](https://doi.org/10.3390/sports12060168)] [Medline: [38921862](https://pubmed.ncbi.nlm.nih.gov/38921862/)]

44. Šlosar L, de Bruin ED, Fontes EB, Plevnik M, Pisot R, Simunic B, et al. Additional exergames to regular tennis training improves cognitive-motor functions of children but may temporarily affect tennis technique: a single-blind randomized controlled trial. *Front Psychol* 2021;12:611382 [FREE Full text] [doi: [10.3389/fpsyg.2021.611382](https://doi.org/10.3389/fpsyg.2021.611382)]
45. Anguera JA, Choudhry A, Seaman M, Fedele D. Assessing the impact of a virtual reality cognitive intervention on tennis performance in junior tennis players: pilot study. *JMIR Form Res* 2025;9:e66979 [FREE Full text] [doi: [10.2196/66979](https://doi.org/10.2196/66979)] [Medline: [39918854](https://pubmed.ncbi.nlm.nih.gov/39918854/)]
46. Barnett ML, Ross D, Schmidt RA, Todd B. Motor skills learning and the specificity of training principle. *Res Q Am Assoc Health Phys Educ* 2013;44(4):440-447. [doi: [10.1080/10671188.1973.10615224](https://doi.org/10.1080/10671188.1973.10615224)]
47. Gioftsidou A, Vernadakis N, Malliou P, Batzios S, Sofokleous P, Antoniou P, et al. Typical balance exercises or exergames for balance improvement? *J Back Musculoskelet Rehabil* 2013;26(3):299-305. [doi: [10.3233/BMR-130384](https://doi.org/10.3233/BMR-130384)] [Medline: [23893145](https://pubmed.ncbi.nlm.nih.gov/23893145/)]
48. Feltz DL. Self-confidence and sports performance. *Exerc Sport Sci Rev* 1988;16(1):423-458. [doi: [10.1249/00003677-198800160-00016](https://doi.org/10.1249/00003677-198800160-00016)]
49. Bandura A. *Self-Efficacy: The Exercise of Control*. New York, NY: WH Freeman/Times Books/Henry Holt & Co; 1997.
50. Heydari A, Soltani H, Mohammadi-Nezhad M. The effect of psychological skills training (goal setting, positive selftalk and Imagery) on self-confidence of adolescent volleyball players. *Pedagog psychol med-biol probl phys train sports* 2018;22(4):189. [doi: [10.15561/18189172.2018.0404](https://doi.org/10.15561/18189172.2018.0404)]
51. Stanmore EK, Mavroei A, de Jong LD, Skelton DA, Sutton CJ, Benedetto V, et al. The effectiveness and cost-effectiveness of strength and balance Exergames to reduce falls risk for people aged 55 years and older in UK assisted living facilities: a multi-centre, cluster randomised controlled trial. *BMC Med* 2019;17(1):49 [FREE Full text] [doi: [10.1186/s12916-019-1278-9](https://doi.org/10.1186/s12916-019-1278-9)] [Medline: [30813926](https://pubmed.ncbi.nlm.nih.gov/30813926/)]
52. McGonigal J. Reality is broken: why games make us better and how they can change the world. *Choice Reviews Online* 2011;49(02):49-0680. [doi: [10.5860/choice.49-0680](https://doi.org/10.5860/choice.49-0680)]
53. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am Psychol* 2000;55(1):68-78 [FREE Full text] [doi: [10.1037//0003-066x.55.1.68](https://doi.org/10.1037//0003-066x.55.1.68)]
54. Grouzet FME, Vallerand RJ, Thill EE, Provencher PJ. From environmental factors to outcomes: a test of an integrated motivational sequence. *Motiv Emot* 2004;28(4):331-346. [doi: [10.1007/s11031-004-2387-z](https://doi.org/10.1007/s11031-004-2387-z)]
55. Kerner C, Goodyear VA. The motivational impact of wearable healthy lifestyle technologies: a self-determination perspective on fitbits with adolescents. *Am J Health Educ* 2017;48(5):287-297. [doi: [10.1080/19325037.2017.1343161](https://doi.org/10.1080/19325037.2017.1343161)]
56. Nugroho D, Hidayatullah MF, Doewes M, Purnama SK. Does grip strength, performance, and hand-eye coordination affect tennis drive skills? In: *Proceedings of the International Conference on Learning and Advanced Education*. 2023 Presented at: ICOLAE 2022; December 6; Surakarta, Indonesia p. 4-11. [doi: [10.2991/978-2-38476-086-2_2](https://doi.org/10.2991/978-2-38476-086-2_2)]
57. Zeng N, Pope Z, Lee J, Gao Z. Virtual reality exercise for anxiety and depression: a preliminary review of current research in an emerging field. *J Clin Med* 2018;7(3):42 [FREE Full text] [doi: [10.3390/jcm7030042](https://doi.org/10.3390/jcm7030042)] [Medline: [29510528](https://pubmed.ncbi.nlm.nih.gov/29510528/)]
58. Lin SC, Lee JY, Yang Y, Fang CC, Fang HL, Hou TH. Exploring the design of upper limb strength training through high-intensity interval training combined with exergaming: usability study. *JMIR Serious Games* 2024;12:e51730 [FREE Full text] [doi: [10.2196/51730](https://doi.org/10.2196/51730)] [Medline: [38632713](https://pubmed.ncbi.nlm.nih.gov/38632713/)]
59. Covassin T, Pero S. The relationship between self-confidence, mood state, and anxiety among collegiate tennis players. *J Sport Behav Sep* 2004;27(3):230 [FREE Full text]
60. Kajtna T, Barić R. Psychological characteristics of coaches of successful and less successful athletes in team and individual sports. *Rev. Psychol* 2009;16(1):47-56 [FREE Full text]
61. Gomes RV, Cunha VCR, Zourdos MC, Aoki MS, Moreira A, Fernandez-Fernandez J, et al. Physiological responses of young tennis players to training drills and simulated match play. *J Strength Cond Res* 2016;30(3):851-858. [doi: [10.1519/JSC.0000000000001159](https://doi.org/10.1519/JSC.0000000000001159)] [Medline: [26382129](https://pubmed.ncbi.nlm.nih.gov/26382129/)]
62. Hahn E. *Emotions in sports*. In: *Anxiety in Sports*. Abingdon, United Kingdom: Taylor & Francis; 2021.
63. Chen SP, Shang Y, Xie LJ, Song D, Chen M. The true value of extrinsic motivation? The effect of intrinsic and extrinsic sports motivation on sports anomie behavior in college students. *Curr Psychol* 2023;43(3):2704-2713. [doi: [10.1007/s12144-023-04524-1](https://doi.org/10.1007/s12144-023-04524-1)]
64. Tuckey M, Brewer N, Williamson P. The influence of motives and goal orientation on feedback seeking. *J Occup Organ Psychol* 2010;75(2):195-216. [doi: [10.1348/09631790260098677](https://doi.org/10.1348/09631790260098677)]
65. Cervelló E, Santos Rosa FJ, Calvo TG, Jiménez R, Iglesias D. Young tennis players' competitive task involvement and performance: the role of goal orientations, contextual motivational climate, and coach-initiated motivational climate. *J Appl Sport Psychol* 2007;19(3):304-321. [doi: [10.1080/10413200701329134](https://doi.org/10.1080/10413200701329134)]
66. Thibodeaux J, Winsler A. Stay positive: self-talk in relation to motivational climate, goal orientation, and self-talk encouragement in tennis. *Res Q Exerc Sport* 2022;93(4):769-780. [doi: [10.1080/02701367.2021.1918323](https://doi.org/10.1080/02701367.2021.1918323)] [Medline: [34723765](https://pubmed.ncbi.nlm.nih.gov/34723765/)]
67. Badau D, Badau A, Ene-Voiculescu C, Larion A, Ene-Voiculescu V, Mihaila I, et al. The impact of implementing an exergame program on the level of reaction time optimization in handball, volleyball, and basketball players. *Int J Environ Res Public Health* 2022;19(9):5598 [FREE Full text] [doi: [10.3390/ijerph19095598](https://doi.org/10.3390/ijerph19095598)] [Medline: [35564993](https://pubmed.ncbi.nlm.nih.gov/35564993/)]

68. Bideau B, Kulpa R, Vignais N, Brault S, Multon F, Craig C. Using virtual reality to analyze sports performance. *IEEE Comput Graph Appl* 2010;30(2):14-21. [doi: [10.1109/MCG.2009.134](https://doi.org/10.1109/MCG.2009.134)] [Medline: [20650707](https://pubmed.ncbi.nlm.nih.gov/20650707/)]
69. Staiano AE, Calvert SL. Exergames for physical education courses: physical, social, and cognitive benefits. *Child Dev Perspect* 2011;5(2):93-98 [FREE Full text] [doi: [10.1111/j.1750-8606.2011.00162.x](https://doi.org/10.1111/j.1750-8606.2011.00162.x)] [Medline: [22563349](https://pubmed.ncbi.nlm.nih.gov/22563349/)]
70. Ochi G, Kuwamizu R, Fujimoto T, Ikarashi K, Yamashiro K, Sato D. The effects of acute virtual reality exergaming on mood and executive function: exploratory crossover trial. *JMIR Serious Games* 2022;10(3):e38200 [FREE Full text] [doi: [10.2196/38200](https://doi.org/10.2196/38200)] [Medline: [36169992](https://pubmed.ncbi.nlm.nih.gov/36169992/)]
71. Anderson-Hanley C, Arciero PJ, Brickman AM, Nimon JP, Okuma N, Westen SC, et al. Exergaming and older adult cognition: a cluster randomized clinical trial. *Am J Prev Med* 2012;42(2):109-119. [doi: [10.1016/j.amepre.2011.10.016](https://doi.org/10.1016/j.amepre.2011.10.016)] [Medline: [22261206](https://pubmed.ncbi.nlm.nih.gov/22261206/)]

Abbreviations

EBTT+OCTT: exergame-based tennis training with on-court tennis training

EM: extrinsic motivation

IM: intrinsic motivation

IPAQ: International Physical Activity Questionnaire

ITT: intention-to-treat

MD: mean difference

OCTT: on-court tennis training

SCI: Sports Confidence Inventory

SMT: Sports Motivation Test

Edited by A Coristine; submitted 17.Mar.2025; peer-reviewed by BCH Huijgen, N Deng; comments to author 23.May.2025; revised version received 15.Dec.2025; accepted 17.Dec.2025; published 02.Feb.2026.

Please cite as:

Ngo JS, Hoe SZ, Mat Rosly M

Effects of Exergaming Tennis on Players' Tennis Skills and Mental State Compared to Regular Tennis in Adult Players: Quasi-Experimental Study

JMIR Serious Games 2026;14:e73732

URL: <https://games.jmir.org/2026/1/e73732>

doi: [10.2196/73732](https://doi.org/10.2196/73732)

PMID:

©Jia Sheng Ngo, See Ziau Hoe, Maziah Mat Rosly. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 02.Feb.2026. 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.

Original Paper

Design and Iterative Development of Serious Exergames for Children With Autism Spectrum Disorder: Formative Multiple-Case Pilot Study

Won Kim^{1,2}, PhD; Minwoo Seong¹, MS; SeungJun Kim¹, PhD

¹Department of AI Convergence, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea

²Department of Computer Engineering, College of IT Convergence, Chosun University, Gwangju, Republic of Korea

Corresponding Author:

SeungJun Kim, PhD

Department of AI Convergence

Gwangju Institute of Science and Technology

Dasan Building 201, Cheomdangwagi-ro 123

Gwangju, 61005

Republic of Korea

Phone: 82 62 715 5352

Email: seungjun@gist.ac.kr

Abstract

Background: Children with autism spectrum disorder (ASD) exhibit cognitive, motor, and social difficulties that affect engagement, causing developmental delays, behavioral challenges, and obesity—interrelated concerns in daily functioning and well-being. Although interactive interventions have incorporated physical activity, they often rely on limited physical involvement and lack iterative, expert-informed design, as built on pre-existing game frameworks. Physical activity is often operationalized as constrained input (eg, gestures or in-place actions) rather than exertion-intensive, whole-body exercise, and design guidance for adapting exercise content under ASD-oriented safety and cognitive-sensory constraints remains limited. These limitations highlight the need for exergames that promote sustained, full-body participation aligned with developmental goals, motivating formative, co-design with expertise and initial field testing in this population.

Objective: We aim to iteratively design exercise-based serious games (SGs) for children with ASD through a structured, expert-informed co-design process involving 21 professionals across special education, adapted physical education, and human-computer interaction, and to examine feasibility and use contexts through an exploratory multiple-case pilot study.

Methods: We derived serious exergames using 4 design methods—stakeholder interview, concept mapping, creative matrix, and visualize the vote. Two exergames—“Fruit Sorting Run” and “Hazard Avoiding Ride”—were developed, integrating full-body running and cycling movements into goal-directed tasks under ASD-oriented constraints. We conducted a multiple-case pilot with 3 children with ASD. During gameplay, caregivers labeled engagement using a binary input interface, and we conducted postsession caregiver interviews to capture complementary observations.

Results: Engagement in both exergames tended to increase over normalized time. Generalized estimating equations with a logit link and an autoregressive working correlation of order 1 (AR1), including participant indicators, showed a statistically significant association between normalized time and engagement in Fruit Sorting Run (per 0.1 increase: $\beta=0.48$; odds ratio 1.62, 95% CI 1.09-2.38; $P=.02$) and Hazard Avoiding Ride (per 0.1 increase: $\beta=0.66$; odds ratio 1.93, 95% CI 1.04-3.60, $P=.04$). Caregiver interviews reinforced these findings, reporting increased attention, motivation, and enjoyment across both activities.

Conclusions: The findings support the applicability of an expert-informed design approach and the viability of the resulting exergames, integrating goal-directed physical activity, virtual agent-based prompting, and stakeholder-informed considerations such as motor-cognitive alignment, interactive scaffolding, and support for daily living skills. Distinct from prior SG approaches that operationalize physical activity through discrete gestures or in-place interactions, the proposed exergames embed sustained, exertion-intensive, whole-body movement within structured gameplay. Within this exploratory multiple-case pilot, engagement trajectories tended to increase over time. These preliminary observations provide an initial basis for a testable hypothesis that exertion-intensive, full-body SGs with virtual agent-based prompting may be associated with increasing engagement over time, meriting further examination in larger samples and applied educational and therapeutic contexts.

KEYWORDS

autism spectrum disorder; children; design research; exergame; physical activity; serious game

Introduction

Background

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized by cognitive, motor, and social difficulties [1], which affect daily functioning [2] and overall well-being [3]. Reduced participation highlights the importance of active engagement in supporting adaptive functioning and skill acquisition. A lack of active engagement may lead children with ASD to experience further developmental delays [4], increased risk of stereotyped and challenging behaviors [5], and a higher incidence of obesity [6]—all of which are interrelated and pose significant concerns. Given the scope of these difficulties, the continued rise in ASD diagnoses over the past decades underscores the importance of creating accessible interventions to support active participation in everyday contexts [7]. For example, serious games (SGs) have been explored as digital interventions that target functional outcomes beyond entertainment [8,9], including social communication, executive functioning, and motor coordination [10,11]. Recent SGs have broadened interaction design through modalities such as motion-tracking [12-15], interactive stimuli [16-18], extended reality [14,19], and socially assistive robots [20-23] or virtual agents (VAs) [24,25], enabling diverse forms of feedback and goal-directed activities [26-28]. Within this design space, SG interventions for children with ASD have positioned movement-based physical activities as a central design emphasis to support active participation.

Related Works

Constrained Physical Interaction in Exergames

While prior interactive systems for children with ASD have incorporated physical activity, often using motion-tracking technologies to facilitate interaction, these have primarily involved constrained forms of bodily input or static posture—particularly basic motion gestures [13,14,17,26,28-30] as well as touch [24] and button presses [18]. For example, Bhattacharya et al [26] used Kinect to design motion-based activities to engage students with ASD in a classroom setting by encouraging them to make specific movements or gestures to animate elements of a story on a screen. To improve eye-body coordination, Caro et al [28] proposed the FroggyBobby game, which used Kinect motion-tracking to control the upper limbs as the tongue of Frog and catch as many flies as possible through limb movements. Bossavit et al [27] presented a natural user interface SG that supported movement-based interaction for high-functioning children with ASD [27]. Similarly, Pena et al [29] introduced “Circus in Motion,” a multimodal exergame designed to support vestibular therapy through in-place actions involving the head, upper limbs, lower limbs, and jumping movements for people with ASD. These systems relied on in-place movements of the upper or lower limbs, using predefined gesture-based inputs that were stylized or scripted,

and structured in ways that may limit the continuity and spontaneity of physical involvement, revealing a gap in current approaches. This study adopts full-body exercises to extend beyond discrete interaction gestures—short, predefined inputs involving single-limb or in-place movements—by embedding physical activities into gameplay through whole-body movements designed to be accessible to children with ASD.

Toward Exertion-Intensive Serious Game Design

As supporting evidence, studies in behavioral intervention have demonstrated that higher-intensity physical activity, even when delivered as simple exercises like jogging or walking, can enhance engagement and reduce stereotyped behaviors among children with ASD. For example, Nicholson et al [31] reported improved on-task behavior following antecedent jogging sessions, and Neely et al [32] found that exercise until satiety produced the greatest behavioral benefits compared with no or brief exercise. These findings suggest that more rigorous and sustained physical activity may have specific benefits for attention and behavior regulation, highlighting the need to design interactive systems that embed such activity within playful and intrinsically motivating formats, such as exertion-intensive SGs, to enhance both engagement and functional outcomes for children with ASD. While some studies have incorporated task-oriented physical activities, these interventions were primarily designed for vocational purposes targeting adolescents or young adults, such as the “Mopping Game,” which used mixed reality to support vocational training [14]. Other examples, such as “PuzzleWalk,” or commercial games like “Pokémon GO [33],” combined walking with pre-existing games and primarily targeted adults [15,34], therefore often lacking iterative expert involvement and offering limited flexibility beyond pre-existing game frameworks. In addition, designing exertion-intensive SGs for children with ASD requires careful selection, adaptation, and progression of exercises to balance capability and cognitive-sensory tolerance while maintaining safety constraints. Design guidance on exercise content for SG-based interventions in this population remains limited [35], motivating a formative co-design process with special education and adapted physical education (APE) professionals.

Objective

The primary objective of this study is to iteratively design serious exergames through a structured co-design process involving professionals from multiple disciplines, including special educators, APE therapists, and human-computer interaction (HCI) researchers. Building on a series of expert interviews and collaboration, the study focuses on exergame development from the initial selection of exercises to the progressive shaping of gameplay. This process spans conceptualization, design, and development, and aims to support active participation and functional development in children with ASD. To explore feasibility and use contexts, we conducted an exploratory multiple-case pilot with 3 children with ASD,

focusing on case-based description of engagement and interaction patterns.

Methods

Ethical Considerations

This study involving autistic participants was conducted following the ethical guidelines and regulations of the Institutional Review Board at Gwangju Institute of Science and Technology (approval number HR-61-04-04). The approved protocol includes the involvement of professionals from multiple disciplines and the participation of children with ASD and their legal guardians (eg, parents).

Participant recruitment was conducted in accordance with Institutional Review Board guidelines. The study was introduced at the “Dream Tree Children Education Center” by therapists; 4 parents voluntarily enrolled their children for the experiment. Prior to participation, the experimental procedure and the role of the proxy user were explained to the parents, and informed consent was obtained. Additionally, each parent provided written assent, acknowledging their voluntary participation and the use of their data for research. They were informed that the experiment could be discontinued at any time upon request by the child, guardian, or therapist. The experiment commenced upon confirmation of the child’s willingness to participate, and participants received monetary compensation equivalent to approximately US \$70.

To maintain confidentiality, all data were deidentified immediately upon collection. Unique identification codes were

assigned to each participant, and personally identifiable information was stored separately in a secure, password-protected database, with access restricted to the primary researcher.

Interdisciplinary Co-Design

Design Considerations

This study involved collaborations with multidisciplinary experts, with participation formats—ranging from semistructured interviews to group sessions—structured around the needs of each phase. Experts were recruited from a pool of professionals who expressed interest in related local research programs (Table 1) and were selected based on their specific domain expertise to facilitate concept refinement, activity framing, and final selection (Figure 1).

To formulate exergames tailored for children with ASD, we applied the LUMA framework, which offers structured methods for synthesizing expert-derived insights into actionable design decisions [36]. Stakeholder interview—a design method for eliciting field-informed input through direct dialogue—was used to refine initial design directions. Concept mapping was then applied to organize therapist-proposed exercises by aligning cognitive and motor capabilities. Creative matrix, a method for generating ideas at the intersections of exercise and game features, was used to conceptualize exergames. Then, visualize the vote, a method for polling collaborators to reveal preferences and opinions, was used to select exergame concepts.

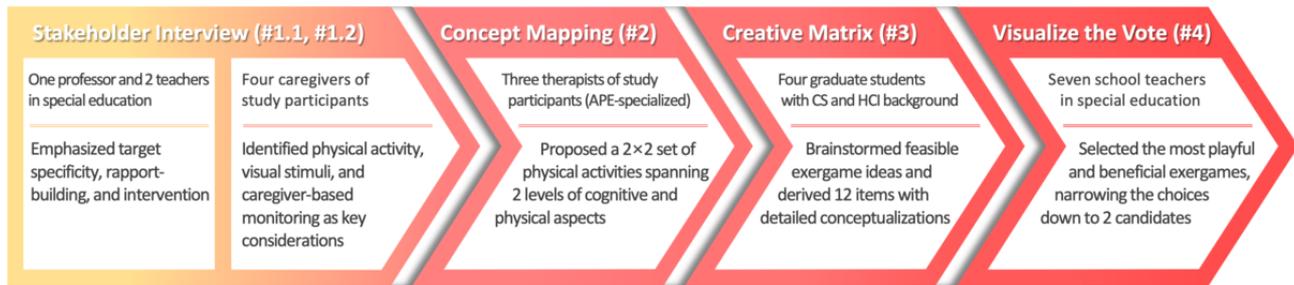
Table 1. Expert configuration across the study process, with numbers representing the composition of experts (n=21) recruited at each stage of the multistage exergame design for children with autism spectrum disorder (ASD).

Stage	Expert group	Expertise (average years of affiliation)	Collaboration type	Expert role
1.1	Professor (n=1); teachers (n=2)	Special education (>20)	Onsite semistructured interview (individual)	ASD heterogeneity, design refinement
1.2	Caregivers (n=4)	Caregiving expertise (>5)	Onsite semistructured interview (individual)	Refining intervention elements
2	Therapists (n=3)	APE ^a (>5)	Online semistructured interview (as a group)	Framing APE content
3	Researchers (n=4)	Computer science/HCI ^b (>7)	Practitioner ideation (as a group)	Exergame conceptualization
4	Teachers (n=7)	Special education (> 15)	Online semistructured interview (Individual)	Final exergame selection

^aAPE: adapted physical education.

^bHCI: human-computer interaction.

Figure 1. Iterative design process to derive exergame candidates, progressing from initial characterization of children with autism spectrum disorder (ASD) to exergame conceptualization and final expert-based selection, as indicated by the arrow direction and increasing color intensity. APE: adapted physical education; CS: computer science; HCI: human-computer interaction.



Design Method 1: Stakeholder Interview

To address the heterogeneity of ASD diagnoses across regions and to refine the initial game design approach, we conducted individual interviews (n=3) with a professor in special education and 2 special education teachers (>20 years of experience). The first expert interview indicated that the broad spectrum of disabilities necessitates specifying target groups or focusing on specific participants (eg, symptom, degree, and age), suggesting a case study approach. The interview also highlighted the importance of designing experiences that foster a sense of achievement through engagement, particularly by building rapport when introducing unfamiliar characters, game elements, or devices. Additionally, experts discussed the role of rapport and the potential use of virtual or realistic agents. To enhance interaction and motivation, they emphasized the incorporation of promptings and reinforcement feedback to support participation for children with low motivation.

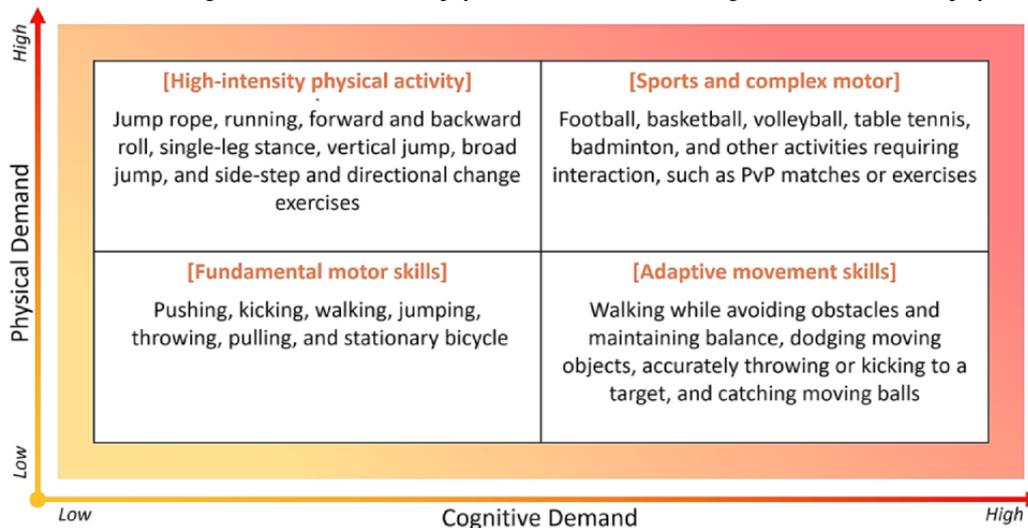
The second interview with the caregivers of our study participants (n=4) highlighted attentional challenges in children with ASD, emphasizing the need for interventions that incorporate physical exertion and visual stimuli (eg, interactive elements) to sustain engagement across varying levels of symptom severity. In particular, these elements were identified

as critical for maintaining attention and eliciting curiosity. The caregivers also noted that children with ASD often exhibit emotional fluctuations and that facial expressions and behaviors may not reliably reflect their internal emotional states (eg, a child may smile despite experiencing anxiety). Given these considerations, the interviews underscored the importance of monitoring engagement through observers with an in-depth understanding of the child’s behavioral cues, such as caregivers, to more accurately interpret emotional and attentional states.

Design Method 2: Concept Mapping

We collaborated with therapists (n=3) who have expertise and experience (>5 years) in APE to derive readily applicable physical activities tailored to children with ASD participating in this study. These experts were APE specialists practicing at local centers, selected to reflect the practical constraints and requirements of the target environment. The therapists proposed a set of exercises that the participants were likely to recognize and engage with. To incorporate both cognitive and physical aspects, they organized these activities within a 2x2 framework, which served as a foundation for identifying and refining exergame ideas (Figure 2). This framework guided exercise selection for children with ASD, which subsequently structured a brainstorming session to develop exergame concepts for implementation.

Figure 2. Therapist-proposed exercises serving as the basis for exergame design, organized in a 2x2 matrix according to cognitive and physical demands (n=3). Cognitive demand increases along the horizontal axis, and physical demand increases along the vertical axis. PvP: player vs player.



Design Method 3: Creative Matrix

Based on suggestions of physical activities spanning low to high physical and cognitive loads, graduate students (n=4) with an HCI background brainstormed feasible exergames that integrate cognitive and physical demands. The criteria included

embedding physical activity scenarios, ensuring accessibility for children aged approximately 10 years, maintaining feasibility for children with ASD, and supporting daily living skill improvement. We derived 12 exergames along with their detailed conceptualizations and prioritized them through a voting process involving special education teachers (Figure 3).

Figure 3. Brainstormed exergame concepts generated through practitioner ideation based on adapted physical education (APE) and human-computer interaction (HCI) expertise, with concepts 1 and 4 selected as final candidates through a voting process involving special education teachers (n=7).



Design Method 4: Visualize the Vote

The final selection process involved special education teachers (n=7) and focused on the following two aspects (1) playfulness, ensuring that children with ASD can actively participate, and (2) the potential developmental benefits of the game, such as supporting daily living skills. Each teacher was asked to select up to 3 exergames they considered the most playful SGs. Two exergames—“Fruit Sorting Run” (1) and “Hazard Avoiding Ride” (4)—received the highest number of votes (5 and 4, respectively) and were the only ones selected by all participants.

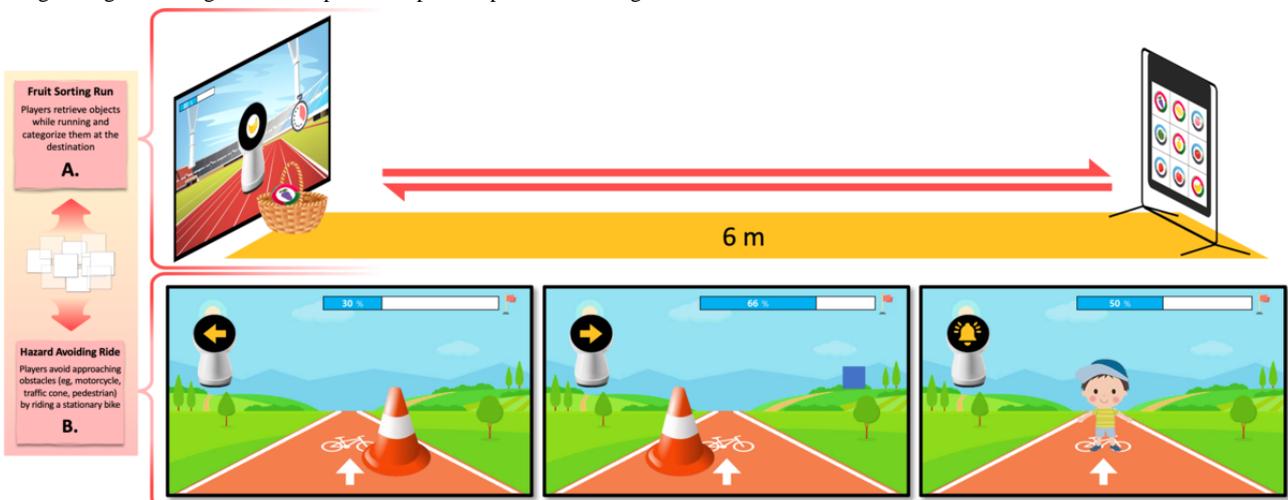
Run and Hazard Avoiding Ride (Figure 4). Rooted in running and bicycling activities, these games integrate full-body dynamics with daily living skills training in accordance with SG design principles. The resulting system supports exertion-intensive gameplay under ASD-oriented constraints through VA-mediated scaffolding and audiovisual feedback for immediate reinforcement [37] within an animated game environment [24,35]. To prioritize predictability and minimize cognitive load, difficulty levels are kept constant; this approach enables children to focus on functional task performance—fruit categorization and hazard avoidance—while mitigating potential anxiety associated with increasing difficulty [38]. The system was implemented using the Unity 2D engine and refined through iterative adjustments following an initial evaluation session involving a therapist and a child with ASD.

Design Principles and Exergame Development

Overview of the 2 Serious Exergames

Through an iterative design process involving multidisciplinary expertise, we derived and finalized 2 exergames: Fruit Sorting

Figure 4. Illustrations of the developed exergames based on running and bicycling, depicting (A) the Fruit Sorting Run game and (B) the Hazard Avoiding Ride game, along with their spatial setup and representative in-game interfaces.

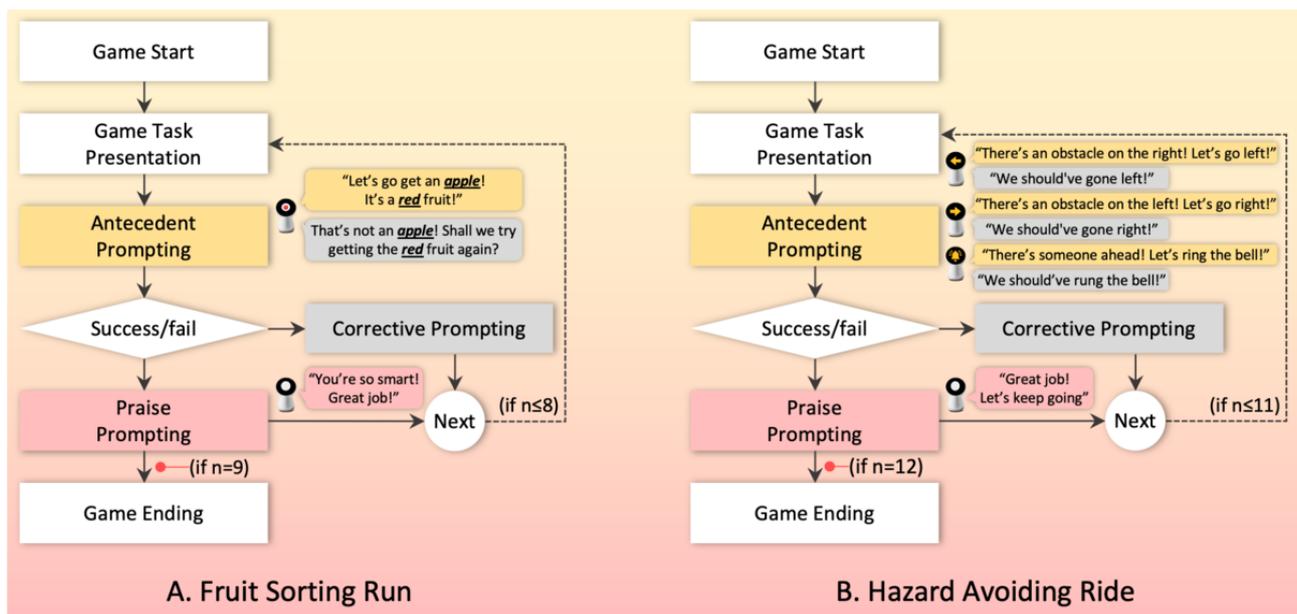


Exergame 1: Fruit Sorting Run

This game aims to enhance visual-perceptual abilities by having children classify velcro balls, each labeled with a fruit image, while walking back and forth across a 6 m straight area (Figure 4A). The game starts in a static game scene where the VA appears at the center of the screen in the running track background. The child observes a fruit image presented on the VA's screen, runs to the back wall to retrieve a ball with the corresponding image, and places it in a basket positioned below

the screen. A total of 9 velcro balls are used, each depicting a different fruit (eg, peach, grape, apple, strawberry, tomato, banana, watermelon, oriental melon, and lemon), and all are randomly placed on the sticky wall. The child is tasked with collecting 9 randomly presented fruit images within a given time limit. The game ends when all balls have been transferred and the progress bar is completely filled. The exergames developed in this study were structured according to the following game logic (Figure 5A).

Figure 5. Game logic and structured prompting flow for the 2 developed exergames, illustrating how tasks progress across stages and branch according to success or failure, while different prompting strategies are applied conditionally throughout gameplay; antecedent prompts appear earlier in the flow, corrective prompts are introduced following unsuccessful attempts, and praise prompts accompany successful task completion, each distinguished by a consistent color scheme.



Exergame 2: Hazard Avoiding Ride

This game aims to enhance visual-perceptual abilities by having children ride a stationary bicycle while avoiding obstacles that appear on the left or right side of the VA's screen through body tilting. Three types of events are randomly presented: (1) a traffic cone on the left, (2) a traffic cone on the right, and (3) a pedestrian ahead (Figure 4B). Participants are required to tilt their bodies in the opposite direction of the obstacle (eg, to the right when a traffic cone appears on the left, and to the left when it appears on the right). When a pedestrian appears in front of the player, the participant must press a bell attached to the bicycle to signal a warning. The game consists of 12 sessions and ends when the progress bar is fully filled according to game logic (Figure 5B).

Virtual Agent Integration

To facilitate sustained engagement and provide prompting support within exergames, a VA was integrated into the system. The VA was modeled after Jibo [23], a human-like robot with a head and torso [37,39] that was used to build rapport with children with ASD before gameplay [24,25]. During gameplay, the VA appeared in the game scene and delivered prompts aligned with 3 established categories [40]. These prompts included not only verbal instructions but also visual cues displayed on the VA's screen to guide task execution.

Specifically, antecedent prompting involved directive cues presented before task initiation to guide the child's actions (eg, "Let's go get an apple"). Corrective prompting was provided after an attempted response to help the child adjust or refine their behavior (eg, "That's not an apple! Shall we try getting the red fruit again?"). Praise prompting consisted of positive reinforcement to encourage continued engagement throughout the tasks (eg, "Great job!" and "Yes, that's correct"). In each gaming context, prompting strategies were tailored based on the task flow and game-specific mechanics of each exergame (Figure 5).

Adjustments Following Initial Evaluation

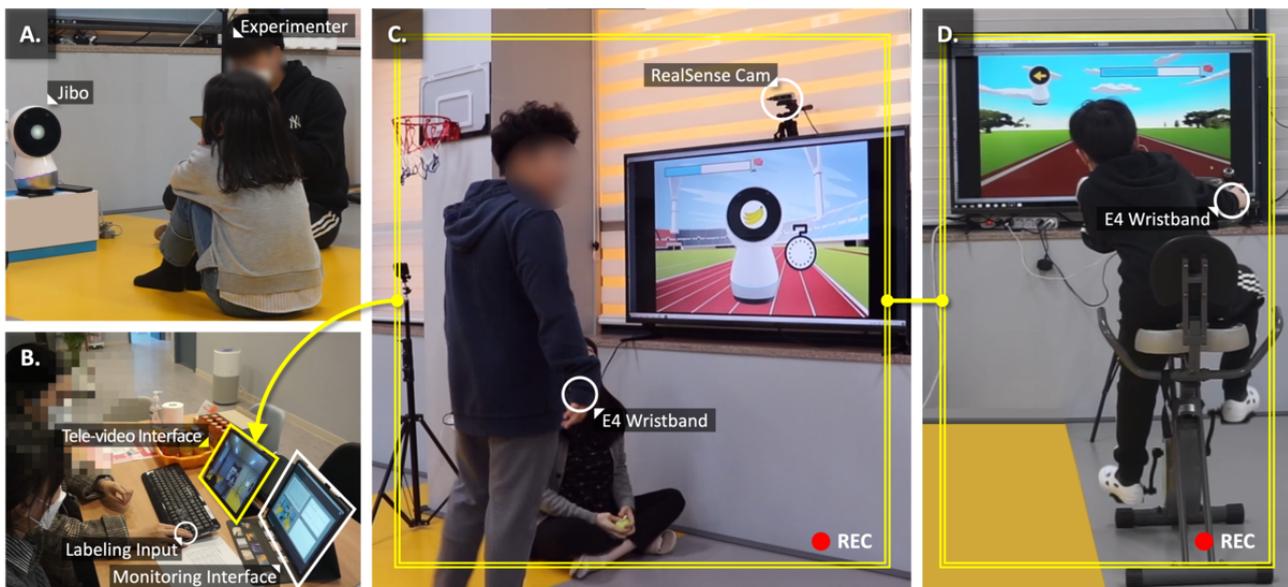
A preliminary study with a child with ASD (a boy, aged 11 years) and one therapist highlighted the need for explicit background audio, auditory prompting, and causal feedback; the volume of prompting and game-related sound effects was increased. Feedback for the final stage was added to support a sense of completion (eg, flag or applause). The study also identified the importance of a progress bar, which enabled the child to track both task progression and their own pace [38]. The progress bar was filled as each session ended. For the Hazard Avoiding Ride, the original 24-session version was reduced to 12 sessions to keep the total duration within 3 minutes, which applies to both exergames.

Exergame Setup

Exergames were implemented at a local special education center that supports mental and physical health, as well as basic motor function development, for children with ASD. A 55-inch screen was placed in front of the players to display the exergames (Figures 6C and 6D). A wearable E4 wristband and 2 cameras were used to minimize external stimuli and accommodate hypersensitivity in children with ASD [41]. The E4 wristband was attached to the participant's wrist and connected to the SG to monitor behavioral and physiological data streams. Two

cameras were positioned at the front and back to capture the player's behavior and record video. Building on this setup, a real-time visualization platform was developed and deployed to facilitate engagement labeling by parents during gameplay. The platform synchronized sensor data with live video and was displayed on two 12.9-inch iPad Pro tablets—one displaying a live video feed of the child, and the other presenting time-series plots visualizing variations in physiological and behavioral signals, including engagement status, accelerometer data, galvanic skin response, heart rate, skin temperature, and motion tracking.

Figure 6. The implemented exergame system at Dream Tree Children Education Center, depicting each study phase: (A) rapport building, (B) engagement labeling and monitoring, (C) Fruit Sorting Run, and (D) Hazard Avoiding Ride. Gameplay video streams from (C) and (D) were transmitted in real time to the engagement labeling interface shown in (B).



Procedure

To examine the practical applicability of the iteratively developed exergames, we conducted a human-participant study involving children with ASD. In line with case-oriented approaches used in ASD-related studies—where high interindividual variability necessitates detailed observation at the individual level—3 children with ASD (P1-P3), aged between 10 years and 13 years (mean 11, SD 1.3 years), were finally recruited and evaluated [42]. P4, who was originally

scheduled to participate, withdrew consent. The refusal stemmed from P4's sense of loss and emotional displacement after observing P3 (a peer identified as a rival) establishing a close and successful play interaction with the Jibo robot in a preceding session, effectively feeling that the friend (robot) had been taken. The demographic details of the participants are summarized in Table 2. All planned engagement labeling data and playtimes were collected without missing entries for the included participants.

Table 2. Pilot study demographics and clinical characteristics of children with autism spectrum disorder (ASD); 4 participants (P1-P4) were initially enrolled, with P4 withdrawing prior to study completion and excluded from subsequent analyses.

Participant (age in years)	Gender	Disability	K-Disability Grade ^a (DSM-5 ^b)	Rapport level ^c	Elementary school grade ^d	Communication capacity
P1 (10)	Female	ASD	G-2 (\approx L-2)	L-2	G-2	Inexpressible
P2 (13)	Male	ASD	G-2 (\approx L-2)	L-1	G-5	Inexpressible
P3 (11)	Male	ASD	G-2 (\approx L-2)	L-5	G-3	Expressible
P4 (11)	Male	ASD	G-2 (\approx L-2)	L-5	G-3	Withdrew

^aThe K-Disability Grades constituted the national clinical classification system for ASD (Republic of Korea). While structurally distinct from the DSM-5, shared criteria like the intelligence quotient (IQ) threshold ($IQ < 70$) allow for cross-reference; specifically, Grade 2 aligns with Level 2 based on this threshold. Corresponding DSM-5 levels are included for cross-reference purposes.

^bDSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.

^cRapport level was evaluated by a therapist using a single item on a 5-point Likert scale (“strongly disagree” to “strongly agree”): “Overall, the interaction felt highly coordinated and established a positive, enjoyable connection.” This measure captures the 2 critical rapport dimensions emphasized in the Gratch et al [43] framework: coordination and connection.

^dKorean elementary school education spans grades 1 to 6.

Before gameplay, participants were introduced to the exergames and informed of their respective roles (Figure 7). A preliminary session was conducted to facilitate rapport building (Figure 6A). During this session, participants interacted with Jibo, a physically embodied robot operated by the experimenter using a Wizard-of-Oz protocol. Children engaged in interactive play with the robot, which included head tilting to navigate a spaceship and responding to movement prompts (eg, dancing or playing guitar), involving audiovisual and physical interaction. The main session began after the therapist confirmed that sufficient engagement and rapport had been established. A practice session followed to assess the child’s readiness to engage with the exergames. Children who were able to follow the game flow—either independently or as determined by the therapist—proceeded to the main game session. The 2 exergames were presented in a counterbalanced order to ensure equitable exposure. The child participated in both games while

each child’s parent, located in a separate room, labeled engagement status using a binary scheme (not engaged=0, engaged=1) based on real-time observation (Figure 6B) [44]. Parents were instructed to press the “Enter” key when the child appeared engaged and to refrain from pressing otherwise, following predefined operational definitions and behavioral indicators (Table 3). Before the main session, the experimenter explained the criteria and conducted a calibration using practice examples. We adopted binary labels to reduce boundary ambiguity that can increase subjectivity in finer-grained schemes (eg, ternary or multilevel labels). If the labeler judged that the annotation was incorrect over a given time-series interval, the label assignments for that interval were revised. After the gameplay session, a parent interview was conducted to gather feedback based on proxy observations of the child’s experience [45,46].

Figure 7. Experimental procedure across child and parent segments, depicting the full study timeline from start to end. The parent (top) and child segments (bottom) were conducted concurrently within the same session; rapport building began at experiment onset, engagement labeling was applied only during main sessions, and postgameplay interviews followed experiment completion. Back-view photos were retained as they do not permit identification; blur was applied where needed to prevent participant identification. Frames are arranged left to right to show the temporal progression; each game session (practice + main) followed a counterbalanced order across participants.

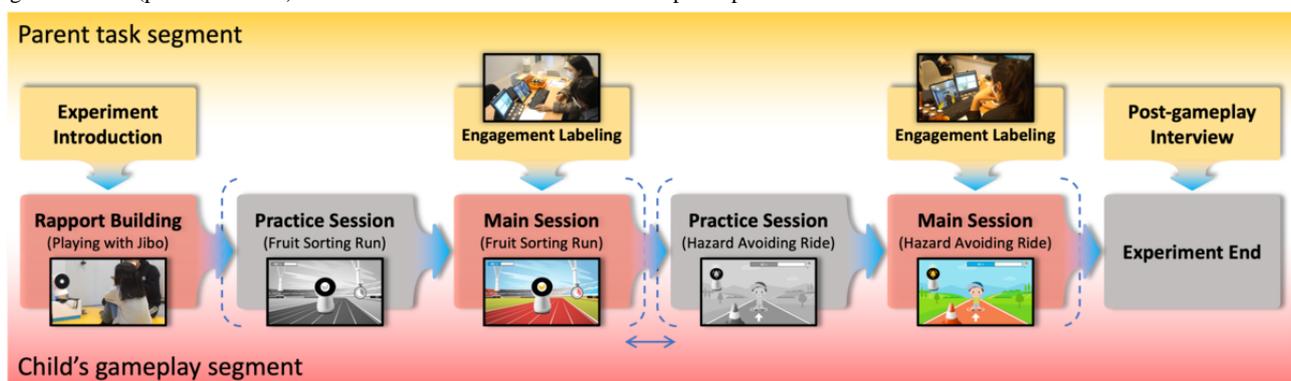


Table 3. Behavioral criteria for caregiver-based engagement labeling during exergame play, including binary engagement labels, their definitions, and corresponding observable behavioral indicators used to guide labeling.

Engagement label	Definition	Behavioral Indicators
Not engaged=0	The child shows no cognitive or motor participation relevant to the exergame.	(1) Looking away from the screen or robot, (2) leaving the play area, (3) stopping movement, or (4) engaging in unrelated behaviors.
Engaged=1	The child demonstrates active participation in the gameplay task.	(1) Orienting toward the screen, (2) following prompts, (3) moving body parts according to task goals, and (4) responding to game feedback

Results

Overview

We analyzed the temporal trends in engagement status of children with ASD, based on labels provided by their caregivers during gameplay. As each child engaged in the game for a different time duration, we first summarized the total playtime for each child (Table 4). While P1 recorded a session of 3 minutes 13 seconds in the Hazard Avoiding Ride, the average playtime across games generally aligned with expert guidance for children with ASD. To support time-aligned comparisons of engagement across participants, the time axis for each child

was normalized by dividing the elapsed time by the total game duration, resulting in a scale from 0 to 1. We then modeled changes in engagement over normalized time using generalized estimating equations (GEE) with a logistic link function to account for repeated measures within participants. Specifically, we fit a GEE logistic model with an autoregressive working correlation of order 1 (AR1), treating each participant as a clustering unit and including participant indicators to control participant-level baseline differences. Engagement status (not engaged=0, engaged=1) served as the outcome variable. Given the small number of participants, we interpret inferential statistics as exploratory and emphasize effect sizes with CIs.

Table 4. Play durations per participant (P1-P3) for each exergame in minutes and seconds, showing participant-level variation in playtime.

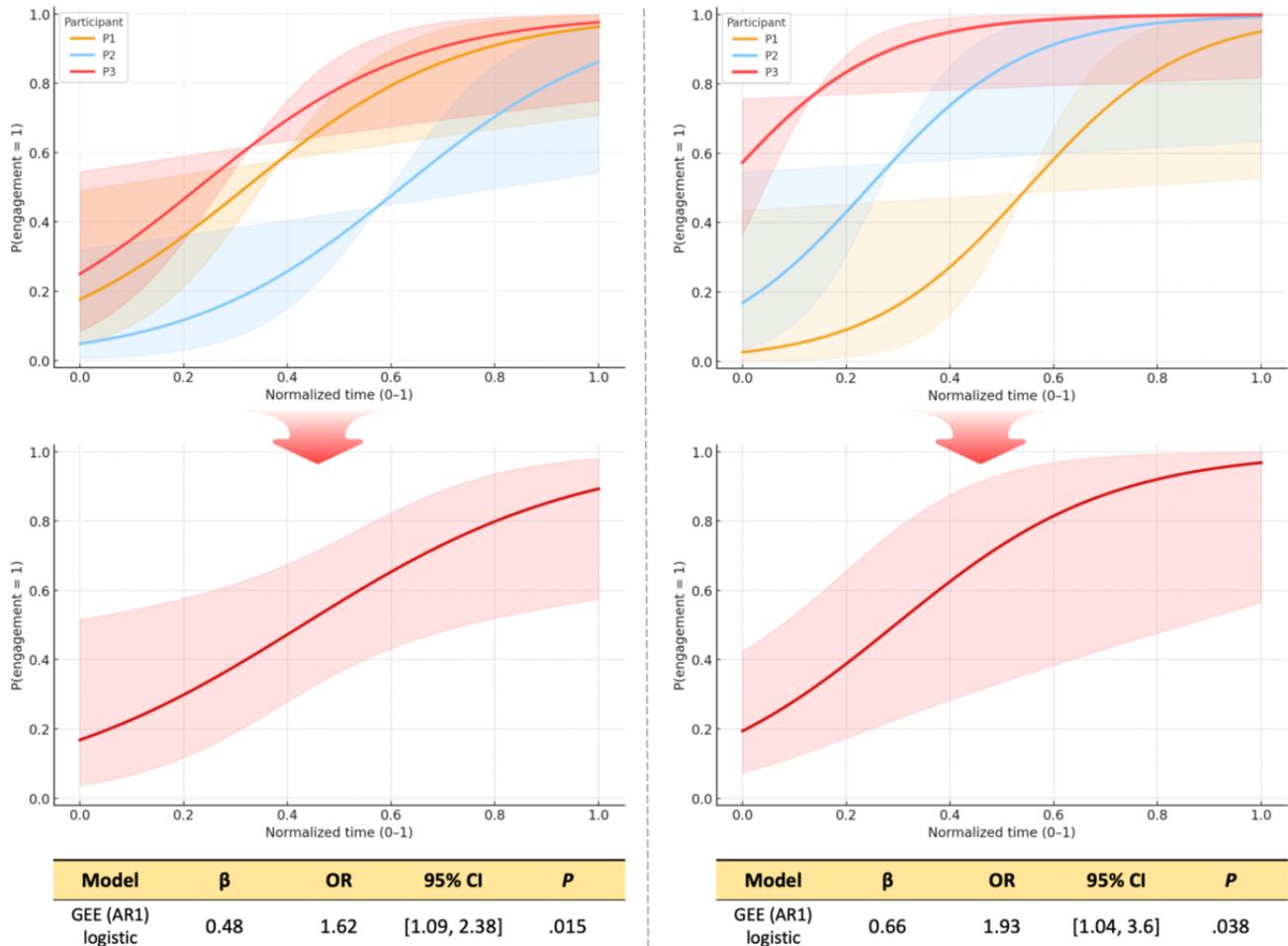
Exergame type	P1	P2	P3	Mean (SD)
Fruit Sorting Run	2 min 25 s	2 min 59 s	2 min 1 s	2 min 28 s (29 s)
Hazard Avoiding Ride	3 min 13 s	2 min 25 s	1 min 55 s	2 min 31 s (39 s)

Engagement Trends Over Time

Engagement in both the Fruit Sorting Run and Hazard Avoiding Ride tended to increase over normalized time (Figure 8). Using GEE with a logistic link to account for repeated measures within participants (AR1 working correlation; participant indicators included), normalized time was significantly associated with a higher probability of engagement status =1 in the Fruit Sorting Run (per 0.1 increase: $\beta=0.48$; odds ratio 1.62, 95% CI

1.09-2.38; $P=.02$). The Hazard Avoiding Ride showed the same direction of association (per 0.1 increase: $\beta=0.66$; odds ratio 1.93, 95% CI 1.04-3.60; $P=.04$). The results present participant-specific predicted probabilities and a population-averaged trend estimated from GEE logistic models with an AR1 working correlation (95% CI). Consistent with these estimates, model-estimated P(engagement=1) tended to be higher at later normalized timepoints in both activities.

Figure 8. Engagement trends over normalized time (generalized estimating equations [GEE] logistic, autoregressive working correlation of order 1 [AR1]). Shaded bands indicate 95% CIs. Top: individual-level predicted probabilities shown to illustrate engagement trends. Bottom: population-averaged predicted probabilities across children with autism spectrum disorder (ASD), forming the basis for the model estimates reported in the accompanying table, with (A) Fruit Sorting Run and (B) Hazard Avoiding Ride. OR: odds ratio.



A. Fruit Sorting Run

B. Hazard Avoiding Ride

Postinterview With Caregiver

Along with the temporal trends observed during gameplay, caregiver interviews with all 3 participating children with ASD (P1-P3) supported the observed engagement patterns. The interviews revealed distinct yet converging increases in engagement, along with perceptions of each game’s strengths and suggestions for improvement. Feedback was analyzed by game type to understand how each activity influenced engagement. In the Fruit Sorting Run, caregivers reported that participants remained attentive and engaged throughout the session. P1 appeared focused and responded positively to praise, suggesting a sense of achievement. The caregiver noted that adaptive verbal cues could further support engagement when the child was not running. P1 also showed difficulty differentiating fruits with similar colors and shapes (eg, apple vs tomato; oriental melon vs lemon), which sometimes interrupted task progression. P2 expressed a stronger preference for the Fruit Sorting Run over the Hazard Avoiding Ride. Although he showed initial hesitation, he maintained effort throughout the activity, even though he typically disengages when bored. The caregiver noted that the task aligned with his

everyday preference for sorting beads by color and type, which may have supported sustained participation. Because communication was challenging, P2 relied primarily on visual prompts. P3, who usually avoids running, remained physically active without breaks and showed enjoyment. His caregiver observed that he took pride in the fruit-sorting task, which may have contributed to sustained engagement. In the Hazard Avoiding Ride, similar patterns of engagement were reported. All 3 children showed attention and engagement patterns that were consistent with an increasing trend over time, despite differences in prior experience. According to P1’s caregiver, the child focused on the task and followed the VA’s instructions, with interest increasing as the session progressed. P2, unfamiliar with bicycling, found the activity novel and remained highly concentrated; the caregiver noted that repeated exposure could promote learning. P3 was initially anxious but became increasingly engaged. Unlike in typical physical or leisure activities—where he often loses interest—he demonstrated more active and enjoyable participation once familiar with the game.

Discussion

Principal Results

The primary goal of this study was to design full-body serious exergames for children with ASD through an iterative, expert-informed process that emphasized both playfulness and potential functional benefits. As a result, 2 exergames were developed—Fruit Sorting Run and Hazard Avoiding Ride—each designed to promote engagement through physical activity aligned with functional skill development. To achieve this, rather than relying on a single design iteration or a pre-existing game framework, we adopted a multiphase participatory approach that integrated input from special education experts, caregivers, therapists, and HCI researchers across 4 structured phases (Figure 1). This process yielded not only 2 target exergames but also a systematic methodology that highlights the role of stakeholder involvement in tailoring game content to the cognitive, motor, and daily needs of children with ASD.

Three design considerations emerged from this process: (1) grounding gameplay in goal-oriented exercises aligned with the cognitive and motor capabilities of the target population. The exergame conceptualization was grounded in a 2×2 activity matrix informed by therapist categorization of physical and cognitive demands—a structured method that reflects prior evidence showing that tasks calibrated to users' cognitive and motor levels support skill development and adaptive functioning in children with ASD [47]. This approach reflects SG design principles that promote aligning game tasks with users' motor and cognitive abilities through intended in-game goals [38]. Recent studies have further shown that exercises aligned with the cognitive and motor abilities of children with ASD not only improve gross motor skills but also help motivate participation among children with ASD [48,49]. (2) Engagement scaffolding mechanisms—such as prompts and feedback from VAs—were incorporated to support sustained interaction. Existing VA-based systems for children with ASD have demonstrated the utility of contingent feedback in guiding user responses. These systems primarily focused on low-level interactions, responding to gaze [24], nonverbal conversational cues [50], or emotional expressions [51]. While such approaches support reciprocal interaction, they typically emphasize momentary cue–response dynamics rather than task-oriented guidance. Our VA expands on these foundations by delivering gameplay tasks directly, combining verbal instructions and visual cues to deliver gameplay tasks. (3) Prioritizing gameplay elements perceived by stakeholders—especially educators and caregivers—as playful, goal-oriented, and compatible with daily routines, the design approach reflects established principles in SG design for children with ASD [38]. Routine exercises that align with daily activities [52], such as jogging and cycling, can promote active engagement of children with ASD [53] and lead to improvements in skills in planning, inhibition, and cognitive flexibility, along with reductions in repetitive behaviors [54].

The pilot field study suggested that engagement trajectories tended to align with an increasing trend over time during gameplay, despite differences among participants in age, school grade, and rapport levels with the VA and occasional distraction

or disengagement reflecting common attentional variability in children with ASD; behavioral labeling and caregiver interviews supported this overall pattern. Caregiver interviews indicated that even children with limited prior experience or low baseline motivation (eg, unfamiliar with bicycling [P2], typically avoidant of running, or prone to early disengagement [P3]) became more involved as they grew accustomed to the game structure. Caregivers noted that children responded positively to praise prompting (P1) and showed behavioral responses to antecedent prompting, following the VA's instructions as interest and task focus increased during gameplay—a pattern similarly emphasized by Bernardini et al [24], who reported that children with ASD showed increased responsiveness to directive prompts delivered by a VA. Caregivers also noted that children with limited expressive communication (P1 and P2) relied more on intuitive, visual prompting; therefore, follow-up work should present options through explicit and intuitive visual media and avoid ambiguous choice boundaries. This design choice aligns with established intervention practices that commonly use visual supports, structured prompting, and reinforcement to sustain task participation [24]. Observations during gameplay further suggested that adaptive prompting, triggered in moments of hesitation or disengagement, could support sustained interaction and re-engagement [55]. Taken together, these findings highlight the VA's role in supporting engagement through prompting during gameplay.

Limitations and Future Directions

This study aimed to develop full-body serious exergames for children with ASD through an iterative, expert-informed design process. As an initial exploration, we tested the games with 3 children to examine engagement patterns during gameplay. Although all participants showed an upward trend in engagement, the small sample size can limit the interpretation of these patterns across the wider population of individuals with ASD. Given the heterogeneity in cognitive, motor, and behavioral characteristics, further studies with larger and more varied samples are required to examine whether similar engagement trends appear across individuals. Examining which child characteristics (eg, communication level, sensory sensitivity, motor proficiency, and baseline interest) are associated with engagement trajectories and responsiveness to prompts may be valuable for informing individualized interaction profiles. This direction may support tailoring prompt timing, modality (visual/verbal), and reinforcement style to the needs of each child rather than applying a single prompting approach. Expanding the dataset would also support future work on adaptive prompting strategies [55].

Physiological signals (eg, heart rate and skin conductance) were streamed during gameplay to assist annotators' contextual interpretation; however, these data were not incorporated into the present analysis. As the study focused on behavioral engagement labeled by caregivers, physiological measures were treated as supplementary rather than analytic components. Future work may integrate these multimodal data to model engagement dynamics more comprehensively. In particular, combining physiological signals with gameplay context (eg, pauses, slowed pace, and errors) may help detect hesitation or overload and trigger context-sensitive prompts.

The engagement labels used a binary scheme, which reduces engagement to an on/off state and may not capture brief fluctuations or multidimensional aspects of engagement. Future work may consider multilevel or continuous labeling, along with complementary measures, to represent engagement dynamics with higher temporal resolution. Such representations can also enable evaluating how different adaptive prompting strategies affect distinct engagement components (eg, attention vs task persistence) across children.

While prompting was delivered through a VA, it was not dynamically tailored to individual behaviors. To enable context-sensitive and engagement-responsive interaction, future studies will aim to recognize engagement status in real time and apply machine learning to model and deliver adaptive prompting, aligned with O'Guinn et al [56]. Building on this direction, future work will extend the current VA-based prompting scheme from a fixed, prespecified schedule to an engagement-responsive strategy. A feasible next step is to evaluate fixed prompting versus adaptive prompting using a single-case experimental design, such as a multiple-baseline design across participants. Baseline sessions would follow the current fixed prompting logic, whereas intervention phases would introduce adaptive prompting that adjusts timing and modality based on predefined behavioral triggers observable in the existing system (eg, sustained disengagement, repeated errors, or prolonged response latency). Engagement trajectories, recovery from disengagement, and prompt responsiveness would be compared within individuals across phases, enabling rigorous individual-level evaluation under substantial interindividual variability in children with ASD.

Finally, although increased engagement was observed, the study did not assess therapeutic or educational outcomes. Future research should investigate whether these exergames can support specific intervention goals, such as improving emotional regulation, reducing stereotyped behaviors, or enhancing functional skills [57]. If such outcomes are demonstrated, the system may also offer practical value to special educators seeking to promote active participation in school-based or clinical settings. To situate the system within the broader intervention landscape, future studies will test deployment across common contexts (school-based sessions, clinical programs, and home practice) and assess feasibility factors relevant to

practitioners (eg, setup time, supervision demands, and alignment with individualized goals).

Conclusions

This study presented the iterative design and initial field evaluation of 2 full-body serious exergames for children with ASD, developed through a multiphase, expert-informed process involving 21 professionals from special education, APE, and HCI. Grounded in goal-directed physical activity and supported by engagement scaffolding mechanisms such as VA-based prompting, the exergames were tailored to promote interaction aligned with developmental needs. The design process was shaped by 3 core considerations: aligning gameplay with children's cognitive and motor capabilities, incorporating scaffolding mechanisms to support engagement, and ensuring playfulness and perceived developmental value, such as supporting daily living skills.

Beyond prior SG approaches that primarily relied on discrete gestures or in-place interactions, this study advances exertion-intensive, whole-body exergame design by embedding sustained physical activity within structured gameplay, shaped through a series of expert collaborations. As an exploratory multiple-case pilot, we tested the 2 developed exergames with 3 children with ASD to examine feasibility and engagement patterns during gameplay, observing increasing engagement trends over time despite inter-individual differences.

By integrating expert-driven design with case-based field evaluation conducted in a special education setting, this work contributes a grounded design and evaluation framework for developing exertion-intensive exergames for children with ASD. This contribution extends prior work on serious exergames for children with ASD by demonstrating how therapist-informed exercises and structured prompting strategies can be systematically translated into playable, full-body game mechanics. While these observations remain preliminary, they provide initial indications that should be examined and strengthened with larger samples. Taken together, this exploratory multiple-case pilot provides an initial basis for a testable hypothesis that full-body serious exergames, paired with VA-based prompting, may be associated with increasing engagement over time, which should be examined in larger samples.

Acknowledgments

We express our sincere appreciation to Director Choongman Lee of the Dream Tree Children Education Center for his role in supporting this work. We also extend our thanks to the children with ASD, their parents, and other caregivers for their participation, as well as to the 21 special education teachers for contributing their time and perspectives. The authors declare the use of generative artificial intelligence (AI) in the research and writing process. According to the GAIDeT taxonomy (2025), generative AI use was limited to proofreading and language editing under full human supervision. The tool used was ChatGPT-5.2 (OpenAI). Responsibility for the final manuscript lies entirely with the authors, and generative AI tools are not listed as authors or responsible for the final outcomes. This declaration is submitted under collective responsibility.

Funding

This work was supported in equal parts by the National Research Foundation of Korea (NRF) grants funded by the Korea government (MSIT) (RS-2024-00343397; RS-2025-16902996).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

WK led the system development, conducted the experiments, and performed data analysis. MS contributed to the experimental sessions and assisted with data organization. SJK supervised the overall study design and provided critical revisions. All authors contributed to manuscript writing and approved the final version.

Conflicts of Interest

None declared.

References

1. Mazurek MO. Social media use among adults with autism spectrum disorders. *Comput Hum Behav* 2013 Jul;29(4):1709-1714. [doi: [10.1016/j.chb.2013.02.004](https://doi.org/10.1016/j.chb.2013.02.004)]
2. Stanish H, Curtin C, Must A, Phillips S, Maslin M, Bandini L. Enjoyment, barriers, and beliefs about physical activity in adolescents with and without autism spectrum disorder. *Adapt Phys Activ Q* 2015;32(4):302-317 [FREE Full text] [doi: [10.1123/APAQ.2015-0038](https://doi.org/10.1123/APAQ.2015-0038)] [Medline: [26485735](https://pubmed.ncbi.nlm.nih.gov/26485735/)]
3. Boerner KE, Pearl-Dowler L, Holsti L, Wharton M, Siden H, Oberlander TF. Family perspectives on in-home multimodal longitudinal data collection for children who function across the developmental spectrum. *J Dev Behav Pediatr* 2023;44(4):e284-e291 [FREE Full text] [doi: [10.1097/DBP.0000000000001183](https://doi.org/10.1097/DBP.0000000000001183)] [Medline: [37074803](https://pubmed.ncbi.nlm.nih.gov/37074803/)]
4. Phillips JM, Uljarević M, Schuck RK, Schapp S, Solomon EM, Salzman E. Development of the stanford social dimensions scale: initial validation in autism spectrum disorder and in neurotypicals. *Mol Autism* 2019;10:48 [FREE Full text] [doi: [10.1186/s13229-019-0298-9](https://doi.org/10.1186/s13229-019-0298-9)] [Medline: [31890146](https://pubmed.ncbi.nlm.nih.gov/31890146/)]
5. Plötz T, Hammerla N, Rozga A, Reavis A, Call N, Abowd GD. Automatic assessment of problem behavior in individuals with developmental disabilities. 2012 Presented at: Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12); September 5-7, 2012; Pittsburgh p. 391-400. [doi: [10.1145/2370216.2370276](https://doi.org/10.1145/2370216.2370276)]
6. McCoy SM, Jakicic JM, Gibbs BB. Comparison of obesity, physical activity, and sedentary behaviors between adolescents with autism spectrum disorders and without. *J Autism Dev Disord* 2016;46(7):2317-2326. [doi: [10.1007/s10803-016-2762-0](https://doi.org/10.1007/s10803-016-2762-0)] [Medline: [26936162](https://pubmed.ncbi.nlm.nih.gov/26936162/)]
7. Choi H, Kim JH, Yang HS, Kim JY, Cortese S, Smith L. Pharmacological and non-pharmacological interventions for irritability in autism spectrum disorder: a systematic review and meta-analysis with the GRADE assessment. *Mol Autism* 2024;15(1):7 [FREE Full text] [doi: [10.1186/s13229-024-00585-6](https://doi.org/10.1186/s13229-024-00585-6)] [Medline: [38263251](https://pubmed.ncbi.nlm.nih.gov/38263251/)]
8. Mazurek MO, Engelhardt CR, Clark KE. Video games from the perspective of adults with autism spectrum disorder. *Comput Hum Behav* 2015;51:122-130. [doi: [10.1016/j.chb.2015.04.062](https://doi.org/10.1016/j.chb.2015.04.062)]
9. Silva GM, Souto JJDS, Fernandes TP, Bolis I, Santos NA. Interventions with serious games and entertainment games in autism spectrum disorder: a systematic review. *Dev Neuropsychol* 2021;46(7):463-485. [doi: [10.1080/87565641.2021.1981905](https://doi.org/10.1080/87565641.2021.1981905)] [Medline: [34595981](https://pubmed.ncbi.nlm.nih.gov/34595981/)]
10. Carneiro T, Carvalho A, Frota S, Filipe MG. Serious games for developing social skills in children and adolescents with autism spectrum disorder: a systematic review. *Healthcare (Basel)* 2024;12(5):508 [FREE Full text] [doi: [10.3390/healthcare12050508](https://doi.org/10.3390/healthcare12050508)] [Medline: [38470619](https://pubmed.ncbi.nlm.nih.gov/38470619/)]
11. Lussier-Desrochers D, Massé L, Simonato I, Lachapelle Y, Godin-Tremblay V, Lemieux A. Evaluation of the effect of a serious game on the performance of daily routines by autistic and ADHD children. *Adv Neurodev Disord* 2023;7(6):1-13. [doi: [10.1007/s41252-023-00319-4](https://doi.org/10.1007/s41252-023-00319-4)] [Medline: [36777795](https://pubmed.ncbi.nlm.nih.gov/36777795/)]
12. Harbin SG, Davis CA, Sandall S, Fettig A. The effects of physical activity on engagement in young children with autism spectrum disorder. *Early Childhood Educ J* 2021;50(8):1461-1473. [doi: [10.1007/s10643-021-01272-4](https://doi.org/10.1007/s10643-021-01272-4)]
13. Finkelstein S, Nickel A, Barnes T, Suma E. Astrojumper: designing a virtual reality exergame to motivate children with autism to exercise. 2010 Presented at: Proceedings of the 2010 IEEE Virtual Reality Conference (VR 2010); March 20-24, 2010; Waltham p. 267-268. [doi: [10.1109/vr.2010.5444770](https://doi.org/10.1109/vr.2010.5444770)]
14. Takahashi I, Oki M, Bourreau B, Suzuki K. Designing interactive visual supports for children with special needs in a school setting. 2018 Presented at: Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18); June 11-14, 2018; Hong Kong p. 265-275. [doi: [10.1145/3196709.3196747](https://doi.org/10.1145/3196709.3196747)]
15. Graham TCN, King N, Coe H, Zabojsnikova P, Gurd BJ, Samdup D. Design and evaluation of an exergaming system for children with autism spectrum disorder: the children's and families' perspective. *Front Virtual Real* 2022;3:817303. [doi: [10.3389/frvir.2022.817303](https://doi.org/10.3389/frvir.2022.817303)]
16. Nonnis A, Bryan-Kinns N. Mazi: tangible technologies as a channel for collaborative play. 2019 Presented at: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19); May 4-9, 2019; Glasgow p. 1-13. [doi: [10.1145/3290605.3300670](https://doi.org/10.1145/3290605.3300670)]

17. Soysa AI, Al MA. Tangible play and children with ASD in low-resource countries: a case study. 2020 Presented at: Proceedings of the 14th International Conference on Tangible, Embedded, and Embodied Interaction (TEI '20); February 9-12, 2020; Sydney p. 219-225. [doi: [10.1145/3374920.3374951](https://doi.org/10.1145/3374920.3374951)]
18. Zhiglova Y. The interactive carpet—smart textile interface for children on autism spectrum disorder. 2018 Presented at: Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '18); March 18-21, 2018; Stockholm p. 712-714. [doi: [10.1145/3173225.3173341](https://doi.org/10.1145/3173225.3173341)]
19. Malihi M, Nguyen J, Cardy RE, Eldon S, Petta C, Kushki A. Short report: evaluating the safety and usability of head-mounted virtual reality compared to monitor-displayed video for children with autism spectrum disorder. *Autism* 2020;24(7):1924-1929. [doi: [10.1177/1362361320934214](https://doi.org/10.1177/1362361320934214)] [Medline: [32615771](https://pubmed.ncbi.nlm.nih.gov/32615771/)]
20. Rudovic O, Lee J, Dai M, Schuller B, Picard RW. Personalized machine learning for robot perception of affect and engagement in autism therapy. *Sci Robot* 2018;3(19):eaao6760. [doi: [10.1126/scirobotics.aao6760](https://doi.org/10.1126/scirobotics.aao6760)] [Medline: [33141688](https://pubmed.ncbi.nlm.nih.gov/33141688/)]
21. Rudovic O, Utsumi Y, Lee J, Hernandez J, Castelló Ferrer E. CultureNet: a deep learning approach for engagement intensity estimation from face images of children with autism. 2018 Presented at: Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2018); October 1-5, 2018; Madrid p. 339-346. [doi: [10.1109/iros.2018.8594177](https://doi.org/10.1109/iros.2018.8594177)]
22. Telisheva Z, Zhanatkyzy A, Turarova A, Rakhymbayeva N, Sandygulova A. Automatic engagement recognition of children within robot-mediated autism therapy. 2020 Presented at: Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20); March 23-26, 2020; Cambridge p. 471-472. [doi: [10.1145/3371382.3378390](https://doi.org/10.1145/3371382.3378390)]
23. Scassellati B, Boccanfuso L, Huang C, Mademtzis M, Qin M, Salomons N. Improving social skills in children with ASD using a long-term, in-home social robot. *Sci Robot* 2018;3(21):eaat7544 [FREE Full text] [doi: [10.1126/scirobotics.aat7544](https://doi.org/10.1126/scirobotics.aat7544)] [Medline: [33141724](https://pubmed.ncbi.nlm.nih.gov/33141724/)]
24. Bernardini S, Porayska-Pomsta K, Smith TJ. ECHOES: an intelligent serious game for fostering social communication in children with autism. *Inf Sci* 2014;264:41-60. [doi: [10.1016/j.ins.2013.10.027](https://doi.org/10.1016/j.ins.2013.10.027)]
25. Razavi SZ, Ali MR, Smith TH, Schubert LK, Hoque ME. The LISSA virtual human and ASD teens: an overview of initial experiments. 2016 Presented at: Intelligent Virtual Agents: 16th International Conference, IVA 2016, Los Angeles, CA, USA, September 20-23, 2016, Proceedings; September 20-23, 2016; Los Angeles p. 460-463. [doi: [10.1007/978-3-319-47665-0_55](https://doi.org/10.1007/978-3-319-47665-0_55)]
26. Bhattacharya A, Gelsomini M, Pérez-Fuster P. Designing motion-based activities to engage students with autism in classroom settings. 2015 Presented at: Proceedings of the 14th International Conference on Interaction Design and Children (IDC '15); June 21-25, 2015; Medford p. 69-78. [doi: [10.1145/2771839.2771847](https://doi.org/10.1145/2771839.2771847)]
27. Bossavit B, Parsons S. “This is how I want to learn”: high functioning autistic teens co-designing a serious game. 2016 Presented at: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16); May 7-12, 2016; San Jose p. 1294-1299. [doi: [10.1145/2858036.2858322](https://doi.org/10.1145/2858036.2858322)]
28. Caro K, Tentori M, Martinez-Garcia AI, Alvelais M. Using the FroggyBobby exergame to support eye-body coordination development of children with severe autism. *Int J Hum-Comput Stud* 2017;105:12-27. [doi: [10.1016/j.ijhcs.2017.03.005](https://doi.org/10.1016/j.ijhcs.2017.03.005)]
29. Peña O, Cibrian FL, Tentori M. Circus in Motion: a multimodal exergame supporting vestibular therapy for children with autism. *J Multimodal User Interfaces* 2021;15(3):283-299. [doi: [10.1007/s12193-020-00345-9](https://doi.org/10.1007/s12193-020-00345-9)]
30. Kosmas P, Ioannou A, Retalis S. Moving bodies to moving minds: a study of the use of motion-based games in special education. *TechTrends* 2018;62(6):594-601. [doi: [10.1007/s11528-018-0294-5](https://doi.org/10.1007/s11528-018-0294-5)]
31. Nicholson H, Kehle TJ, Bray MA, Heest JV. The effects of antecedent physical activity on the academic engagement of children with autism spectrum disorder. *Psychol Sch* 2010;48(2):198-213. [doi: [10.1002/pits.20537](https://doi.org/10.1002/pits.20537)]
32. Neely L, Rispoli M, Gerow S, Ninci J. Effects of antecedent exercise on academic engagement and stereotypy during instruction. *Behav Modif* 2015;39(1):98-116. [doi: [10.1177/0145445514552891](https://doi.org/10.1177/0145445514552891)] [Medline: [25271070](https://pubmed.ncbi.nlm.nih.gov/25271070/)]
33. Serino M, Cordrey K, McLaughlin L, Milanaik RL. Pokémon Go and augmented virtual reality games: a cautionary commentary for parents and pediatricians. *Curr Opin Pediatr* 2016;28(5):673-677. [doi: [10.1097/MOP.0000000000000409](https://doi.org/10.1097/MOP.0000000000000409)] [Medline: [27479151](https://pubmed.ncbi.nlm.nih.gov/27479151/)]
34. Kim B, Lee D, Min A, Paik S, Frey G, Bellini S. PuzzleWalk: a theory-driven iterative design inquiry of a mobile game for promoting physical activity in adults with autism spectrum disorder. *PLoS One* 2020;15(9):e0237966 [FREE Full text] [doi: [10.1371/journal.pone.0237966](https://doi.org/10.1371/journal.pone.0237966)] [Medline: [32911501](https://pubmed.ncbi.nlm.nih.gov/32911501/)]
35. Graham TCN, King N, Coo H, Zaboynikova P, Gurd BJ, Samdup D. Design and evaluation of an exergaming system for children with autism spectrum disorder: the children's and families' perspective. *Front Virtual Real* 2022;3:817303. [doi: [10.3389/frvir.2022.817303](https://doi.org/10.3389/frvir.2022.817303)]
36. *Innovating for People: Handbook of Human-Centered Design Methods*. Pittsburgh: Luma Institute, LLC; 2012.
37. Gamage V, Ennis C. Examining the effects of a virtual character on learning and engagement in serious games. 2018 Presented at: Proceedings of the 11th Annual ACM SIGGRAPH Conference on Motion, Interaction and Games (MIG '18); November 8-10, 2018; Limassol Cyprus p. 1-9. [doi: [10.1145/3274247.3274499](https://doi.org/10.1145/3274247.3274499)]
38. Whyte EM, Smyth JM, Scherf KS. Designing serious game interventions for individuals with autism. *J Autism Dev Disord* 2015;45(12):3820-3831. [doi: [10.1007/s10803-014-2333-1](https://doi.org/10.1007/s10803-014-2333-1)] [Medline: [25488121](https://pubmed.ncbi.nlm.nih.gov/25488121/)]

39. Ali S, Abodayeh A, Dhuliawala Z, Breazeal C, Park HW. Towards inclusive co-creative child-robot interaction: can social robots support neurodivergent children's creativity. 2025 Presented at: Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction (HRI '25); March 4-6, 2025; Melbourne p. 321-330. [doi: [10.1109/hri61500.2025.10974006](https://doi.org/10.1109/hri61500.2025.10974006)]
40. Gilson CB, Carter EW. Promoting social interactions and job independence for college students with autism or intellectual disability: a pilot study. *J Autism Dev Disord* 2016;46(11):3583-3596. [doi: [10.1007/s10803-016-2894-2](https://doi.org/10.1007/s10803-016-2894-2)] [Medline: [27573857](https://pubmed.ncbi.nlm.nih.gov/27573857/)]
41. Baron-Cohen S, Ashwin E, Ashwin C, Tavassoli T, Chakrabarti B. Talent in autism: hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philos Trans R Soc Lond B Biol Sci* 2009;364(1522):1377-1383 [FREE Full text] [doi: [10.1098/rstb.2008.0337](https://doi.org/10.1098/rstb.2008.0337)] [Medline: [19528020](https://pubmed.ncbi.nlm.nih.gov/19528020/)]
42. Kim S, Jo E, Ryu M, Cha I, Kim YH, Yoo H, et al. Toward becoming a better self: understanding self-tracking experiences of adolescents with autism spectrum disorder using custom trackers. 2019 Presented at: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '19); May 20-23, 2019; Trento p. 169-178. [doi: [10.1145/3329189.3329209](https://doi.org/10.1145/3329189.3329209)]
43. Gratch J, Wang N, Gerten J, Fast E, Duffy R. Creating rapport with virtual agents. 2007 Presented at: Intelligent Virtual Agents: 7th International Working Conference, IVA 2007, Paris, France, September 17-19, 2007, Proceedings; September 17-19, 2007; Paris p. 125-138. [doi: [10.1007/978-3-540-74997-4_12](https://doi.org/10.1007/978-3-540-74997-4_12)]
44. Jain S, Thiagarajan B, Shi Z, Clabaugh C, Matarić MJ. Modeling engagement in long-term, in-home socially assistive robot interventions for children with autism spectrum disorders. *Sci Robot* 2020;5(39):eaaz3791. [doi: [10.1126/scirobotics.aaz3791](https://doi.org/10.1126/scirobotics.aaz3791)] [Medline: [33022604](https://pubmed.ncbi.nlm.nih.gov/33022604/)]
45. Boyd LE, Gupta S, Vikmani SB, Gutierrez CM, Yang J, Linstead E, et al. VRSocial: toward immersive therapeutic VR systems for children with autism. 2018 Presented at: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18); April 21-26, 2018; Montreal p. 1-12. [doi: [10.1145/3173574.3173778](https://doi.org/10.1145/3173574.3173778)]
46. Kim W, Seong M, Kim KJ, Kim S. Engagnition: a multi-dimensional dataset for engagement recognition of children with autism spectrum disorder. *Sci Data* 2024;11(1):299 [FREE Full text] [doi: [10.1038/s41597-024-03132-3](https://doi.org/10.1038/s41597-024-03132-3)] [Medline: [38491000](https://pubmed.ncbi.nlm.nih.gov/38491000/)]
47. Suárez-Manzano S, Ruiz-Ariza A, de Loureiro NEM, Martínez-López EJ. Effects of physical activity on cognition, behavior, and motor skills in youth with autism spectrum disorder: a systematic review of intervention studies. *Behav Sci (Basel)* 2024;14(4):330 [FREE Full text] [doi: [10.3390/bs14040330](https://doi.org/10.3390/bs14040330)] [Medline: [38667126](https://pubmed.ncbi.nlm.nih.gov/38667126/)]
48. Castaño PRL, Suárez DPM, González ER, Robledo-Castro C, Hederich-Martínez C, Cadena HPG. Effects of physical exercise on gross motor skills in children with autism spectrum disorder. *J Autism Dev Disord* 2024;54(8):2816-2825. [doi: [10.1007/s10803-023-06031-5](https://doi.org/10.1007/s10803-023-06031-5)] [Medline: [37410256](https://pubmed.ncbi.nlm.nih.gov/37410256/)]
49. Xing Y, Wu X. Effects of motor skills and physical activity interventions on motor development in children with autism spectrum disorder: a systematic review. *Healthcare (Basel)* 2025;13(5):489 [FREE Full text] [doi: [10.3390/healthcare13050489](https://doi.org/10.3390/healthcare13050489)] [Medline: [40077051](https://pubmed.ncbi.nlm.nih.gov/40077051/)]
50. Ali M, Razavi SZ, Langevin R, Mamun AI, Kane B, Rawassizadeh R, et al. A virtual conversational agent for teens with autism spectrum disorder: experimental results and design lessons. 2020 Presented at: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents (IVA '20); October 19-23, 2020; Virtual Event p. 1-8. [doi: [10.1145/3383652.3423900](https://doi.org/10.1145/3383652.3423900)]
51. Mower E, Black MP, Flores E, Williams M, Narayanan S. Rachel: design of an emotionally targeted interactive agent for children with autism. 2011 Presented at: 2011 IEEE International Conference on Multimedia and Expo (ICME); July 11-15, 2011; Barcelona p. 1-6. [doi: [10.1109/icme.2011.6011990](https://doi.org/10.1109/icme.2011.6011990)]
52. Kern L, Koegel RL, Dyer K, Blew PA, Fenton LR. The effects of physical exercise on self-stimulation and appropriate responding in autistic children. *J Autism Dev Disord* 1982;12(4):399-419. [doi: [10.1007/BF01538327](https://doi.org/10.1007/BF01538327)] [Medline: [7161239](https://pubmed.ncbi.nlm.nih.gov/7161239/)]
53. Anderson-Hanley C, Tureck K, Schneiderman RL. Autism and exergaming: effects on repetitive behaviors and cognition. *Psychol Res Behav Manag* 2011;4:129-137 [FREE Full text] [doi: [10.2147/PRBM.S24016](https://doi.org/10.2147/PRBM.S24016)] [Medline: [22114543](https://pubmed.ncbi.nlm.nih.gov/22114543/)]
54. Fang Q, Aiken CA, Fang C, Pan Z. Effects of exergaming on physical and cognitive functions in individuals with autism spectrum disorder: a systematic review. *Games Health J* 2019;8(2):74-84. [doi: [10.1089/g4h.2018.0032](https://doi.org/10.1089/g4h.2018.0032)] [Medline: [30332294](https://pubmed.ncbi.nlm.nih.gov/30332294/)]
55. Kim W, Seong M, DelPreto J, Matusik W, Rus D, Kim S. Exploring potential application areas of artificial intelligence-infused system for engagement recognition. 2024 Presented at: Companion of the 2024 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp/ISWC '24); October 5-9, 2024; Athens p. 803-808. [doi: [10.1145/3675094.3678376](https://doi.org/10.1145/3675094.3678376)]
56. O'Guinn KN, Akers J, Gerencser K. Interventions targeting interactive play in individuals with autism: a systematic review. *Rev J Autism Dev Disord* 2023;12(3):422-436. [doi: [10.1007/s40489-023-00392-4](https://doi.org/10.1007/s40489-023-00392-4)]
57. Meindl JN, Delgado D, Casey LB. Increasing engagement in students with autism in inclusion classrooms. *Child Youth Serv Rev* 2020;111:104854. [doi: [10.1016/j.childyouth.2020.104854](https://doi.org/10.1016/j.childyouth.2020.104854)]

Abbreviations

APE: adapted physical education
AR1: autoregressive working correlation of order 1
ASD: autism spectrum disorder
GEE: generalized estimating equations
HCI: human-computer interaction
SG: serious game
VA: virtual agent

Edited by S Brini; submitted 19.May.2025; peer-reviewed by T Krasovsky, D Bethere; comments to author 26.Nov.2025; revised version received 11.Jan.2026; accepted 13.Jan.2026; published 02.Mar.2026.

Please cite as:

Kim W, Seong M, Kim S

Design and Iterative Development of Serious Exergames for Children With Autism Spectrum Disorder: Formative Multiple-Case Pilot Study

JMIR Serious Games 2026;14:e77727

URL: <https://games.jmir.org/2026/1/e77727>

doi:10.2196/77727

PMID:

©Won Kim, Minwoo Seong, SeungJun Kim. Originally published in JMIR Serious Games (<https://games.jmir.org>), 02.Mar.2026. 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.

Immersive Tai Chi for Home-Based Exercise in Older Adults: Usability and Feasibility Study

XiaCheng Song¹, PhD; Nazlena Mohamad Ali¹, PhD; Mohamad Hidir Mhd Salim¹, PhD; Muhammad Yudhi Rezaldi^{2,3}, PhD

¹Institute of Visual Informatics, National University of Malaysia, Aras 1, Bangi, Selangor, Malaysia

²Research Centre for Data and Information Sciences, National Research and Innovation Agency, Bandung, Indonesia

³Faculty of Creative Industry, Telkom University, Bandung, Indonesia

Corresponding Author:

Nazlena Mohamad Ali, PhD

Institute of Visual Informatics, National University of Malaysia, Aras 1, Bangi, Selangor, Malaysia

Abstract

Background: Longer life expectancy makes physical exercise crucial for active aging. However, adherence to traditional exercise among community-dwelling older adults is generally low. Virtual reality (VR) and mixed reality (MR) Tai Chi exergames, as novel health promotion tools, show significant potential, particularly for older adults exercising in a home setting.

Objective: This study aimed to evaluate the usability and feasibility of a VR and MR Tai Chi exergame for community-dwelling older adults, focusing on subjective experience, physiological comfort, and objective interaction performance. The study also explored the relationships between key usability factors and sought to quantify links between objective accuracy and subjective experience (the Game Experience Questionnaire or Virtual Reality Sickness Questionnaire [VRSQ]) to inform choices of display mode, feedback strength, and session length.

Methods: Of the 86 community-dwelling older adults recruited for this study, data from 70 participants were considered valid after an initial screening, during which 16 (18.6%) were excluded due to issues with VR adaptation. Participants were sequentially assigned in a rotating order to 1 of 4 variants (VR/MR×soothing/intense) to balance exposure; however, primary analyses were preplanned to be collapsed across variants, focusing on whole-sample usability and feasibility rather than confirmatory between-group hypotheses. The primary outcome measures included the Game Experience Questionnaire, VRSQ, and objective gameplay logs.

Results: The VR/MR Tai Chi game demonstrated good overall usability and acceptability among the screened participants. Subjective experience was highly positive, with median scores for “positive affect” (median 4.0) and “competence” (median 3.8) being significantly high, whereas the median for “Challenge” (median 1.4) was significantly low ($P<.001$ for all). Physiological comfort in the postscreening sample was acceptable, with the most common mild symptoms being dizziness with eyes closed (20.0%) and vertigo (18.6%), both of low severity; however, the initial exclusion of 18.6% of participants due to VR discomfort is noteworthy. Therefore, generalizability is limited because the analyzed sample overrepresents older adults who tolerate immersive displays. Accuracy showed significant positive correlations with flow ($\rho=0.342$) and competence ($\rho=0.322$), whereas the VRSQ total score was significantly negatively correlated with positive affect ($\rho=-0.334$, $P=.005$).

Conclusions: Tai Chi exergames based on immersive technologies offer a feasible and attractive pathway for promoting physical exercise among community-dwelling older adults, particularly within the home environment, supporting the goal of aging in place. As a single, laboratory-based session, the reported satisfaction may partly reflect a novelty effect; therefore, longer-term, home-based follow-ups are needed to assess durability. Analysis of the key usability factors provides guidance for specific design choices, while also indicating directions for future research, such as longitudinal evaluations, extension to more diverse populations, and application in real-world home settings.

(*JMIR Serious Games* 2026;14:e79453) doi:[10.2196/79453](https://doi.org/10.2196/79453)

KEYWORDS

virtual reality; video games; aged; Tai Chi; usability testing; motion sickness

Introduction

The trend of people getting older around the world is accelerating. By 2050, there will be more than 2 billion people

over the age of 60 years, and older adults will make up a larger share of the world's population, especially in developing areas [1]. This change in the population puts more pressure on the economy and shows how important programs such as “Healthy

Aging” and letting older people “age in place” are. Regular exercise is very important for healthy aging [2]. It helps keep people healthy, helps them stay independent, makes them feel better mentally, and improves their quality of life [3]. These benefits make exercise a good and useful way to help older people all over the world. However, older adults often face numerous barriers that make it hard for them to participate in traditional exercise programs. These problems include not being motivated, being afraid of falling, and not being able to get to the right facilities [4]. These challenges show that conventional exercise models often fail to meet the emotional and experiential needs of older adults. Recent studies suggest that digitally delivered, personalized interventions, such as programs incorporating wearable technology, remote coaching, and motivational messaging, can significantly improve adherence and reduce fear of falling [5,6]. This points to the need for approaches that improve enjoyment and intrinsic motivation in older adults.

To tackle this issue, immersive exergames that combine virtual reality (VR) and mixed reality (MR) technologies are a viable solution. This research concentrates on the integration of Tai Chi with VR/MR technologies. Tai Chi is particularly suitable for older adults due to its proven efficacy in enhancing dynamic balance and substantially decreasing both the frequency of falls and the fear of falling, which is a significant psychological impediment to physical activity engagement in this demographic [7]. These advantages render Tai Chi a pragmatic and psychologically beneficial approach for enhancing mobility and self-assurance in older adults. VR/MR Tai Chi exergames are meant to be a safe, easy, and fun way to work out at home. These exergames can address many limitations of traditional exercise by providing virtual practice environments that are controllable. They may also boost intrinsic motivation by enhancing feelings of autonomy and competence. Although extended reality technologies hold substantial promise for helping older adults, several challenges remain, such as the high cost of equipment, steep learning curves, and cybersickness. Notably, previous research has emphasized that cybersickness is a frequent and underestimated barrier that may lead to user dropout and restrict broader adoption in unsupervised home settings [8,9]. These usability and safety issues should be addressed before extended reality solutions can be widely implemented among older populations.

Although VR and MR exergames show promise for helping older adults avoid falling, there is insufficient systematic research comparing how different display modes and feedback systems affect this group. Some early evidence suggests that augmented reality, a type of MR, may be better for older adults because it gives them a better sense of their body and space, which makes them feel less likely to fall [10]. This points to a research gap in comprehending the preferences and reactions of older users to various immersive environments. Recognizing the distinct interactive experiences of older adults is essential for effective design, particularly in unsupervised home contexts. Hence, distinguishing between VR and MR modes goes beyond being a mere technical matter; it has direct implications for the safety, comfort, and acceptance of these technologies among older adults using them independently at home.

This work provides evidence to guide the development of immersive health care applications to enable older individuals to practice independently at home, a key component within the “aging in place” plan. With professional guidance lacking within home contexts, intrinsic motivational abilities and efficacy and safety of the game within itself become paramount [11]. Thus, this work’s investigation into aspects such as cybersickness and “accuracy-competence paradox” remains a key focal point within this work to aid in determining if such technology can be disseminated and leveraged effectively.

Specifically, this study will seek to answer the following core questions:

1. Overall usability and acceptance: How does age affect subjective experience (eg, “empowerment” feelings), physiological comfort (ie, cybersickness symptom occurrence), and objective interaction efficacy (eg, the “accuracy-competence paradox”) while playing the VR/MR Tai Chi game?
2. Internal links between fundamental factors: How does objective gameplay performance (eg, pointing precision) influence subjective psychological experiences (eg, flow, competence)? Do users’ physiological symptoms of discomfort affect positive emotions? Are positive emotions correlated to users’ inherent traits, such as age?

Through systematic investigation of these questions, this study aimed to contribute empirical evidence to be used in designing immersive health applications among older people [12].

Methods

Ethical Considerations

The study received ethics approval and adhered to guidelines for research involving human participants.

Research Ethics Approval

This study received ethics approval from the Institute of Visual Informatics, Universiti Kebangsaan Malaysia. The approval was granted on July 2, 2024, with the reference number UKM.IVI.600-4/6/P130610.

Informed Consent

All participants provided written informed consent before the start of any experimental procedures. The researchers clarified the study’s objectives, methods, possible risks (including cybersickness symptoms), and potential benefits. They also made it clear to participants that they could withdraw from the study at any point without facing any consequences.

Privacy and Confidentiality

All personally identifiable information and harvested data were fully anonymized to ensure respondent privacy. Data were stored securely on a password-protected server, and only central research team staff had access.

Compensation

Participants did not receive any monetary or nonmonetary compensation for their participation. Their involvement in this study was entirely voluntary.

Study Design

Overview

This study was designed as a pilot study using a 2×2 factorial design (display mode: VR/MR×feedback: soothing/intense) to effectively explore the feasibility of varying exergame configurations before a large trial. Participants were sequentially assigned to 1 of the 4 conditions using a rotating 1:1:1:1 schedule upon enrollment. The primary goal of this multicondition design was not to conduct formal hypothesis testing between groups, which would require a much larger sample. Instead, its purpose was to ensure that our usability and feasibility assessment covered a range of potential interaction styles. This approach allows us to identify any major negative outcomes (eg, excessive cybersickness or frustration) that might be unique to a specific condition, providing critical insights for optimizing a single, refined protocol for future, larger-scale home-based studies.

Recruitment, Screening, and Final Sample

A research team initially recruited 86 community-dwelling older people who were capable of living independently, via media such as community promotions at community centers and newspaper advertising locally. Such a process was used to maintain the sample's characteristics closely aligned with those of the target end user group for future home-based VR/MR Tai Chi exergames [13].

In this pilot phase of research, screening led to a dropout of 16 participants (16/86, 18.6% of the original sample) who could not accommodate to the virtual environment or who experienced severe dizziness. Notably, this screening process suggests that the final sample was skewed to some degree to VR tolerance, that is, comprised individuals better predisposed to accommodate virtual environment exposures. This characteristic is relevant in estimating the extent to which the findings generalize to a broader community of older individuals.

We performed an analysis of the 16 participants who were excluded due to a failure to adapt to VR. This group comprised 11 (68.8%) women and 5 (31.2%) men, a gender distribution similar to that of the retained sample. However, a systematic

difference was evident in age: 43.8% (7/16) of participants excluded were aged older than 75 years, a proportion more than double that of the sample maintained (15/70, 21.4%). This finding suggests that older age could be a determining factor in intolerance to VR in this population.

Inclusion and Exclusion Criteria

The aforementioned recruitment and screening processes adhered to clearly defined criteria. The inclusion criterion was enrollment of participants aged 60 years and older [14,15]. Exclusion criteria encompassed any known adverse reactions to VR experiences and severe cardiorespiratory or musculoskeletal diseases that would significantly impede participation in the experimental procedures, severe vision loss, and other health conditions unsuitable for the use of immersive devices (eg, vertigo and epilepsy).

Self-report alone was considered insufficient; therefore, a brief experiential screening was recommended before unsupervised use.

Demographic Characteristics, Health Status, and Exercise Habits

Data regarding demographic information, health background, and daily exercise habits were collected from the 70 valid participants using a self-developed "basic information questionnaire." The detailed characteristics are presented in Table 1.

The 70 valid participants included in the final analysis exhibited certain demographic tendencies: they were predominantly female (52/70, 74.29%) and largely comprised "younger" older adults aged 60 to 75 years (55/70, collectively 78.56%). In terms of health status, although a considerable proportion (40/70, 57.14%) self-reported chronic conditions such as cardiopulmonary diseases, they were also individuals who actively managed their health, with a high percentage (66/70, 94.29%) undergoing regular physical examinations. This characteristic of "living actively with illness" suggests that the sample group might hold a more open attitude toward health-promoting interventions, but it also necessitates that such interventions ensure safety and appropriateness.

Table . Participant demographic characteristics, health status, and exercise habits (N=70).

Characteristics	Values, n (%)
Sex	
Male	18 (26)
Female	52 (74)
Age distribution (y)	
60 - 65	32 (46)
65 - 70	11 (16)
71 - 75	12 (17)
75 - 80	11 (16)
>80	4 (6)
Height distribution (cm)	
140 - 150	13 (19)
150 - 160	36 (51)
160 - 170	20 (29)
>170	1 (1)
Weight distribution (kg)	
30 - 40	2 (3)
40 - 50	8 (11)
50 - 60	36 (51)
60 - 70	18 (26)
70 - 80	4 (6)
>80	2 (3)
Self-reported illness status	
No reported cardiopulmonary illness	30 (43)
With cardiopulmonary disease	40 (57)
Regular physical examination	
Yes	66 (94)
No	4 (6)
Weekly exercise frequency	
Exercise daily	66 (94)
Exercise 3 - 5 times per week	1 (1)
Rarely exercise	3 (4)

This pilot group reported very high levels of physical activity, such that the vast majority (66/70, 94.29%) used a daily exercise regimen and participated in a broad range of activities, such as walking, square dancing, and Tai Chi. Overall, this sample is a subgroup of predominantly female, aged between 60 and 75 years, who exercise on a regular basis, approach health care management in an active manner, and have good tolerance to VR environment (as a consequence of preliminary screening that eliminated 18.6% (16/86) who reported dizziness while experiencing VR). These characteristics are relevant to explain why this sample reports such a high level of interest and positive response to the VR Tai Chi game. While these same characteristics will limit how widely the investigation can be

generalized to broader older groups, such as those who exercise less or who possess poorer VR tolerance.

VR Tai Chi Exergame and Experimental Equipment

Intervention and Equipment

The VR Tai Chi exergame used in this study was designed for older individuals. Its fundamentals can be equally analyzed using the MDA (mechanics, dynamics, and aesthetics) approach [16,17].

Mechanics

It guides users through 16 selected Tai Chi exercises. It is interactive via a head-mounted display (HMD) and hand

controllers reacting to head and hand movements. Users must synchronize their pose to 3 hotspots (head and both hands) provided in the on-screen instruction. A movement only proceeds after a successful success window, and a movement can save up to 3 hotspot holds (left hand, right hand, and HMD) before advancing. If users complete the entire sequence of 16 movements ahead of schedule, they will repeat the sequence immediately, beginning with the first movement. This generous success condition was a conscious design choice intended to provide a sense of accomplishment. It has 2 display modes, VR and MR, and can come equipped with either soothing or upbeat music and vibration feedback. Additionally, players can skip the current move or pause at will during gameplay.

Dynamics

Playing involves players' efforts at mirroring exemplified postures, receiving immediate multisensory feedback, and regulating progression during a trial lasting a maximum of 5 minutes. One key dynamic reveals itself in the gamification mechanics: participants can choose to meet only the bare minimum threshold needed for progress ("single hotspot for 3 s") to progress effortlessly or shoot for immaculate matching across all hotspots to progress to a higher "accuracy" level. This allows participants to balance challenge against victory on a preference and ability level. It is this dynamic that later results regarding user experience build upon.

Aesthetics

The game aims to evoke multidimensional aesthetic experiences, such as a "sense of competence" from successfully completing movements, a "sense of challenge" from improving performance, feelings of "relaxation or motivation" engendered by different musical feedback, and a "sense of immersion" provided by the VR/MR environments. These subjective experiences are intended to encourage physical exercise in older users in an engaging manner, and their achievement was assessed using the Game Experience Questionnaire (GEQ) [18,19]. Design philosophy appears to be interested in desirable affective effects and perceived competence corresponding to self-determination theory principles, such that satisfying the need for competence can improve intrinsic motivation.

Experimental Equipment

This study used the Meta Quest 3 (Meta Platforms, Inc) HMD for presenting both VR and MR conditions [20,21]. The choice of specific, commercially available hardware, such as the Meta Quest 3, enhances the transparency and potential replicability of the study.

Activity Area

Exercise took place in an area of approximately 5 m², comparable to a typical living room, to ensure home feasibility and safety [22,23].

Experimental Procedure

Group allocation followed a prespecified rotating schedule (1:1:1:1) prepared before data collection; participants had no influence on assignment.

Experimental Preparation and Adaptation Training

When participants arrived at the laboratory, they first completed an informed consent process and a general information questionnaire. Next, a short adaptation training session was completed to enable participants (and especially those without any training in VR/MR) to become familiar with both the immersive environment and how to use the equipment. This adaptation step is a valuable step in acclimating older adults to new technology, such as VR, because it can be used to help reduce technology anxiety and achieve a minimum baseline level of comfort and competence in how to use the equipment before beginning the intervention [24,25].

Exergame Intervention

A visual summary of the participant flow and 2×2 design (VR/MR×soothing/intense) appears in Figure 1. This approach ensured that group sizes remained balanced and that potential confounding factors, such as personal preference or prior experience, were minimized. Participants followed a fixed 16-movement sequence; each 5-minute session could end before all movements were finished due to the time limit or restart at the first movement if the sequence was completed early. After this orientation, each participant performed 2 formal 5-minute gameplay sessions under their assigned experimental condition (VR/MR×soothing/intense feedback). The better performance out of the 2 sessions was recorded for analysis. All single-session metrics (eg, posture, accuracy) were computed per session; the higher of the 2 sessions was retained for analysis. This procedure ensured that every participant experienced the complete set of movements while also minimizing the impact of first-trial unfamiliarity on the final outcome [26].

VR provided a fully virtual scene; MR overlaid guidance onto the real room (Figure 2). "Soothing" used gentle cues; "intense" used more salient multimodal cues; task rules were identical [27].

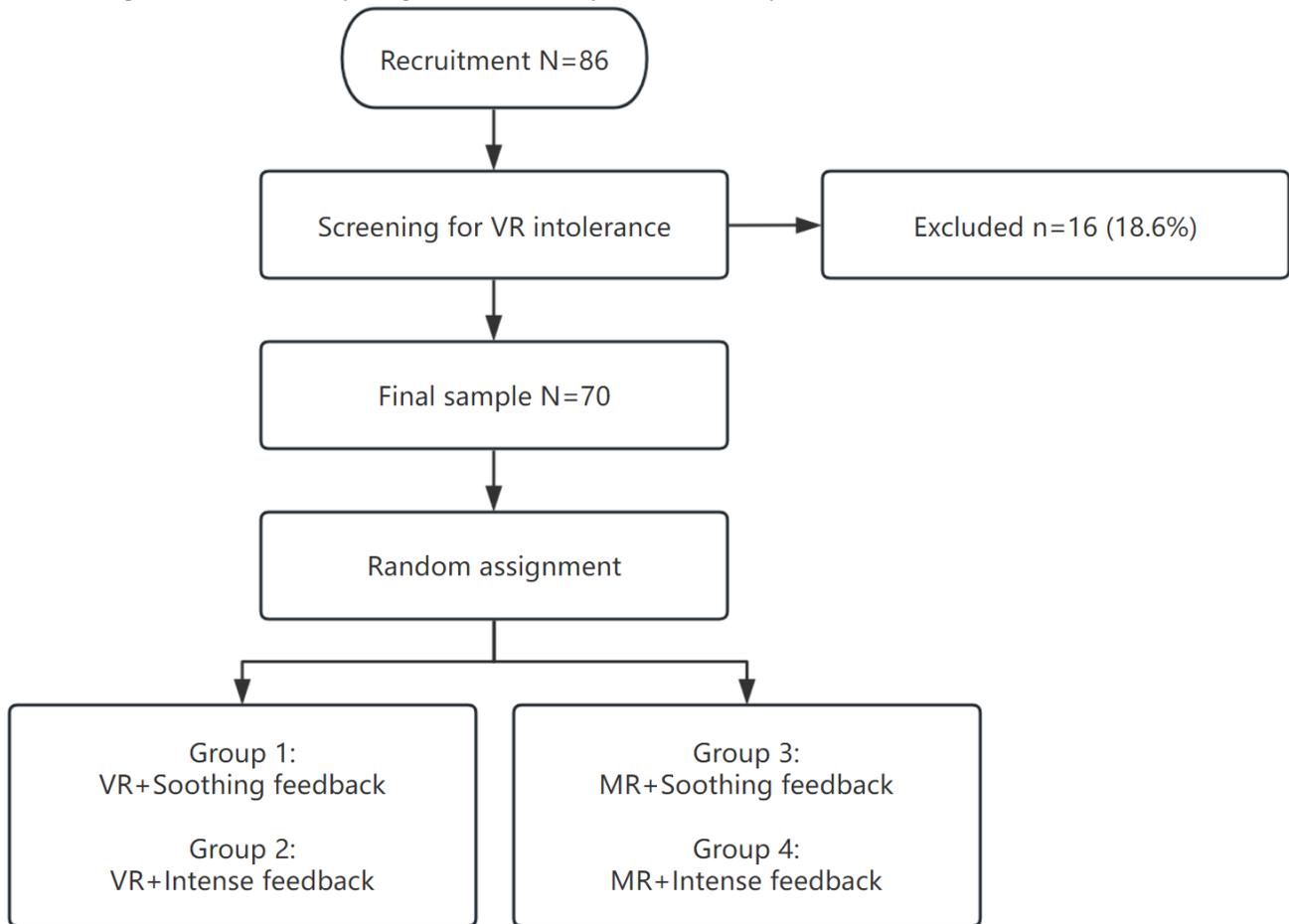
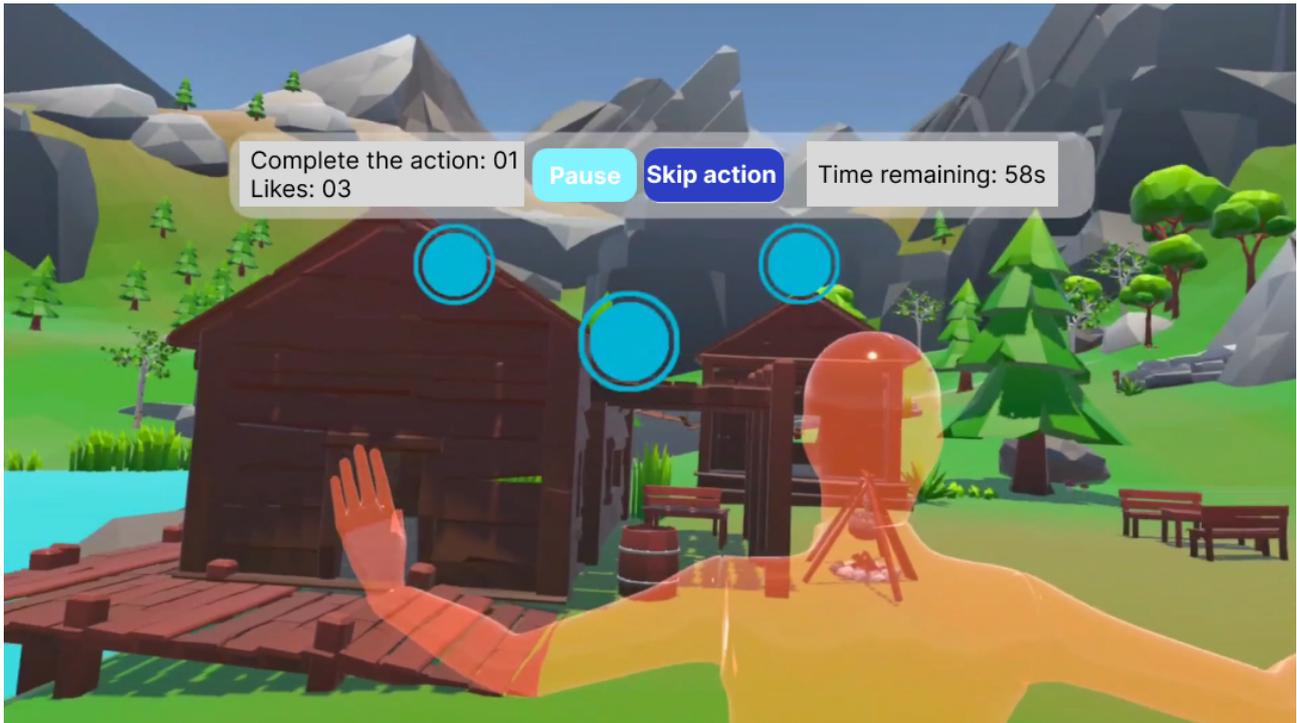
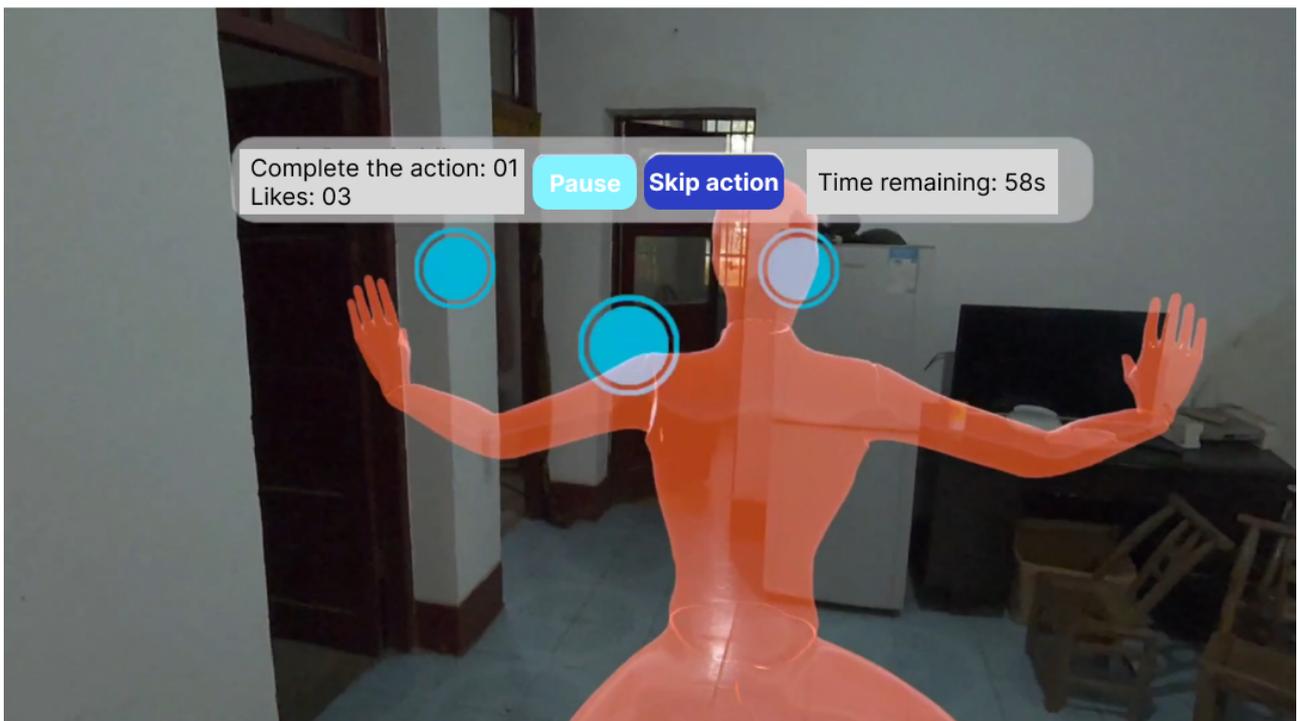
Figure 1. Participant Flow and 2x2 Study Design. MR: mixed reality; VR: virtual reality.

Figure 2. Screenshots of the VR and MR display modes. MR: mixed reality; UI: user interface; VR: virtual reality.



VR mode UI



MR mode UI

Safety and Supervision

The experimental protocol was observed and supervised in a one-on-one capacity by trained researchers to be able to control and respond to any new sensory stimulus or sensation of disorientation. Subjects were informed they could withdraw from participation at any time if they experienced any type of

discomfort or abnormal level of fatigability. Such precautions are particularly relevant in research among older individuals and potentially disorienting technologies such as VR. To regulate fatigue and potential dizziness between sessions, exposure was scheduled in 2 approximately 5-minute sessions, and participants could rest and pause whenever needed.

This was a single-visit, supervised laboratory study, consisting of the 2 gameplay blocks described earlier, with no home-based or longitudinal follow-up. Postsession questionnaires captured immediate responses; repeated exposures to control for novelty were not implemented.

Outcome Measures

Overview

This research took a multidimensional perspective to assess usability and feasibility, where primary data were drawn from:

1. GEQ [18]: An adapted 32-item version of the basic module of the GEQ (skipping item 3 in the original 33-item) was used to measure seven aspects of game experience: flow, competence, positive affect, negative affect, tension/annoyance, challenge, and sensory/imaginative immersion. Item 3 (“I was interested in the game’s story”) was excluded from the component immersion because our own game did not have a story. Component flow and competence scores were of particular importance to intrinsic motivation. Component scores were calculated as the means of their items on a scale ranging from 0 to 4.
2. The Virtual Reality Sickness Questionnaire (VRSQ) [28]: It was used to detect and quantify physiological symptom indicators of distress that users may experience while controlling VR/MR hardware, such as eye strain, nausea, headache, disorientation, and dizziness. It is a 9-item measure scored on a 0- to 3-point scale, and the VRSQ total score was calculated as the mean of these items, which provided a final score between 0 and 3. It has been shown to be a valid measure in the original investigation of the

VRSQ for cybersickness because it was very highly correlated with existing motion sickness indices.

3. Objective game performance logs: Primary objective metrics were derived from automatically collected gameplay data provided by the game’s backend system. They were “game score,” “movement accuracy,” “mean time per round,” “task completion rate,” “movement level (extracted from hand movement and head movement trajectory),” and “cumulative game time.” These data were direct evidence to substantiate the assessment of whether users could use the game in an efficient manner.

We took an integrated approach, triangulating subjective experience (GEQ), physical comfort (VRSQ), and actual performance (game logs) to provide a sound foundation for a unified and rigorous assessment of the game’s usability and feasibility.

We adopted a target engineering level of 80% movement precision to be used as an internal and exploration benchmarking.

Psychometric Properties of Outcome Scales

Internal consistency was evaluated on the present sample using Cronbach α and Spearman-Brown split-half coefficients. As several components (tension/annoyance, challenge, and negative affect) exhibited pronounced floor effects with near-zero variance, internal consistency coefficients for these scales were not estimable; for these, we report distributional summaries only. Reliability results for all measures are presented in Table 2. Per the GEQ guidelines, we dropped the story-related immersion item (item 3), leaving 5 items for this component; reliability values in Table 2 reflect this adapted set.

Table . Internal consistency of Game Experience Questionnaire and Virtual Reality Sickness Questionnaire (VRSQ) subscales.

Measure	k (items)	Cronbach α	Split-half (Spearman-Brown)
Sensory and imaginative immersion	5 ^a	0.657	0.583
Flow	5	0.546	0.539
Tension/annoyance	3	N/A ^b (floor effect)	N/A ^b (floor effect)
Challenge	5	N/A ^b (floor effect)	N/A ^b (floor effect)
Negative affect	4	N/A ^b (floor effect)	N/A ^b (floor effect)
Positive affect	5	0.86	0.85
VRSQ total	9	0.782	0.689

^aGame Experience Questionnaire core scored 0 - 4; immersion used items 12, 18, 19, 27, 30 (item 3 removed). Split-half denotes Spearman-Brown.

^bN/A indicates nonestimable due to floor effect (near-zero variance).

Operational Definitions

The key game performance indicators were defined as follows: score was the system-generated composite points; accuracy (%) was the proportion of frames where hotspots were within tolerance; and posture (count) represented the total number of successful hotspot holds. Per movement, the player aligns the left and right controllers and the HMD with 3 predefined hotspots; holding any hotspot for ≥ 3 seconds is counted as one successful hotspot hold. The 16-movement sequence loops whenever it is completed before the 5-minute limit; therefore,

posture is an open-ended count (no fixed cap) within the time window. For interpretability, 48 holds correspond to one full loop (16 movements \times 3 hotspots).

Data Preparation and Statistical Analysis

Overview

Consistent with the a priori plan, no confirmatory between-condition hypothesis testing (main effects or interaction) was conducted. To enhance transparency,

per-condition descriptive summaries (medians and IQRs for the 4 cells) are provided in [Table 3](#).

Table 3. Per-condition descriptive statistics.

Outcome ^a	Group A: VR ^b soothing (n=17), median (IQR)	Group C: VR intense (n=19), median (IQR)	Group B: MR ^c soothing (n=17), median (IQR)	Group D: MR intense (n=17), median (IQR)
Positive affect	4 (3.8 - 4)	4 (3.8 - 4)	4 (3.4 - 4)	4 (3.8 - 4)
Flow	4 (3.2 - 4)	3.8 (3.6 - 4)	4 (3.2 - 4)	4 (3.2 - 4)
Competence	3.6 (3.1 - 4)	3.8 (3.2 - 4)	3.6 (3.4 - 4)	4 (3.3 - 4)
Challenge	1.4 (1.2 - 1.6)	1.4 (1 - 1.8)	1.6 (1.1 - 1.7)	1.6 (0.8 - 1.7)
VRSQ ^d total	0 (0 - 1)	0 (0 - 2)	0 (0 - 1.5)	0 (0 - 2)
Accuracy (%)	75 (56-84)	65 (54-81)	65 (54-77)	79 (54-94)

^aPer-condition n shown in column headers. No inferential tests were planned; descriptive summaries only.

^bVR: virtual reality.

^cMR: mixed reality.

^dVRSQ: Virtual Reality Sickness Questionnaire.

Data Cleaning

Handled sporadic missingness via median imputation only; no multiple imputation was performed [29].

Analysis Plan

Data normality was examined using the Shapiro-Wilk test [30,31]. Most key variables were found not to be normally distributed ([Table 4](#)). Nonparametric methods were consequently unanimously selected across hypothesis testing to yield robust

and valid inferences. That is, median dimension scores on the GEQ were compared against a theoretical midpoint score of 2.0 using 1-sample Wilcoxon signed-rank tests. Wilcoxon tests correspondingly compared the median against a theoretical score of 0 on the VRSQ symptom scores (rated 0 - 3) to determine whether symptoms were distinctly present. With the exploratory nature of this item-level analysis, no adjustment for multiple comparisons was made on these 9 tests. Spearman ρ was used to examine associations between key factors. A level of significance was predetermined at an α of .05.

Table . Shapiro-Wilk test of normality for Game Experience Questionnaire (GEQ) dimensions, Virtual Reality Sickness Questionnaire (VRSQ) symptoms, and game performance indicators.

	W	N	P value
Overall GEQ dimension score			
Competence	0.798	70	<.001
Sensory and imaginative immersion	0.777	70	<.001
Flow	0.719	70	<.001
Tension/annoyance	0.275	70	<.001
Challenge	0.974	70	.16
Negative affect	0.573	70	<.001
Positive affect	0.551	70	<.001
VRSQ main symptom			
General discomfort	0.205	70	<.001
Fatigue	0.28	70	<.001
Headache	0.158	70	<.001
Eye fatigue	0.431	70	<.001
Difficulty concentrating	0.153	70	<.001
Head fullness	0.346	70	<.001
Blurred vision	0.349	70	<.001
Dizziness when closing eyes	0.512	70	<.001
Dizziness/vertigo	0.494	70	<.001
Key game performance			
Score	0.967	70	.06
Posture	0.967	70	.06
Accuracy	0.962	70	.03
Average time spent	0.796	70	<.001
Movement distance (m)	0.979	70	.29

Overall Usability Analysis

Descriptive statistics summarized the GEQ, the VRSQ, and game log indicators as median (IQR) [19,32].

Results

Participant Flow and Analysis Overview

In total, 86 older adults were recruited; 16 (18.6%) discontinued during acclimation due to intolerance, and data from 70 (81.4%) participants were analyzed. Participant flow and group allocation are shown in Figure 1. Per-condition medians (IQRs) for the 4 cells are available in Table 3.

Scale Reliability

Internal consistency was good for positive affect ($\alpha=.860$; split-half=0.850) and acceptable for the VRSQ total ($\alpha=.782$; split-half=0.689). Competence ($\alpha=.667$; split-half=0.750) and sensory/imaginative immersion ($\alpha=.657$; split-half=0.583) were in the moderate range, whereas flow was borderline ($\alpha=.546$; split-half=0.539). Coefficients for tension/annoyance, challenge,

and negative affect were not estimable due to floor effects (near-zero variance; Table 2).

Before the main analysis, the assumption of normality for the key-dependent variables was assessed using the Shapiro-Wilk test. The results are detailed in Table 4. While a few indicators did not significantly deviate from normality (eg, *challenge*, $P=.16$; game log *score*, $P=.06$; *posture*, $P=.06$; *movement distance (m)* $P=.29$), the majority of variables substantially deviated (eg, *accuracy*, $P=.032$; most GEQ/VRSQ dimensions, $P<.001$). Therefore, to maintain consistency and ensure robust inference across all analyses, nonparametric procedures were uniformly adopted.

Descriptive Statistics

We reported per-condition descriptive statistics for key outcomes across the 2×2 design (VR/MR×feedback intensity). As planned, no between-condition inferential tests were conducted; values are summarized as median (IQR) for interpretability with nonnormal distributions (see Table 3 for the 4 cells).

Overall Usability and Acceptability of the VR Tai Chi Game

To evaluate the overall usability and acceptability of the VR Tai Chi game for older users, this study collected and analyzed survey data from GEQ, VRSQ, and objective game performance log data.

Subjective Game Experience (GEQ)

As reflected in Table 5, subjective gaming experience as measured using the GEQ was highly positive. The median levels

on positive aspects like “positive affect” (median 4.0), “flow” (median 4.0), “competence” (median 3.8), and “sensory/imaginative immersion” (median 3.8) all far surpassed the theoretical midpoint of the scale ($P < .001$ in all). Medians on negative aspects such as “tension/annoyance” (median 0) and “negative affect” (median 0) were below midpoint to a statistically significant degree ($P < .001$ in both). The median level on “challenge” was low to a statistically significant degree as well (median 1.4, $P < .001$).

Table . Descriptive statistics and one-sample Wilcoxon signed-rank test for Game Experience Questionnaire dimensions.

Dimension	Values, median (IQR ^a)	<i>P</i> value ^b
Competence	3.8 (3.2 - 4)	<.001
Sensory/imaginative immersion	3.8 (3.2 - 4)	<.001
Flow	4 (3.2 - 4)	<.001
Tension/annoyance	0 (0 - 0)	<.001
Challenge	1.4 (1 - 1.6)	<.001
Negative affect	0 (0 - 0.25)	<.001
Positive affect	4 (3.8 - 4)	<.001

^aGame Experience Questionnaire scores were rated on a 0 - 4 scale.

^b*P* values reflect a one-sample Wilcoxon signed-rank test comparing the median score against the theoretical midpoint of 2.0.

Physiological Comfort (VRSQ)

The physiological comfort of the screened participants was assessed using the VRSQ. Overall, the incidence and severity of reported symptoms were low. Detailed item-level medians,

statistical test results, incidence rates, and severity scores for reporters are presented in Table 6. The *P* values in the table should be interpreted as exploratory signals of which symptoms were most frequently endorsed, rather than as confirmatory hypothesis tests.

Table . Overall Virtual Reality Sickness Questionnaire main symptom statistical analysis.

Symptom item	Median	<i>P</i> value	Incidence rate (%)	Mean severity for reporters
General discomfort	0 (0-0)	.08	4.3	1
Fatigue	0 (0-0)	.04	7.1	1.4
Headache	0 (0-0)	.16	2.9	1
Eye fatigue	0 (0-0)	.002	15.7	1.273
Difficulty concentrating	0 (0-0)	.18	2.9	1.5
Head fullness	0 (0-0)	.01	10.0	1.286
Blurred vision	0 (0-0)	.01	10.0	1.143
Dizziness when closing eyes	0 (0-0)	<.001	20.0	1.214
Dizziness/vertigo	0 (0-0)	<.001	18.6	1

Game Performance

A look at the in-game objective performance data listed in Table 7 reveals wide individual variations among older adult participants. Participants demonstrated a fluid interaction rhythm with a median time spent per action of 6.4 seconds. However,

large variations across major indicators were apparent due to wide IQRs seen across score and accuracy measurements (Table 7). Median movement accuracy was 68.5% (IQR 54.2% - 82.3%). Scores higher than 48 indicate that participants completed more than a single loop (or accrued numerous hotspot holds within a movement) within a 5-minute period.

Table . Key game performance indicators.

Indicator	Values, median (IQR)
Score	96 (70.25-133.5)
Posture ^a (count; no fixed cap, 48=one loop)	47 (39.75-55)
Accuracy (%)	68.5 (54.2 - 82.3)
Average time spent (s)	6.4 (5.5-7.5)
Movement distance (m)	273.5 (231.5-317.25)

^aPosture counts the total hotspot holds within the retained 5-min session (best-of-two). The 16-movement sequence loops if completed early.

Correlation Analysis of Key Usability Factors

Correlation between major usability factors was examined using Spearman rank correlation analysis, and the results were presented in Table 8. Flow was significantly positively

correlated with accuracy ($\rho=.342$; $P=.004$), and competence was significantly positively correlated with accuracy ($\rho=.322$; $P=.007$). VRSQ total score was significantly inversely correlated to positive affect ($\rho=-0.334$; $P=.005$). Competence was not significantly correlated with age ($\rho=-0.179$; $P=.14$).

Table . Spearman ρ correlation analysis (n=70).

Hypothesis	Variable 1	Variable 2	Correlation coefficient	P value
H1	Flow	Accuracy	0.342 ^a	.004
H2	VRSQ ^b total score	Positive affect	-0.334 ^a	.005
H3	Accuracy	Competence	0.322 ^a	.007
H4	Age	Competence	-0.179	.14

^aCorrelation is significant at the 0.01 level (2 tailed).

^bVRSQ: Virtual Reality Sickness Questionnaire.

Discussion

Subjective Experience Profile: High Perceived Competence With Low Perceived Challenge

As presented in Table 5, medians for positive affect, flow, and competence were above the scale midpoint, whereas challenge was markedly below (all $P<.001$). We observe a distinct user experience profile: high perceived competence and low challenge. For older adults who begin using exergames at home, this combination may be most relevant in reducing frustration and promoting first-time engagement [33,34]. At the same time, future research should gradually increase difficulty in an adaptive manner such that motivational support is balanced against accurate skill acquisition in free-living situations [35,36].

The internal consistency profile reveals adequate internal consistency on positive affect and satisfactory consistency on total VRSQ and moderate-to-borderline level in several components of the GEQ. Borderline α on flow is likely owing to constrained variance and the brevity and heterogeneity of item material in this first-stage application involving a single session; such findings on these components should consequently be interpreted correspondingly cautiously.

Physiological Comfort (VRSQ): Acceptability in Screened Sample and Potential Challenges

Regarding physiological comfort, the VR Tai Chi game demonstrated good tolerability within the screened participant sample, but this conclusion must be interpreted with caution [37]. The initial exclusion of a high 18.6% of recruits due to

inability to adapt or strong VR-related sickness signifies a considerable bias in the VR tolerance of the final analyzed sample. Consequently, the currently observed low levels of cybersickness might overestimate the game's physiological comfort among the general older adult population. Although within the screened "more tolerant" sample, the average reported levels and perceived severity of various physiological discomfort symptoms (mainly mild "dizziness when closing eyes," "vertigo," and "eye fatigue") were low, 15% to 20% of participants still reported these mild symptoms [38]. This implies that cybersickness continues to be a primary barrier that has to be cleared ahead of VR technology deployment to older age categories. This 18.6% exclusion rate translates directly to current technological or design limitations in making VR universally comfortable among older adults and places a realistic limit on such interventions' immediate impact absent substantive improvements. For plans to promote such VR/MR Tai Chi games to the home environments of community-dwelling older adults, this high initial exclusion rate reveals key practical challenges. In home settings lacking immediate professional guidance, effective user screening, clear adaptation training guidelines, and easy-to-operate personalized comfort settings (eg, mode switching, visual effect adjustments) will be prerequisites for ensuring user safety and continued use [15,39]. To mitigate cybersickness at home, key strategies include starting with MR or low-vection scenes, using brief (≈ 5 min) sessions, ensuring proper headset fit with stable frame rates, and avoiding artificial locomotion. Furthermore, even mild symptoms, when experienced during unsupervised long-term home use, could affect user experience and adherence, thus

necessitating continuous design optimization to minimize discomfort as much as possible [40].

Accuracy-Competence Paradox

Despite a median movement accuracy of 68.5% (IQR 54.2% - 82.3%), which was below our predefined exploratory benchmark of 80%, participants reported high perceived competence (median 3.8), reflecting a rule set that credits partial hotspot matches ($\geq 1/3$ for ≥ 3 s). This system focuses on older users' positive affective experiences to improve long-term exercise adherence while reducing frustration and elevating confidence but without concomitant increases in stricter movement norms. However, this dilemma positions an additional discussion regarding such a game's effectiveness as a very precise device for training within corresponding trade-offs between motivation and skill fidelity.

The permissive rule (≥ 1 hotspot for ≥ 3 s) retains motivation but tends to overcredit partial matches. Future iterations will include adaptive multipoint matching and more informative corrective cues to put verifiable and perceived competence into alignment while retaining engagement [41]. It presents a foundational tension: the probable discrepancy between perceived benefit/competence and verifiable physiological/skill-based benefit [42]. While perceived competence is central to adherence, if it is not adequately correlated to correct movement execution, long-term benefit to health can be undermined and signal a need for "verifiable competence" feedback loops that remain motivating. For a mass market product like this, selling independent home-based exercise to older people, such a motivational design can serve to buffer against users experiencing "anxiety alone" if having trouble and will invite continued participation [35]. It does raise a valuable question, though: without professional instruction in a home environment setup, how can it be ensured that movements/movement effectiveness is appropriate? Therefore, future iterations of the gameplay might involve implementing tiered success criteria or optional stepwise movement guidance and feedback modules such that users can reach a personally optimal balance between enjoying playing the game and checking movement correctness in accordance with their own needs [43].

Insights From Correlation Analysis of Key Usability Factors

The interplay between gameplay performance, user comfort levels, and subjective experience yields direct design recommendations for successful home-based exergames. It was found that "accuracy" was strongly positively correlated with "flow" and "competence." This indicates that more precise movement performance can guide users into a state of deep, immersive focus, forming a positive feedback loop of "successful execution \rightarrow flow experience \rightarrow enhanced participation." This also offers a more nuanced interpretation of the "accuracy-competence paradox": the game provides a high baseline of perceived competence for all participants through its lenient judgment, and on top of this, higher accuracy leads to an even stronger sense of competence, cleverly balancing universal motivation with individual rewards. This suggests that users do derive satisfaction from objective mastery; while a safety net of easy success is important, completely

removing skill development or clear performance feedback might make the game less engaging long term for some [44,45]. Furthermore, the analysis also highlighted the impact of users' individual states: the degree of physiological discomfort (VRSQ total score) was significantly negatively correlated with "positive affect." This warns that even mild discomfort can negatively affect older users' enjoyment, which is particularly critical in home settings lacking immediate help, thereby re-emphasizing the necessity of minimizing cybersickness risks [46]. Finally, a positive signal was that the study found no significant relationship between "age" and "competence," suggesting that the game possesses good age inclusivity within the 60 to 75 years age group, and users' perceived competence did not systematically diminish with increasing age. This has positive implications for its promotion among diverse community-dwelling older adult groups, although caution is advised when generalizing this conclusion to very old adults [39,47].

Limitations

A number of limitations can be applied to this work. A selection bias in this sample exists since a large number of preliminary participants were lost due to side effects associated with VR, and the resulting group was predominantly female, younger, and physically active. This extremely high level of preexisting physical activity may not be characteristic among this typically sedentary population that such home-based exergames aim to reach but may not serve adequately and hence might limit how much our findings generalize to less active older individuals in this regard. This may inflate physiological comfort but restrict findings' generalizability to a wider range of heterogeneous older people, especially older individuals who may be less technology-savvy or physically constrained. Finally, a single, brief VR intervention cannot eliminate any novelty effects or long-term adherence or long-term effects on health and retention of skill in a naturalistic home environment.

Future Direction

Future research will need to proceed beyond virtual-simulated laboratory settings to home-based settings and use strategies such as Ecological Momentary Assessment to assess long-term use, adherence, skill learning, and actual resultant health effects. It will be important to diversify the samples of participants to include a broader age range, a wider range of health statuses, and a wider range of technology backgrounds to assess the generalizability of such findings. Studies will need to work toward developing adaptive VR systems capable of shifting difficulty and feedback in a dynamic fashion, dependent upon user progress. We will assess fixed versus adaptive thresholds and include per-user calibration routines within a home-based study.

This will aid in connecting, keeping participants motivated, and ensuring appropriate skill acquisition to promote successful unsupervised home-based practice. Although this current system used the Meta Quest 3 headset and hand tracking, future work could include alternative systems (eg, Xbox/Kinect) or outward-facing full-body tracking systems to better assess posture in unsupervised home-based settings. Ecological Momentary Assessment is described as a future longitudinal

assessment methodology to assess adherence and in-the-wild engagement but not a single-session assessment of emotion.

Conclusions

This study demonstrates that well-designed VR/MR Tai Chi exercises can be a realistic and inspiring alternative exercise among community-dwelling older adults, supporting aging in place initiatives. A self-reported profile of high competence but low challenge perception implies early initiation but will need

variable difficulty to maintain motivation in alignment with acuity of therapy for long-term use at home. Although physiologically tolerable in this prescreened population, a high entry exclusion rate due to cybersickness is a major barrier to wider use. A future direction to wider use success is a balance between motivational design of content and therapy acuity and modifications to system usability and comfort to provide net health gains in homeowner unsupervised use.

Acknowledgments

The authors are very grateful to the older residents of Chang'an Community in Changde City, Hunan Province, for participating in the experiment.

Funding

Funding was partly supported by a research grant from Universiti Kebangsaan Malaysia (GUP-2019-066).

Data Availability

The datasets generated and/or analyzed during this study are not publicly available due to participant privacy and confidentiality considerations and institutional ethics restrictions. Deidentified data are available from the corresponding author upon reasonable request for noncommercial research purposes, subject to approval by the relevant ethics committee where applicable and the signing of a data use agreement. The analysis code used in this study is available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: XS, NMA, MHMS, MYR

Methodology: XS

Formal analysis: XS

Investigation: XS

Data curation: XS

Writing – original draft preparation: XS

Writing – review & editing: all authors

Supervision: NMA, MHMS, MYR

Funding acquisition: NMA

Resources: NMA

Conflicts of Interest

None declared.

References

1. Au A, Lai S, Ng E. Introduction: mental health of older adults in Asia. *Clin Gerontol* 2015 May 27;38(3):187-189. [doi: [10.1080/07317115.2015.1008350](https://doi.org/10.1080/07317115.2015.1008350)]
2. Guastafierro E, Rocco I, Quintas R, et al. Identification of determinants of healthy ageing in Italy: results from the national survey IDAGIT. *Ageing Soc* 2022 Aug;42(8):1760-1780. [doi: [10.1017/S0144686X20001671](https://doi.org/10.1017/S0144686X20001671)]
3. Penger S, Grah B. Healthy aging and physical activity: the qualitative research. *ENTRENOVA* 2024;10(1):126-139. [doi: [10.54820/entrenova-2024-0012](https://doi.org/10.54820/entrenova-2024-0012)]
4. Mohd Johari NF, Mohamad Ali N, Mhd Salim MH, Abdullah NA. Factors driving the use of mobile health app: insights from a survey. *mHealth* 2025;11. [doi: [10.21037/mhealth-24-44](https://doi.org/10.21037/mhealth-24-44)] [Medline: [40248757](https://pubmed.ncbi.nlm.nih.gov/40248757/)]
5. Bajdek N, Latham NK, Dishaw M, et al. Feasibility of a multicomponent digital fall prevention exercise intervention for at-risk older adults. *J Frailty Aging* 2024;13(4):349-358. [doi: [10.14283/jfa.2024.35](https://doi.org/10.14283/jfa.2024.35)] [Medline: [39574254](https://pubmed.ncbi.nlm.nih.gov/39574254/)]
6. Huang X, Ali NM, Sahrani S. Evolution and future of serious game technology for older adults. *Information* 2024;15(7):385. [doi: [10.3390/info15070385](https://doi.org/10.3390/info15070385)]
7. Zhang W, Sun J, Feng X, Zhang H, Zhang Y, Zhao M. Effectiveness of Tai Chi exercise on fear of falling and balance in older adults: a meta-analysis. *Geriatr Nurs* 2023;51:194-201. [doi: [10.1016/j.gerinurse.2023.03.019](https://doi.org/10.1016/j.gerinurse.2023.03.019)] [Medline: [37011491](https://pubmed.ncbi.nlm.nih.gov/37011491/)]
8. Boot W, Dilanchian AT. Potential and pitfalls in the use of extended reality to support older adults. *Innov Aging* 2022;6(Suppl 1):216. [doi: [10.1093/geroni/igac059.860](https://doi.org/10.1093/geroni/igac059.860)]

9. Wang PG, Ali NM, Sarker MR. A bibliometric analysis exploring the acceptance of virtual reality among older adults: a review. *Computers* 2024;13(10):262. [doi: [10.3390/computers13100262](https://doi.org/10.3390/computers13100262)]
10. Chen M, Tang Q, Xu S, Leng P, Pan Z. Design and evaluation of an augmented reality-based exergame system to reduce fall risk in the elderly. *Int J Environ Res Public Health* 2020 Oct 1;17(19). [doi: [10.3390/ijerph17197208](https://doi.org/10.3390/ijerph17197208)] [Medline: [33019759](https://pubmed.ncbi.nlm.nih.gov/33019759/)]
11. Huang X, Ali NM, Sahrani S. Haptic-driven serious card games for older adults: user preferences study. *JMIR Serious Games* 2025 May 27;13:e73135. [doi: [10.2196/73135](https://doi.org/10.2196/73135)] [Medline: [40424615](https://pubmed.ncbi.nlm.nih.gov/40424615/)]
12. Dyke G. The logical structure of a research paper in clinical healthcare. *Biosis Biol Syst* 2020;1(4):126-133. [doi: [10.37819/biosis.001.04.0088](https://doi.org/10.37819/biosis.001.04.0088)]
13. Song X, Ali NM, Mhd Salim MH, Rezaldi MY. A literature review of virtual reality exergames for older adults: enhancing physical, cognitive, and social health. *Appl Sci (Basel)* 2025;15(1):351. [doi: [10.3390/app15010351](https://doi.org/10.3390/app15010351)]
14. Benham S, Pennell K, Wyncarczuk K. The meaning of immersive virtual reality from the perspectives of older adults: a qualitative analysis. *Arch Phys Med Rehabil* 2022 Mar;103(3):e17-e18. [doi: [10.1016/j.apmr.2022.01.047](https://doi.org/10.1016/j.apmr.2022.01.047)]
15. Brown JA, Dinh AT, Oh C. Safety and ethical considerations when designing a virtual reality study with older adult participants. 2022 Presented at: Proceedings of the 24th International Conference on Human-Computer Interaction (HCII 2022); Jun 26 to Jul 1, 2022. [doi: [10.1007/978-3-031-05581-2_2](https://doi.org/10.1007/978-3-031-05581-2_2)]
16. Hunicke R, LeBlanc M, Zubek R. MDA: a formal approach to game design and game research. Presented at: Proceedings of the AAAI-04 Workshop on Challenges in Game AI; Jul 25-26, 2004 URL: <https://users.cs.northwestern.edu/~hunicke/MDA.pdf> [accessed 2026-01-10]
17. Fu Y, Hu Y, Sundstedt V, Forsell Y. Conceptual design of an extended reality exercise game for the elderly. *Appl Sci* 2022;12(13):6436. [doi: [10.3390/app12136436](https://doi.org/10.3390/app12136436)]
18. IJsselsteijn WA, Kort YAW, Poels K. The game experience questionnaire. : Technische Universiteit Eindhoven; 2013 URL: https://pure.tue.nl/ws/portalfiles/portal/21666907/Game_Experience_Questionnaire_English.pdf?utm [accessed 2026-01-10]
19. Singh M, Sun D. Evaluating Minecraft as a game-based metaverse platform: exploring gaming experience, social presence, and STEM outcomes. *Interact Learn Environ* 2025;33(7):4272-4294. [doi: [10.1080/10494820.2025.2459200](https://doi.org/10.1080/10494820.2025.2459200)]
20. Fuentes J, Restrepo L, Reyes B, Sánchez R, Figueroa PA. Answering anya: a virtual and mixed reality game centered on nurturing. Presented at: Entertainment Computing - ICEC 2024 - 23rd IFIP TC 14 International Conference; Sep 30 to Oct 3, 2024. [doi: [10.1007/978-3-031-74353-5_28](https://doi.org/10.1007/978-3-031-74353-5_28)]
21. Delgado F, Greenberg J. Cognitive and practice effects of immersive virtual reality use in older adults: preliminary results. *Innov Aging* 2024 Dec 31;8(Suppl 1):1231. [doi: [10.1093/geroni/igae098.3939](https://doi.org/10.1093/geroni/igae098.3939)]
22. Ali NA, Hudson J, Stratton G, Mullins J. O.1.2-7 Optimization of older adults home spaces. *Eur J Public Health* 2023 Sep 11;33(suppl 1). [doi: [10.1093/eurpub/ckad133.094](https://doi.org/10.1093/eurpub/ckad133.094)]
23. Shewaga R, Uribe-Quevedo A, Kapralos B, Alam F. A comparison of seated and room-scale virtual reality in a serious game for epidural preparation. *IEEE Trans Emerg Topics Comput* 2020;8(1):218-232. [doi: [10.1109/TETC.2017.2746085](https://doi.org/10.1109/TETC.2017.2746085)]
24. Chau PH, Kwok YYJ, Chan MKM, et al. Feasibility, acceptability, and efficacy of virtual reality training for older adults and people with disabilities: single-arm pre-post study. *J Med Internet Res* 2021 May 4;23(5):e27640. [doi: [10.2196/27640](https://doi.org/10.2196/27640)] [Medline: [33944795](https://pubmed.ncbi.nlm.nih.gov/33944795/)]
25. Xu TB, Mostafavi A, Boot WR, Czaja S, Kalantari S. Assessing the feasibility and efficacy of virtual reality navigational training for older adults. *Innov Aging* 2024;9(1):igae099. [doi: [10.1093/geroni/igae099](https://doi.org/10.1093/geroni/igae099)] [Medline: [39781245](https://pubmed.ncbi.nlm.nih.gov/39781245/)]
26. Küntzer L, Schwab S, Spaderna H, Rock G. Measuring user experience of older adults during virtual reality exergaming. Presented at: 2024 16th International Conference on Quality of Multimedia Experience (QoMEX); Jun 18-20, 2024; Karlshamn, Sweden p. 153-159. [doi: [10.1109/QoMEX61742.2024.10598263](https://doi.org/10.1109/QoMEX61742.2024.10598263)]
27. Tian F, Ni S, Zhang X, et al. Enhancing Tai Chi training system: towards group-based and hyper-realistic training experiences. *IEEE Trans Vis Comput Graph* 2024 May;30(5):1-11. [doi: [10.1109/TVCG.2024.3372099](https://doi.org/10.1109/TVCG.2024.3372099)] [Medline: [38457324](https://pubmed.ncbi.nlm.nih.gov/38457324/)]
28. Kim HK, Park J, Choi Y, Choe M. Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. *Appl Ergon* 2018 May;69:66-73. [doi: [10.1016/j.apergo.2017.12.016](https://doi.org/10.1016/j.apergo.2017.12.016)] [Medline: [29477332](https://pubmed.ncbi.nlm.nih.gov/29477332/)]
29. Borrohou S, Fissoune R, Badir H. Data cleaning survey and challenges – improving outlier detection algorithm in machine learning. *J Smart Cities Soc* 2023 Oct 9;2(3):125-140. [doi: [10.3233/SCS-230008](https://doi.org/10.3233/SCS-230008)]
30. Ahmed HS. Inferential statistics for cardiothoracic surgeons: part 3 - drawing valid conclusions from clinical data. *Indian J Thorac Cardiovasc Surg* 2025 Feb;41(2):233-247. [doi: [10.1007/s12055-024-01867-7](https://doi.org/10.1007/s12055-024-01867-7)] [Medline: [39822865](https://pubmed.ncbi.nlm.nih.gov/39822865/)]
31. Habibzadeh F. Data distribution: normal or abnormal? *J Korean Med Sci* 2024 Jan 22;39(3):e35. [doi: [10.3346/jkms.2024.39.e35](https://doi.org/10.3346/jkms.2024.39.e35)] [Medline: [38258367](https://pubmed.ncbi.nlm.nih.gov/38258367/)]
32. Boletsis C. A user experience questionnaire for VR locomotion: formulation and preliminary evaluation. 2020 Presented at: Augmented Reality, Virtual Reality, and Computer Graphics, AVR 2020; Sep 7-9, 2020. [doi: [10.1007/978-3-030-58465-8_11](https://doi.org/10.1007/978-3-030-58465-8_11)]
33. Pacheco TBF, de Medeiros CSP, de Oliveira VHB, Vieira ER, de Cavalcanti FAC. Effectiveness of exergames for improving mobility and balance in older adults: a systematic review and meta-analysis. *Syst Rev* 2020 Jul 18;9(1). [doi: [10.1186/s13643-020-01421-7](https://doi.org/10.1186/s13643-020-01421-7)] [Medline: [32682439](https://pubmed.ncbi.nlm.nih.gov/32682439/)]

34. Wong AKC, Bayuo J, Su JJ, et al. Effectiveness of the support from community health workers and health care professionals on the sustained use of wearable monitoring devices among community-dwelling older adults: feasibility randomized controlled trial. *J Med Internet Res* 2024 Nov 18;26:e52435. [doi: [10.2196/52435](https://doi.org/10.2196/52435)] [Medline: [39556810](https://pubmed.ncbi.nlm.nih.gov/39556810/)]
35. Mehrabi S, Muñoz JE, Basharat A, et al. Immersive virtual reality exergames to promote the well-being of community-dwelling older adults: protocol for a mixed methods pilot study. *JMIR Res Protoc* 2022 Jun 13;11(6):e32955. [doi: [10.2196/32955](https://doi.org/10.2196/32955)] [Medline: [35700014](https://pubmed.ncbi.nlm.nih.gov/35700014/)]
36. Martins AC, Quatorze J, Guia D. Exergaming as a tool to enhance strength, balance, gait, mobility, participation, self-efficacy for exercise and adherence in older adults. *Eur J Public Health* 2020 Jun 1;30(Suppl 2):024. [doi: [10.1093/eurpub/ckaa040.024](https://doi.org/10.1093/eurpub/ckaa040.024)]
37. Kim J, Kim Y, Chang PS, Min Oh S, Han S. A pilot study of virtual reality (VR) Tai Chi program on mental health among older adults during the COVID-19 pandemic. *Am J Health Behav* 2022 Oct 17;46(5):576-585. [doi: [10.5993/AJHB.46.5.8](https://doi.org/10.5993/AJHB.46.5.8)] [Medline: [36333829](https://pubmed.ncbi.nlm.nih.gov/36333829/)]
38. Pau M, Arippa F, Leban B, et al. Cybersickness in people with multiple sclerosis exposed to immersive virtual reality. *Bioengineering (Basel)* 2024 Jan 24;11(2):115. [doi: [10.3390/bioengineering11020115](https://doi.org/10.3390/bioengineering11020115)] [Medline: [38391601](https://pubmed.ncbi.nlm.nih.gov/38391601/)]
39. Mehrabi S, Basharat A, Mazen S, et al. Age in place, move in VR: investigating the feasibility and usability of a custom-made virtual reality exergame to promote the well-being of community-dwelling older adults. *Alzheimer's & Dementia* 2023 Dec;19(S23):e071610 [FREE Full text] [doi: [10.1002/alz.071610](https://doi.org/10.1002/alz.071610)]
40. Rmadi H, Maillot P, Artico R, et al. Tolerance of immersive head-mounted virtual reality among older nursing home residents. *Front Public Health* 2023;11:1163484. [doi: [10.3389/fpubh.2023.1163484](https://doi.org/10.3389/fpubh.2023.1163484)] [Medline: [37538272](https://pubmed.ncbi.nlm.nih.gov/37538272/)]
41. Ong DSM, Weibin MZ, Vallabhajosyula R. Serious games as rehabilitation tools in neurological conditions: a comprehensive review. *Technol Health Care* 2021;29(1):15-31. [doi: [10.3233/THC-202333](https://doi.org/10.3233/THC-202333)] [Medline: [32804107](https://pubmed.ncbi.nlm.nih.gov/32804107/)]
42. Tondorf DF, Hounsell M, Pereira VA. Evaluating a serious game for math: is it useful and fun? *JIS* 2025;16(1):199-212. [doi: [10.5753/jis.2025.4656](https://doi.org/10.5753/jis.2025.4656)]
43. Guimarães V, Oliveira E, Carvalho A, et al. An exergame solution for personalized multicomponent training in older adults. *Appl Sci* 2021;11(17):7986. [doi: [10.3390/app11177986](https://doi.org/10.3390/app11177986)]
44. Guzmán DE, Rengifo CF, Guzmán JD, Garcia Cena CE. Virtual reality games for cognitive rehabilitation of older adults: a review of adaptive games, domains and techniques. *Virtual Real* 2024;28(2). [doi: [10.1007/s10055-024-00968-3](https://doi.org/10.1007/s10055-024-00968-3)]
45. Borrego G, Morán AL, Meza V, Orihuela-Espina F, Sucar LE. Key factors that influence the UX of a dual-player game for the cognitive stimulation and motor rehabilitation of older adults. *Univ Access Inf Soc* 2021 Nov;20(4):767-783. [doi: [10.1007/s10209-020-00746-3](https://doi.org/10.1007/s10209-020-00746-3)]
46. Vlahovic S, Skorin-Kapov L, Suznjevic M, Pavlin-Bernardic N. Not just cybersickness: short-term effects of popular VR game mechanics on physical discomfort and reaction time. *Virtual Real* 2024;28(2):1-30. [doi: [10.1007/s10055-024-01007-x](https://doi.org/10.1007/s10055-024-01007-x)]
47. Wang YH. Understanding senior adults' needs, preferences, and experiences of commercial exergames for health: Usability study. *JMIR Serious Games* 2024 Apr 5;12:e36154. [doi: [10.2196/36154](https://doi.org/10.2196/36154)] [Medline: [38578674](https://pubmed.ncbi.nlm.nih.gov/38578674/)]

Abbreviations

- GEQ:** Game Experience Questionnaire
- HMD:** head-mounted display
- MDA:** mechanics, dynamics, and aesthetics
- MR:** mixed reality
- VR:** virtual reality
- VRSQ:** Virtual Reality Sickness Questionnaire

Edited by N Ahmadpour; submitted 23.Jun.2025; peer-reviewed by JAA Bakar, T Baranowski, X Liang; revised version received 08.Sep.2025; accepted 30.Oct.2025; published 28.Jan.2026.

Please cite as:

Song X, Ali NM, Salim MHM, Rezaldi MY
Immersive Tai Chi for Home-Based Exercise in Older Adults: Usability and Feasibility Study
JMIR Serious Games 2026;14:e79453
URL: <https://games.jmir.org/2026/1/e79453>
doi: [10.2196/79453](https://doi.org/10.2196/79453)

© XiaCheng Song, Nazlena Mohamad Ali, Mohamad Hidir Mhd Salim, Muhammad Yudhi Rezaldi. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 28.Jan.2026. 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.

Original Paper

Effects of Game-Based Learning on Piano Music Knowledge Among Elementary School Pupils: Pretest-Posttest Quasi-Experimental Study

Yun Wang^{1,2*}, BA, MA, PhD; Wee Hoe Tan^{1,3*}, BMm, MMM, PhD; Qiaoling Ye^{4*}, BA; Tingting Gu⁵, BA, MA

¹Faculty of Social Sciences and Liberal Arts, UCSI University, Cheras, Kuala Lumpur, Malaysia

²College of Music and Education, Chuzhou University, Chuzhou, Anhui, China

³Chongqing University of Posts and Telecommunications, Chongqing, Chongqing, China

⁴Chengbei Primary School, Chaohu, Anhui, China

⁵Faculty of Teacher Education, Chaohu University, Chaohu, Anhui, China

*these authors contributed equally

Corresponding Author:

Wee Hoe Tan, BMm, MMM, PhD

Faculty of Social Sciences and Liberal Arts

UCSI University

1, Jalan UCSI, UCSI Heights (Taman Connaught)

Cheras, Kuala Lumpur, 56000

Malaysia

Phone: 60 391018880

Fax: 60 391023606

Email: Tanwh@ucsiuniversity.edu.my

Abstract

Background: Music education is central to holistic child development; yet, traditional piano instruction in China often emphasizes rote memorization at the expense of meaningful understanding. Elementary school pupils, particularly novices, frequently struggle with rhythm, melody, and music theory. Game-based learning (GBL), which applies game elements to teaching, may address these challenges by enhancing engagement, providing immediate feedback, and fostering long-term learning.

Objective: This study aimed to examine the effects of a GBL module for piano education on elementary school pupils' music knowledge in Anhui Province, China, compared to traditional instruction.

Methods: A quasi-experimental, nonequivalent control group pretest-posttest design was used. Participants were allocated nonrandomly to experimental and control groups based on scheduling feasibility and teacher availability. A total of 60 novice piano learners (mean age 8.8 years, SD 1.16 years; 16 boys and 44 girls) from 3 elementary schools were assigned to either the GBL group (n=30) or the control group (n=30). Music knowledge was measured using the standardized Level 1 Basic Music Written Test before and after an 8-week intervention. Nonparametric analyses were applied, including Mann-Whitney *U*, Wilcoxon signed-rank, and McNemar item-level analyses. Rank-based effect sizes (*r*) and 95% CIs were reported where applicable.

Results: Baseline differences were present, with the control group scoring higher at pretest (median 52, IQR 24-76) than the GBL group (median 28, IQR 16-64; Mann-Whitney $U=265.50$; $r=-0.35$; 95% CI -0.39 to -0.32 ; $P=.006$). After the intervention, the GBL group significantly outperformed controls (median 100, IQR 88-100 vs median 60, IQR 40-92; Mann-Whitney $U=4.0$; $r=-0.87$; 95% CI -0.90 to -0.83 ; $P<.001$). Within-group analyses confirmed significant pre-post improvements for both groups (control $Z=-3.24$; $r=-0.59$; $P=.001$; and GBL $Z=-4.81$; $r=-0.88$; $P<.001$). Item-level McNemar tests showed significant improvements ($P<.05$) in 5 of 25 items for the GBL group. Missing data were negligible (<2%) and handled via listwise deletion after Little's missing completely at random (MCAR) confirmation ($P=.08$).

Conclusions: The GBL module significantly improved pupils' music knowledge, overcoming baseline disparities and producing posttest score gains with consistent mastery. The innovation of the study lies in the systematic integration of gamification with Orff and Dalcroze pedagogy through the Sidek Module Development Model, which distinguishes it from previous music education studies that examine gamification in isolation. By providing a validated, cost-effective, and scalable instructional module, the study contributes empirical evidence to the field of game-based music education and other practical implications for improving piano instruction in resource-constrained elementary school settings.

KEYWORDS

quality education; gamified learning; piano education; music knowledge; school pupils

Introduction

Music education is a cornerstone of holistic development for elementary school pupils in China, nurturing not only artistic literacy but also cognitive, emotional, and social growth. The piano is highly regarded among musical instruments due to its versatility and extensive skill set. In recent decades, young pupils have taken up piano, a phenomenon known as the “piano heat” [1]. This increase shows a societal commitment to arts education, as piano skills enhance a child’s cultural capital and allow for creative self-expression. The setting of this study, Anhui Province, has seen piano education grow along with national trends and demonstrate regional commitment to early artistic development. Despite the excitement for piano learning, it remains difficult to ensure that pupils develop a profound, meaningful understanding of music beyond mechanical proficiency.

Traditional piano teaching methods that emphasize rote memorization and mechanical practice typically fail to interest and comprehend young learners [2,3]. Due to their cognitive and emotional growth, elementary school pupils require pedagogical approaches that match their capacities. Their short attention spans and abstract nature make them struggle to grasp rhythm, melody, and basic music theory [4-6]. The lack of musical knowledge might also prevent pupils from connecting emotionally with the subject, reducing their motivation and interest [7]. These obstacles highlight the need for innovative teaching strategies that improve music knowledge and engagement.

Contemporary research identifies critical shortcomings in traditional piano beginner teaching, including a reliance on single-minded evaluation models, a neglect of process-based instruction, and a teaching focus that often prioritizes technical results over genuine artistic development. A key challenge remains the comprehensive cultivation of musical literacy, which extends far beyond the accurate reproduction of notes and rhythms [8]. To address these pedagogical gaps, game-based learning (GBL) has emerged as a promising intervention. GBL represents a fusion of informal and formal education, leveraging the interactive nature of games to deliver structured content. Scholarly reviews confirm that interactive mechanics of GBL, such as the use of challenges, rewards, and progress tracking, provide learners with a sense of autonomy and competence [9]. Furthermore, well-designed educational games can deepen a pupil’s understanding of abstract content, encouraging them to find diverse solutions and train their creative and critical thinking skills. These mechanisms align directly with the goal of developing comprehensive musical literacy that transcends simple rote technique [10].

Gamified learning, which uses game design to boost engagement and learning, is a transformative solution. Reimagining piano exercises as engaging, entertaining experiences makes abstract

concepts more approachable and fosters continuous involvement in music instruction [11]. Gamified learning works well in music education with obstacles, rapid feedback, and incentive systems [12]. Gamified learning also fits well with music pedagogy like the Orff and Dalcroze Eurhythmics, which emphasize active participation, rhythm, and movement to meet young learners’ developmental needs [13-16].

In Anhui, elementary school pupils’ piano music knowledge is still lacking despite research showing the benefits of gamified learning in other educational settings and music contexts. Regional specificity is essential for meaningful research since local teaching practices and resource availability may affect gamified intervention efficacy [17]. Gamified music education literature emphasizes short-term outcomes, like skill acquisition, but lacks data on long-term knowledge retention and conceptual understanding [10]. Few studies have examined how gamified learning can be integrated with traditional piano teaching methods or tailored to individual learner differences.

This study examines how a tailored GBL module affects elementary school pupils’ music knowledge in Anhui. Innovative, evidence-based strategies to meet these learners’ unique issues are used to promote music instruction. The design of the GBL module was grounded in established pedagogical principles of Orff-Schulwerk and Dalcroze Eurhythmics, which provided the theoretical foundation for integrating movement, rhythm, and embodied cognition into game tasks. These theories collectively informed the development of a multimodal, developmentally appropriate, and musically grounded GBL intervention for elementary pupils.

The findings should make theoretical and practical contributions to the field. The study expanded on Orff and Dalcroze frameworks to better understand how GBL can promote cognitive, emotional, and creative development in piano teaching. It will provide a validated GBL module for educators to enhance teaching, particularly in resource-constrained or outdated curriculum situations. The findings may also inform policymakers about how gamified learning might improve music education and offer a scalable approach. Accordingly, this study aims to examine the effects of a GBL module on piano music knowledge among elementary school pupils in Anhui province, China, and hypothesizes that pupils receiving the GBL intervention demonstrate greater improvements than those receiving traditional instruction.

Methods

Study Design

This study used a quasi-experimental, nonequivalent control group pretest-posttest design to examine the effects of a GBL module for piano education on elementary school pupils’ music knowledge in Anhui Province, China. This design allowed comparisons between an intervention group receiving gamified piano instruction and a control group receiving traditional

instruction. The report follows the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) statement to enhance transparency and reproducibility [18].

Setting

The study was conducted at a private piano training center located in Anhui Province, China. Three elementary schools partnered with the training center to provide access to novice piano learners. Data collection took place between February and April 2025, with 8 consecutive weekly sessions delivered during this period. Teacher workshops to standardize instructional methods were conducted in January 2025 before the intervention.

Participants

A total of 60 elementary school pupils aged 7 to 12 years were recruited through a series of briefing sessions held at the training center and affiliated schools. Eligibility criteria included (1) no

previous formal piano training, (2) enrollment in the first year of piano lessons at the training center, and (3) parental consent and child assent for participation. Pupils with known hearing impairments or learning disabilities that could substantially affect music learning were excluded. Participants were systematically assigned to an experimental or a control group using a quasi-random sampling procedure: the first 60 pupils whose parents or guardians signed consent forms were enrolled, and every alternate consenting child was assigned to one of the two groups. This resulted in 30 participants in the gamified learning group and 30 in the control group. The final sample comprised 16 boys and 44 girls, consistent with broader trends in higher female participation in music education. [Table 1](#) summarizes demographic characteristics. Although participants were alternately allocated, the procedure did not constitute true randomization. To minimize allocation bias, baseline demographics and pretest scores were compared across groups to assess equivalence.

Table 1. Demographic characteristics of elementary school pupils participating in a quasi-experimental study of game-based piano learning in Anhui, China.

Demographic variable	Experimental group, n	Control group, n	Total, n
Age (years)			60
7	2	4	6
8	6	17	23
9	7	6	13
10	11	3	14
11	3	0	3
12	1	0	1
Sex			60
Male	11	5	16
Female	19	25	44
Year in school			60
Year 1	3	9	12
Year 2	3	12	15
Year 3	13	9	22
Year 4	7	0	7
Year 5	3	0	3
Year 6	1	0	1

Intervention

The experimental group received an 8-week GBL module for piano education designed using the Sidek Module Development Model (SMDM). The module consisted of eight 45-minute sessions incorporating music games, including the “Central C Knocking Game,” “Rhythm Clapping Game,” and “Music Notation Song Game.” Content followed a progressive trajectory from basic note recognition to more complex rhythm and score interpretation, guided by Orff and Dalcroze pedagogical principles. All lessons were delivered on an individual basis, with hands-on keyboard practice embedded within structured teacher-led activities. Attendance was recorded weekly, and

teachers completed session-by-session fidelity checklists to ensure consistent adherence to the intended instructional sequence for both the gamified and traditional teaching conditions.

The control group received traditional piano instruction over the same period. Both groups were taught by the same 5 piano teachers, who attended a 3-day workshop to ensure standardized delivery of both instructional methods. Teachers applied gamified methods for the intervention group and conventional techniques for the control group. Neither participants nor teachers were blinded due to the nature of the educational intervention; however, outcome assessors scoring the written test were blinded to group allocation.

GBL Module Development

Overview

The SMDM, proposed by Sidek and Jamaludin [19], emphasizes systemic structure and the operational feasibility of instructional activities, making it particularly suitable for developing educational modules. In the context of a GBL module for piano education for elementary school pupils, the SMDM iterative and user-centered approach aligns seamlessly with GBL principles, fostering an engaging and effective learning experience tailored to young learners' needs. The structured yet flexible model of SMDM supports the creation of a GBL module for piano education by prioritizing user engagement and iterative refinement. The SMDM was used to create the GBL module to enhance elementary school pupils' piano music knowledge in Anhui [20]. The SMDM is a user-centered instructional design framework that emphasizes comprehensive needs analysis and iterative feedback to create effective multimedia learning tools. The model prioritizes user involvement and adaptability, which are essential for addressing the dynamic challenges in piano education, such as motivation, posture, and skill progression. While the original SMDM framework consists of 2 main stages, the study adapted it into 5 phases (design, development, assembly, validation, and evaluation) to provide a more granular and systematic approach tailored to the development of a GBL module for piano education. This expansion was necessary to accommodate the complexities of integrating technology, gamification, and pedagogical scaffolding in piano learning for elementary school pupils. The needs analysis stage was refined into the design phase for thorough user requirement gathering, while the multimedia development stage was subdivided into development, assembly, and validation phases to ensure iterative refinement and quality assurance. The evaluation phase was added as a distinct step to assess the effectiveness of the module after its implementation, aligning with empirical validation of SMDM. This 5-phase adaptation mirrors the scaffolding process in one-to-one piano lessons, where instruction progresses systematically from foundational to advanced skills, and supports targeted interventions. This systematic approach made the module pedagogically robust, engaging, and customized to young learners. Design, development, assembly, validation, and evaluation built on each other to create a cohesive educational tool.

Design Phase

The main goal of the module was to develop a dynamic GBL approach that elevates piano music knowledge for 7- to 12-year-olds. The aim was to craft a challenging but fun learning experience that deepened musical understanding. A game-based piano learning experience for elementary school pupils should be engaging and appealing [21]. The module emphasizes engaging experiences, clear goals, immediate feedback, diverse gaming techniques, and adaptive difficulty levels. Games that reinforce piano music knowledge use these principles. A comprehensive needs analysis identified learning challenges and ensured the module addressed them through targeted gamification. The 4Keys2Fun framework [22] was used to keep pupils engaged by incorporating social interaction, challenge,

exploration, and purposeful learning to improve musical knowledge retention.

Development Phase

Content and instructional strategies were carefully selected during development to support a progressive learning trajectory. Music was used to teach specific knowledge outcomes, starting with note recognition and rhythm in "Please Play" and advancing to complex structures in "Church Organ" and "Yankee Doodle." Teacher demonstration, GBL, guided practice, rhythmic chanting, and formative feedback were used to teach piano. "Rhythm Clapping Game" (pupils competed to replicate rhythmic patterns in pairs, turning practice into a lively and collaborative challenge that reinforced knowledge while fostering peer learning), "Musical Notation Song Game," and "Central C Knocking Game" were embedded to improve rhythm, note positions, and key identification. A piano and copies of John Thomson's Easiest Piano Course 1 were needed for 45-minute sessions, ensuring the module is accessible. The teacher demonstrated real-time modeling to explain music.

Assembly and Draft Completion

A backward design model prioritized piano music knowledge outcomes before aligning assessments, content, and activities [23]. This concentrated and unified the framework. A total of 8 successive lessons with objectives, assessment tools, and resources guided pupils from basic note recognition to sophisticated musical interpretation in the final edition. This systematic assembly ensured that every part improved musical comprehension.

Validation and Pilot Testing

Validation was essential for module reliability and suitability. Three music pedagogy specialists critically assessed its content and design, followed by semistructured interviews with 2 piano teachers. Their suggestions improved content alignment and game dynamics. After expert approval, a pilot test with 17 students over 8 weeks (November 19, 2024, to January 14, 2025) provided practical input and revisions to ensure the readiness of the module for broader use. The 17 participants were novice learners aged 7 to 11 (mean age 8.8, SD 1.16) years, with 11 girls and 6 boys, mostly in Year 2 and Year 3 of elementary school.

Piano Education GBL Module

This basic school package has 8 lessons that establish musical principles through individual works. It uses games like "Central C Knocking Game," "Five-Line Speech Practice," and "Right-Hand Shooting Game" (students aimed to strike correct notes under time pressure, and their excitement translated into improved recognition and accuracy) to enhance learning. The curriculum systematically teaches piano key identification, note values, rests, and specific notes (C, D, B, E, A, G, and F). It emphasizes practical exercises to improve finger independence, rhythmic precision, and technical fluency on the piano. It also promotes curiosity, self-confidence, and music appreciation through interactive games and collaborative activities. Rhythm accuracy, hand coordination, and note and rest recognition are used to assess student progress, ensuring evidence-based musical progression. This structured technique improves musical

proficiency and makes learning fun, supporting game-based music education.

Assessments and Outcomes

The primary outcome was music knowledge, assessed using the Level 1 Basic Music Written Test developed by the Chinese Music Academy. This 25-item multiple-choice test evaluates aural recognition (tones, melodic contours, dynamics, tempo, and rhythmic structures) and theoretical knowledge (pitch, solfège, intervals, meter, and musical symbols). The test was administered at baseline (pretest) and at the end of the intervention (posttest). Exposures included the gamified learning intervention, while potential confounders such as age, grade level, and baseline knowledge were recorded.

Study Size

The study included 60 participants (30 per group). The sample size was determined pragmatically based on the capacity of the training center and the number of eligible consenting pupils during the recruitment window. Although no formal power calculation was conducted, a sample of 30 per group has been shown in previous quasi-experimental studies to detect medium-to-large effects in educational outcomes.

Recruitment Procedures

Recruitment was conducted in collaboration with school administrators and the piano training center. Parents and guardians were provided with written study information, and informed consent was obtained prior to enrollment. Children also provided verbal assent. Recruitment was based on voluntary self-selection following briefing sessions conducted at the partner schools and training center. All eligible pupils whose parents provided consent during the 2-week recruitment window were included, and no incentives were offered.

A total of 86 pupils were screened for eligibility; 60 met the inclusion criteria and consented to participate. All 60 pupils were allocated to one of the two study conditions (30 experimental and 30 control). All participants received the assigned intervention, completed the 8 sessions, and completed both pretest and posttest assessments. No participants were lost to follow-up, and no protocol deviations occurred.

Data Analysis

All analyses were conducted using SPSS (version 29.0; IBM Corp). The normality of the pretest was assessed using the

Shapiro-Wilk test, which indicated nonnormal distributions; hence, nonparametric statistical tests were chosen. Mann-Whitney U tests were carried out to compare the control and experimental groups in every pretest and posttest. Rank-based effect sizes (r) and 95% CIs were measured where applicable. Wilcoxon signed-rank tests examined changes from pretest to posttest within each group. McNemar tests assessed changes in correct responses for individual items on the 25-item test [24]. Data completeness was assessed, in which only less than 2% of responses were missing, and Little's missing completely at random (MCAR) test indicated data were missing completely at random ($P=.078$). As the proportion was small, listwise deletion was applied. No imputations were performed. Interpretations emphasized both statistical significance and precision of effect estimates, with CI reporting throughout.

Ethical Considerations

This study was approved by the Institutional Ethics Committee, UCSI University, under approval number IEC-2024-FOSSLA-0167, dated November 19, 2024. Written informed consent was obtained from all legal guardians, and verbal assent was provided by children before participation.

Confidentiality was ensured by anonymizing all datasets and storing them securely with restricted access. No personal identifiers were included in analyses or publications. Participants were not given financial incentives, but small educational gifts (eg, notebooks and stationery) were provided as tokens of appreciation. No identifiable photographs were taken; all supplementary materials were free of personal identifiers. Where identifiable images were unavoidable, written consent was obtained from both pupils and guardians.

Results

The pretest results showed a baseline piano music knowledge difference between groups (Table 2). The control and experimental groups had considerable piano music knowledge gaps before the intervention. The control group recorded a median pretest score of 52 (range 24-76) and a mean of 47.73 (SD 14.59), while the experimental group had 28 (range 16-64) and 36.13 (SD 12.92). In Table 3, a Mann-Whitney U test revealed a statistically significant difference (Mann-Whitney $U=265.50$; $P=.006$), indicating that the control group began with a higher level of knowledge.

Table 2. Pretest and posttest piano music knowledge scores among elementary school pupils receiving an 8-week game-based learning module versus traditional piano instruction in Anhui, China.

	Pretest scores			Posttest scores		
	Control ^a	Experimental ^b	Both	Control ^a	Experimental ^b	Both
n	30	30	60	30	30	60
Median (IQR)	52 (32-60)	32 (28-45)	40 (28-56) ^c	60 (51-65)	100 (95-100)	90 (60-100) ^d
Mode	60	28	28	60	100	100
Mean (SD)	47.73 (14.59)	36.13 (12.92)	41.93 (14.86)	58.67 (11.71)	97.33 (3.68)	78 (21.31)
Minimum	24	16	16	40	88	40
Maximum	76	64	76	92	100	100

^aControl group (n=30).^bExperimental group (n=30).^c $P=.006$.^d $P<.001$.**Table 3.** Between-group comparison of piano music knowledge at pretest and posttest using Mann-Whitney U tests in a quasi-experimental study of elementary school pupils in Anhui, China.

	Mann-Whitney <i>U</i>	<i>Z</i>	<i>r</i> (95% CI)	<i>P</i> value
Pretest	265.50	-2.75	0.35 (-0.39 to -0.32)	.006
Posttest	4.00	-6.71	0.87 (-0.90 to -0.83)	<.001

To evaluate the effectiveness of an 8-week GBL intervention on piano music knowledge, Mann-Whitney *U* tests were conducted to compare pretest and posttest scores between the experimental group (n=30) and the control group (n=30). Descriptive statistics and test results are presented in [Table 3](#). No adverse events or unintended effects were observed in either study condition throughout the 8-week intervention.

In the pretest, a significant difference was observed between the experimental and control groups (Mann-Whitney $U=265.50$; $Z=-2.75$; $P=.006$), with the experimental group demonstrating higher scores compared to the control group. The effect size was medium ($r=0.35$, 95% CI -0.39 to -0.32), indicating a moderate baseline advantage for the experimental group.

In the posttest, the experimental group significantly outperformed the control group (Mann-Whitney $U=4.00$; $Z=-6.71$; $P<.001$). The experimental group achieved a median score of 100 (range 88-100; mean 97.33, SD 3.69), while the control group had a median score of 60 (range 40-92; mean

58.67, SD 11.71). The effect size was large ($r=0.87$; 95% CI -0.90 to -0.83), reflecting a substantial improvement in the performance of the experimental group. These results suggest that the 8-week GBL module effectively reversed the baseline performance patterns, closing the initial knowledge gap and significantly enhancing piano music knowledge in the experimental group compared to the control group.

The posttest scores achieved by the experimental group indicate a high degree of knowledge mastery, with reduced variability, in contrast to the more varied outcomes observed in the control group. The decreased SD (SD 3.69) in the experimental group suggests more consistent learning gains, reinforcing the ability of the GBL module to achieve uniform learning outcomes. [Table 4](#) shows pretest and posttest score disparities demonstrating the effect of the intervention. The experimental group experienced a median gain of 64 points (range 36-84) with a mean improvement of 61.20 (SD 12.61). The pronounced disparity in knowledge gains shows that the GBL module promotes significant learning progress.

Table 4. Differences in piano music knowledge scores between pretest and posttest among elementary school pupils receiving game-based versus traditional piano instruction in Anhui, China

	Score difference between pretest and posttest		
	Control group	Experimental group	Control group and experimental group
n	30	30	60
Median (IQR)	8 (4-16)	64 (55-68)	36 (8-64)
Mode	4	64	64
Mean (SD)	10.93 (11.64)	61.20 (12.61)	36.07 (28.05)
Minimum	-4	36	-4
Maximum	64	84	84

To assess the effect of an 8-week GBL intervention on piano music knowledge, the Wilcoxon Signed Ranks test was conducted to compare pretest and posttest scores within the

experimental group (n=30) and the control group (n=30). Results are presented in [Table 5](#).

Table 5. Within-group comparison of pretest and posttest piano music knowledge using Wilcoxon signed-rank tests among elementary school pupils in Anhui, China.

Group	Z	r (95% CI)	P value
Control group	-3.24	-0.59 (-0.66 to -0.53)	.001
Experimental group	-4.81	-0.88 (-0.94 to -0.82)	<.001

Both groups demonstrated significant improvements in piano music knowledge from pretest to posttest. For the control group ($Z=-3.24$; $P=.001$), a substantial increase was observed ($Z=-3.24$; $P=.001$), with a large effect size ($r=-0.59$, 95% CI -0.66 to -0.53). The experimental group exhibited a more robust improvement ($Z=-4.81$; $P<.001$), with a larger effect size ($r=-0.88$, 95% CI -0.94 to -0.81). The higher Z value and lower P value in the experimental group with a larger effect size indicate that the GBL intervention led to greater knowledge growth compared to the control group. These findings, combined with the between-group comparison (see [Table 3](#)), suggest that the GBL module significantly enhanced learning outcomes, effectively closing the baseline knowledge gap in the experimental group.

McNemar tests were conducted on pretest and posttest scores for 25 test items to assess significant changes in performance. The item-level McNemar tests were exploratory ancillary analyses intended to examine differential item difficulty and the distributional impact of the intervention across 25 outcomes.

For the experimental group, McNemar tests were applied to each of the 25 piano knowledge items ([Table 6](#)). The analysis revealed statistically significant improvements ($P<.05$) in 5 items, highlighting the targeted effectiveness of the GBL module [25]. The remaining 20 items showed no statistically significant changes ($P>.05$).

Table 6. Item-level McNemar analyses of changes in piano music knowledge test items following an 8-week game-based learning (GBL) intervention among elementary school pupils in Anhui, China.

	Test items	Counts		Post ^a		Mode	P value
				W ^b	C ^c		
1	Listen and discern which sound is higher.	Pre ^d	W	13	3	W-W	.04
		— ^e	C	12	2		
2	Listen and discern which sound has a shorter duration.	Pre	W	11	6	W-W	≥.99
		—	C	7	6		
3	Listen and discern, choose the correct pitch.	Pre	W	16	2	W-W	.11
		—	C	8	4		
4	Listen and discern which one is the Concord interval.	Pre	W	21	0	W-W	.008
		—	C	8	1		
5	Comparing melody intervals, which interval is farther away?	Pre	W	10	2	C-W	.001
		—	C	16	2		
6	Listen and discern the direction of melody.	Pre	W	10	4	W-W	.27
		—	C	9	7		
7	Listen and discern the melody, choose the correct beat number.	Pre	W	5	6	C-C	.79
		—	C	8	11		
8	Listen and discern, choose the rhythm within the square.	Pre	W	16	0	W-W	N/A ^f
		—	C	14	0		
9	Listen and discern the melody, choose the correct note within the square.	Pre	W	18	3	W-W	.73
		—	C	5	4		
10	Listen to the melody. What is the style of this piece of music?	Pre	W	9	9	W-W; W-C	.80
		—	C	7	5		
11	Which note has a lower vocal range?	Pre	W	11	8	W-W	.79
		—	C	6	5		
12	Which one is a harmony interval?					W-W	.27

Test items	Counts		Post ^a		Mode	P value	
			W ^b	C ^c			
			Pre	W			
13	Choose the correct singing name.	Pre	W	13	4	W-C; C-C	N/A
		—	C	9	4		
14	Choose the correct interval degree.	Pre	W	0	15	C-W; C-C	.09
		—	C	0	15		
15	Which one is the dissonant interval?	Pre	W	7	3	C-C	.23
		—	C	10	10		
16	Which one is the sixteenth rest?	Pre	W	9	3	W-C; C-C	N/A
		—	C	8	10		
17	Choose the correct symbol meaning.	Pre	W	0	15	W-C	N/A
		—	C	0	15		
18	According to the spectrum example, choose the correct beat number.	Pre	W	0	16	W-C	<.001
		—	C	0	14		
19	According to the beat sign, how many bars of the score are shared?	Pre	W	1	21	W-C	.01
		—	C	1	7		
20	Choose the correct sound name and complete the C natural major scale.	Pre	W	5	12	C-C	.11
		—	C	2	11		
21	The instrument in the picture is _____.	Pre	W	3	8	C-C	N/A
		—	C	2	17		
22	The instrument in the picture is _____.	Pre	W	0	12	C-C	N/A
		—	C	0	18		
23	The instruments belonging to the woodwind group are_____.	Pre	W	0	12	W-C; C-C	N/A
		—	C	0	18		
24	Representatives of the “classical music” period.	Pre	W	0	15	W-C	N/A
		—	C	0	15		
		Pre	W	0	22		
		—	C	0	8		

Test items	Counts		Post ^a		Mode	P value
			W ^b	C ^c		
			Pre	W		
25 “Sunset Flute and Drum” belongs to _____.	—	C	0	18	C-C	NA
	—	—	0	18		

^aPost: posttest.

^bW: wrong.

^cC: correct.

^dPre: pretest.

^eNot applicable.

^fN/A: not applicable.

For the control group, McNemar tests were similarly conducted on the 25 items. Five items (#1, #4, #5, #18, and #19) showed statistically significant improvements ($P < .05$), though the gains were less pronounced than in the experimental group. Both groups improved significantly in aural recognition (#1, #4, and #5) and rhythmic analysis (#18 and #19). However, the experimental group showed greater gains in complex tasks like beat identification (#18) and score interpretation (#19), suggesting the GBL module excels at engaging pupils in challenging musical concepts. The improvements in the control group were more stable in simpler tasks (eg, #1 and #4), indicating retention rather than new learning, consistent with the limited scope of traditional approaches. The item-level analysis underscores that the GBL module improves piano music knowledge, particularly in rhythm and score analysis, where the experimental group outperformed the control group. Both groups mastered several things, but the larger and more substantial gains of the experimental group show that the GBL module can enhance music education outcomes.

Discussion

Summary of the Main Study

This study examined whether a GBL module for piano education could improve music knowledge among elementary school pupils in Anhui, China, compared with traditional piano instruction. The findings demonstrated that pupils in the gamified learning group significantly outperformed their peers in the control group after the 8-week intervention, achieving near-perfect posttest scores with reduced variability. This suggests that gamification not only facilitated mastery of theoretical and aural skills but also promoted more consistent learning outcomes across participants. The between-group Mann-Whitney U test showed a large effect size, while the within-group Wilcoxon analysis confirmed significant pre-post improvements in both groups, with stronger gains in the gamified condition. These results validate the hypothesis of the study that gamified learning would yield superior knowledge acquisition compared to conventional methods, addressing the gap identified in the Introduction. In summary, the evidence confirms that gamification is an effective and scalable pedagogical approach for strengthening music knowledge in early piano education.

A tailored GBL module improved elementary school pupils' piano music knowledge in Anhui, compared to the control group. Postintervention, GBL module students outperformed the control group by a significant margin. Gamified approaches improve music knowledge, supporting previous research on GBL for music education engagement [11,12]. The near-perfect scores of the experimental group and low variability indicate that the module boosted knowledge acquisition and ensured consistent mastery, a challenge in traditional piano pedagogy.

A comprehensive needs analysis indicated that Anhui elementary pupils struggle with motivation and foundational skills when learning piano. A dynamic, entertaining GBL approach was used to improve piano knowledge, skills, and attitudes. The module promoted musical competency and positive learning habits through gamification principles, including fun, clear objectives, immediate feedback, and adaptability [26]. Addressing these challenges created an engaging experience targeted to young learners' interests, establishing the groundwork for the research findings.

Starting with “Please Play” to build foundational skills and on to “Church Organ” to refine expressive techniques, the module material was progressive. Structured repetition improved cognitive and motor skills in this sequencing. The instructional design integrated traditional piano pedagogy with gamified activities like the “Right-Hand Shooting Game” and “Five-Line Speech Practice” to improve retention and motivation [27]. These entertaining activities improved rhythm, note recognition, and expression, directly benefiting the understanding of the experimental group.

The lesson design required only a piano and basic multimedia tools, ensuring accessibility across schools in Anhui. Teacher-led demonstrations emphasized real-time feedback and interpersonal connection, which young learners need, above digital media. Tailored feedback allowed pupils to correct errors and refine performance immediately [4], boosting motivation and addressing individual needs [27]. While technology provided aural and visual signals, research suggests that interleaving teacher demonstrations with student imitation is better than audio-only or blocked observation [28]. Thus, the module balanced traditional and technological approaches in a practical, learner-centered environment [28].

Comparisons to Existing Literature

Quasi-experimental assessments confirmed the efficacy of the module. The pupils improved their piano knowledge and skills from a poor baseline to mastery by the end of the study, outperforming peers in traditional settings. Item-level analysis demonstrated targeted gains in aural recognition (pitch differentiation) and rhythmic skills (beat identification), aligning with the gamified focus [29]. Game-based elements, including challenges, real-time feedback, and rewards, increased the median score of the experimental group much more than the control group. Similar to Molloy et al [21], gamified piano instruction enhanced note accuracy and technical performance compared to traditional techniques, possibly due to greater engagement and skill mastery. Fadhli et al [30] found in a meta-analysis that gamified instruction improves children's musical knowledge and technical proficiency. These findings support the idea that GBL approaches can significantly improve musical learning outcomes.

Both groups started with similar baselines, but the experimental group improved by posttest. Traditional teaching methods yielded marginal, inconsistent gains in the control group. GBL improves engagement and retention, according to Robert et al [31] and Qian and Jiang [32], who linked it to piano performance. The interactive design of the module addressed the focus of traditional pedagogy on mechanical execution over holistic understanding, resulting in limited development in the control group [33,34].

The success of the experimental group reflects the affective and cognitive benefits of GBL. Gamification increases motivation and self-efficacy, while narrative-driven GBL deepens emotional connections to information [35]. These mechanisms likely drove the experimental group to engage and master. Molero et al [36] observed that gamified systems simplify challenging concepts, which promotes integration of Orff and Dalcroze approaches [15,37]. The focus of the module on aural and rhythmic skills addressed music education gaps, where an outdated curriculum hinders creativity [32]. Beyond immediate outcomes, the performance of the experimental group in analytical tasks (eg, interval comparison) implies GBL fosters higher-order thinking and skill transferability, areas understudied [17,38].

Implementation fidelity was generally high, as teachers followed the structured lesson plans and completed weekly fidelity checklists. However, minor barriers were observed, particularly in maintaining pupil focus during transitions from game activities to reflective discussions. These challenges suggest that future iterations of the module should incorporate more structured debriefing routines to support smoother transitions and deepen conceptual understanding. This study builds on the Orff approach, Dalcroze Eurhythmics, and the SMDM to better understand how gamified learning improves music knowledge in elementary school pupils. The interactive, rhythm-focused exercises of the game-based module increased engagement and musical comprehension, following the Orff approach [13]. The kinesthetic features of the module follow Dalcroze Eurhythmics, which integrates movement and music to improve rhythmic and expressive skills [14], improving pupils' competency and sensitivity.

Practically, the validated game-based module offers Anhui educators a scalable, resource-efficient tool to modernize piano teaching in often underresourced settings. The experimental group outperformed the control group significantly in knowledge and skills underscores the efficacy of integrating game elements, such as points, levels, and immediate feedback, into instruction [4]. By fostering enjoyment and reducing anxiety, the module also sustains long-term engagement, addressing dropout risks and challenging traditional methods' focus on technical drills over emotional connection [39,40]. Curriculum designers should thus prioritize holistic, learner-centered approaches.

Limitations

Despite the promising outcomes of this study, several limitations warrant consideration to contextualize the findings and guide future research. First, the quasi-experimental design used a nonequivalent control group without true random assignment, which may have introduced selection bias, as evidenced by initial differences in pretest scores [24]. Future studies should use true random assignment to enhance internal validity and ensure baseline equivalence between groups. Additionally, the sample size of 60 participants, while sufficient for detecting significant effects, limits the generalizability of findings. The study focused on novice piano learners in a private training center in Anhui, which may not fully represent the diverse educational contexts across China or other regions. Expanding the sample size and including public school settings could strengthen the applicability of the GBL module.

The 8-week intervention period provided valuable insights into short-term knowledge gains but was insufficient to assess long-term retention or the sustained impact of GBL on musical development. Previous research highlights the need for longitudinal data to evaluate deep conceptual understanding and skill retention [10]. Future studies should extend the intervention duration and incorporate follow-up assessments to examine whether the observed gains persist over time. While gamification enhanced engagement, teachers occasionally reported difficulty redirecting pupils' attention from gameplay to reflective discussion. To mitigate this, future iterations may include structured debrief sessions after each activity.

Another limitation lies in the assessment tool, the Level 1 Basic Music Written Test, which exhibited ceiling effects in the experimental group through their posttest scores. These effects suggest that the test may not have been sufficiently challenging to fully capture knowledge gains, potentially underestimating the impact of the intervention. Moreover, the study primarily measures aural and theoretical knowledge. This focus may overlook other critical aspects of piano learning, such as technical proficiency, emotional expression, or creativity, which are integral to holistic music education [7]. Incorporating multifaceted assessment tools, such as performance evaluations or qualitative measures of student engagement, would provide a more comprehensive understanding of the impact of GBL. Future studies may supplement written tests with recorded piano performances assessed using standardized rubrics, teacher observations, and student portfolios, enabling a more holistic evaluation of musical skill development.

The study did not account for individual learner differences, such as cognitive abilities, learning styles, or previous musical exposure, which may influence the efficacy of GBL [41,42]. The GBL module's one-size-fits-all approach, while effective overall, may not equally benefit all pupils. For instance, item-level analysis revealed variability in performance on complex tasks like interval comparison, suggesting potential confusion for some learners. Future research should explore adaptive GBL designs that tailor content to individual needs. Additionally, the reliance on teacher-led instruction, while resource-efficient, may have introduced variability in implementation fidelity. Although teachers underwent training, differences in their delivery styles could have influenced outcomes.

Finally, this study focuses on a specific context of Anhui, while a strength for regional relevance, but limits its generalizability to other cultural or educational settings. Local teaching practices and resource constraints in Anhui shaped the design of the module, but these may differ elsewhere. Comparative studies across diverse regions could elucidate how contextual factors influence the effectiveness of the GBL module. Furthermore, its integration with traditional pedagogy like Orff and Dalcroze was effective, but its compatibility with other music education frameworks remains underexplored. Investigating hybrid models that combine GBL with varied pedagogical approaches could broaden its applicability.

To address these limitations, researchers should prioritize longitudinal studies with larger, more diverse samples to validate and extend the findings. Randomizing group assignments and incorporating varied assessment methods will enhance the robustness of future investigations. Educators in Anhui and similar regions are encouraged to adopt the validated GBL module, adapting it to local needs while ensuring teacher training emphasizes consistent implementation. Policymakers should consider investing in scalable GBL frameworks, potentially integrating digital tools to enhance accessibility and personalization in resource-constrained settings. By addressing these limitations and implementing these recommendations, the field can further harness the potential of GBL to transform piano education, fostering both musical proficiency and lifelong engagement among young learners.

Conclusion

The design and validation of the module underscore key implications for music education. Gamification enhanced engagement and knowledge acquisition, suggesting its potential for broader application. Multimodal instruction catered to diverse learning styles, while collaborative practices extended learning beyond the classroom. In conclusion, this GBL module provides a robust, validated approach to early music education

for Anhui pupils. By integrating structured content, playful activities, and reflective teaching, it strengthens musical competencies and fosters a positive learning ecosystem, contributing to the evidence base for innovative pedagogy.

This study highlights that integrating gamification into piano pedagogy is both innovative and impactful for early music education. By embedding elements of play, challenge, feedback, and progression into traditional piano lessons, the GBL module transformed abstract musical concepts into engaging learning experiences that fostered consistent mastery among elementary school pupils. This approach is particularly valuable in resource-constrained environments, as the module was designed to be cost-effective, requiring minimal technology beyond a piano and simple teaching aids, while still producing large and meaningful learning gains. The innovation lies in bridging established pedagogical traditions, such as Orff and Dalcroze, with systematic gamification through the SMDM, thereby creating a hybrid framework that supports motivation, knowledge retention, and skill transfer. While traditional methods often emphasize technical drills at the expense of holistic understanding, this study shows that gamification sustains engagement and cultivates deeper conceptual knowledge. Practically, the validated GBL module can be adopted widely by music educators and curriculum designers to modernize piano teaching, reduce dropout risks, and increase accessibility of music learning across diverse school settings. Ultimately, the study demonstrates that gamified learning is not only a pedagogical enhancement but also a scalable and sustainable solution for advancing children's cultural capital, creativity, and long-term musical engagement.

In conclusion, this study demonstrates that integrating gamification into elementary piano education represents both an innovative and effective pedagogical advancement. The study differs from existing research by systematically embedding GBL within established music education frameworks, namely Orff-Schulwerk and Dalcroze Eurhythmics, and operationalizing this integration through the SMDM. This approach moves beyond previous studies that primarily focus on engagement or short-term skill gains by providing robust empirical evidence of consistent knowledge mastery and reduced performance variability among young learners. The contribution of this study lies in offering a validated, theoretically grounded, and developmentally appropriate GBL module that advances understanding of how gamification can support foundational music knowledge acquisition. From a real-world perspective, the module provides educators and curriculum designers with a scalable, low-cost, and practical solution for modernizing piano instruction, particularly in underresourced educational contexts, while fostering sustained engagement, reducing learner attrition, and strengthening long-term musical development.

Data Availability

The data generated and analyzed during this study are not publicly available due to institutional and participant privacy considerations, but are available from the corresponding author upon reasonable request. Deidentified datasets, analysis code, and supplementary materials can be provided to qualified researchers for academic and noncommercial purposes.

Conflicts of Interest

None declared.

References

1. Bai B. Piano learning in the context of schooling during China's 'Piano craze' and beyond: motivations and pathways. *Music Education Research* 2021;23(4):512-526. [doi: [10.1080/14613808.2021.1929139](https://doi.org/10.1080/14613808.2021.1929139)]
2. Liu J. Compare the effect of different teaching methods in piano education. *HBEM* 2024;43:661-666. [doi: [10.54097/nt1nbh54](https://doi.org/10.54097/nt1nbh54)]
3. Lin C, Liu EZ, Chen Y. Game-based remedial instruction in mastery learning for upper-primary school students. *J Educ Tech Soc* 2013;271-281 [FREE Full text]
4. Shi M. The exploration and application of children's music education in piano teaching. *artpl* 2023;4(7). [doi: [10.23977/artpl.2023.040707](https://doi.org/10.23977/artpl.2023.040707)]
5. Sun Y. Comparative analysis of methods for teaching piano in various cultural traditions: from the classical European school to modern American approaches. *Educ Manag Rev* 2023;13 (12-1):191-197. [doi: [10.25726/j2371-4388-1702-x](https://doi.org/10.25726/j2371-4388-1702-x)]
6. Garn AC, Jolly JL. High ability students' voice on learning motivation. *J Advanced Academics* 2013;25(1):7-24. [doi: [10.1177/1932202x13513262](https://doi.org/10.1177/1932202x13513262)]
7. Reinsalu H, Timostuk I, Ruokonen I. Literature review about the learning engagement in preschool and primary school music education. *PMP* 2022;21(2):35-69. [doi: [10.59893/pmp.21\(2\).004](https://doi.org/10.59893/pmp.21(2).004)]
8. Yang L. Analysis of the current situation and improvement strategies for piano beginner teaching methods. *Int J Educ Humanities* 2025;20(3):9-12. [doi: [10.54097/rbn4f060](https://doi.org/10.54097/rbn4f060)]
9. Alotaibi MS. Game-based learning in early childhood education: a systematic review and meta-analysis. *Front Psychol* 2024;15:1307881 [FREE Full text] [doi: [10.3389/fpsyg.2024.1307881](https://doi.org/10.3389/fpsyg.2024.1307881)] [Medline: [38629045](https://pubmed.ncbi.nlm.nih.gov/38629045/)]
10. Weatherly KICH, Wright B, Lee EY. Digital game-based learning in music education: a systematic review between 2011 and 2023. *Res Studies Music Educ* 2024. [doi: [10.1177/1321103x241270819](https://doi.org/10.1177/1321103x241270819)]
11. Liu Y. The application of music games in elementary school classes. *Int J Educ Humanities* 2023;9(2):205-208. [doi: [10.54097/ijeh.v9i2.9934](https://doi.org/10.54097/ijeh.v9i2.9934)]
12. Phunsa S, Pawala T. Learning through gameplay: creating memorable learning experiences with games. *J Educ Learn Rev* 2024;1(1):39-46. [doi: [10.60027/jelr.2024.678](https://doi.org/10.60027/jelr.2024.678)]
13. Orff C. The Schulwerk: its origin and aims. *Music Educ J* 1963;49(5):69-74. [doi: [10.2307/3389951](https://doi.org/10.2307/3389951)]
14. Jaques-Dalcroze, Rubinstein H. *Rhythm, Music and Education*. London: Chatto & Windus; 1921.
15. Beegle A. Playing with Orff and classroom instruments. In: Abril CR, Gault BM, editors. *General Music: Dimensions of Practice*. New York, NY: Oxford University Press; 2022.
16. Daly DK. Creativity, autonomy and Dalcroze eurhythmics: an arts practice exploration. *Int J Music Educ* 2021;40(1):105-117. [doi: [10.1177/02557614211028600](https://doi.org/10.1177/02557614211028600)]
17. Liu L. Research on piano education in contemporary china. *J Contemp Educ Res* 2021;5(2). [doi: [10.26689/jcer.v5i2.1836](https://doi.org/10.26689/jcer.v5i2.1836)]
18. Des Jarlais DC, Lyles C, Crepaz N, TREND Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004;94(3):361-366. [doi: [10.2105/ajph.94.3.361](https://doi.org/10.2105/ajph.94.3.361)] [Medline: [14998794](https://pubmed.ncbi.nlm.nih.gov/14998794/)]
19. Sidek MN, Jamaludin A. *Pembinaan modul: bagaimana membina modul latihan dan modul akademik*. Serdang: Penerbit Universiti; 2005.
20. Madihie A, Noah SM. An application of the Sidek module development in rebt counseling intervention module design for orphans. *Procedia - Social and Behavioral Sciences* 2013;84:1481-1491. [doi: [10.1016/j.sbspro.2013.06.777](https://doi.org/10.1016/j.sbspro.2013.06.777)]
21. Molloy W, Huang E, Wünsche B. Mixed reality piano tutor: a gamified piano practice environment. 2019 Presented at: Proceedings of the International Conference on Electronics, Information, and Communication (ICEIC); May 6, 2019; Auckland, New Zealand p. 22-25. [doi: [10.23919/elinfocom.2019.8706474](https://doi.org/10.23919/elinfocom.2019.8706474)]
22. Lazzaro N. The 4 keys 2 fun. XEODesign. URL: <http://www.xeodesign.com/research/> [accessed 2025-06-28]
23. Mohammed J, Schmidt K, Williams J. Designing a new course using backward design. 2022 Presented at: Proceedings of the ASEE Annual Conference & Exposition; June 26, 2022; Minneapolis, USA p. 26-29. [doi: [10.18260/1-2--41806](https://doi.org/10.18260/1-2--41806)]
24. Hinton PR. Mann-Whitney U test. In: Salkind NJ, editor. *The SAGE Encyclopedia of Social Science Research Methods*. Thousand Oaks, CA: SAGE Publications; 2010:748.
25. Hupe M. StatPearls. *J Electron Res Med Libraries* 2024;21(3):147-152. [doi: [10.1080/15424065.2024.2389325](https://doi.org/10.1080/15424065.2024.2389325)]
26. Purinton EF, Burke MM. Engaging online students: using a multisensory exercise for deeper, active learning. *Marketing Educ Rev* 2019;30(1):29-42. [doi: [10.1080/10528008.2019.1677478](https://doi.org/10.1080/10528008.2019.1677478)]
27. Taran I. Musical activities and games in the process of learning to play the piano. *Musical art in the educological discourse* 2024(8):36-40. [doi: [10.28925/2518-766x.2023.86](https://doi.org/10.28925/2518-766x.2023.86)]
28. Simones L, Rodger M, Schroeder F. Seeing how it sounds: observation, imitation, and improved learning in piano playing. *Cognition and Instruction* 2017;35(2):125-140. [doi: [10.1080/07370008.2017.1282483](https://doi.org/10.1080/07370008.2017.1282483)]

29. Yang Z. Modern piano teaching technologies: accessibility, effectiveness, the need for pedagogues. *Elementary Educ Online* 2020;19(3):1812-1819 [[FREE Full text](#)]
30. Fadhli M, Brick B, Setyosari P, Ulfa S, Kuswandi D. A meta-analysis of selected studies on the effectiveness of gamification method for children. *Int J Instruction* 2020;13(1):845-854. [doi: [10.29333/iji.2020.13154a](#)]
31. Robert D, Bt. Jamri N, Sandra Hazel Ling, Bt. Amin AA, Bt. Yazid FA. Gamified learning intervention to promote music literacy and creativity in elementary music education. *JCSHD* 2023;9(1):18-41. [doi: [10.33736/jcs hd.5481.2023](#)]
32. Qian C, Jiang M. Exploring the effects of digital game-based learning on music education. *Studies in Soc Sci Humanities*. URL: <https://www.paradigmexpress.org/SSSH/article/view/1118> [accessed 2025-12-23]
33. Smith S. Expressive piano technic: basics of healthy movement, technic and interpretation. *Rast Müzikoloji Dergisi*. 2021;9(2):2757-2768. [doi: [10.12975/rastmd.2021923](#)]
34. Martinović Bogojević J, Rotar Pance B. Musical creativity in the teaching practice in Montenegrin and Slovenian primary schools. *Brit J Music Ed* 2022;39(2):169-182. [doi: [10.1017/s0265051722000018](#)]
35. Yeboah R, Amponsah KD, Commey Mintah P, Sedofia J, Kwarteng Donkor (Opare) PB. Game-based learning in Ghanaian primary schools: listening to the views of teachers. *Education 3-13* 2023;53(1):99-113. [doi: [10.1080/03004279.2023.2171269](#)]
36. Molero D, Schez-Sobrino S, Vallejo D, Glez-Morcillo C, Albusac J. A novel approach to learning music and piano based on mixed reality and gamification. *Multimed Tools Appl* 2020;80(1):165-186. [doi: [10.1007/s11042-020-09678-9](#)]
37. Guan Y, Meeson P. The application of Dalcroze eurhythmics in primary school music education. *Educ Rev* 2024;8(2):340-344. [doi: [10.26855/er.2024.02.027](#)]
38. Safian A, Nasrifan MN. Higher-order thinking skills in music education: a comprehensive analysis. *Grenek: j seni musik* 2013;13(2):182-191. [doi: [10.24114/grenek.v13i2.62780](#)]
39. Du S. What teaching methods can be used by piano teachers to improve learning outcomes of child and adult beginners? 2022 Presented at: Proceedings of the International Conference on Education, Language and Art; Nov 26-28, 2022; Sanya, China p. 26-28. [doi: [10.2991/assehr.k.220131.020](#)]
40. Ou C. Analysis and comparison of methods of teaching piano art in classical and modern educational systems: from traditions to innovations [Article in Russian]. *Educ Manag Rev* 2024;14(1-2):256-264. [doi: [10.25726/y5657-0174-6659-o](#)]
41. Ningtyas DP, Setyosari P, Kuswandi D, Ulfa S. Enhancing early childhood problem-solving abilities through game-based learning and computational thinking: the impact of cognitive styles. *Golden Age SCI J ECD* 2024;9(3):409-419. [doi: [10.14421/jga.2024.93-04](#)]
42. Avci A. Piano education through educational games: piano in the box game. *Base for Electronic Educ Sci* 2020;1(1):27-39. [doi: [10.29329/bedu.2020.253.3](#)]

Abbreviations

GBL: game-based learning

MCAR: missing completely at random

SMDM: Sidek Module Development Model

TREND: Transparent Reporting of Evaluations with Nonrandomized Designs

Edited by S Brini; submitted 16.Jul.2025; peer-reviewed by H Haruna, S Papadakis, Y Asada; comments to author 05.Oct.2025; revised version received 15.Dec.2025; accepted 21.Dec.2025; published 19.Jan.2026.

Please cite as:

Wang Y, Tan WH, Ye Q, Gu T

Effects of Game-Based Learning on Piano Music Knowledge Among Elementary School Pupils: Pretest-Posttest Quasi-Experimental Study

JMIR Serious Games 2026;14:e80766

URL: <https://games.jmir.org/2026/1/e80766>

doi: [10.2196/80766](#)

PMID:

©Yun Wang, Wee Hoe Tan, Qiaoling Ye, Tingting Gu. Originally published in JMIR Serious Games (<https://games.jmir.org>), 19.Jan.2026. 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.

Feasibility and User Experience of Immersive Virtual Reality–Based Rehabilitation in Patients With Stroke: Single-Arm Pretest-Posttest Pilot Study

Sara Arlati¹, PhD; Marta Mondellini^{1,2}, MS; Isabella Martinelli^{3,4}, MS; Eleonora Guanziroli⁵, PhD; Mauro Rossini⁶, MSc; Marco Sacco¹, MS; Franco Molteni⁵, MD

¹Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, National Research Council, via Previati 1/E, Lecco, Italy

²Department of Psychology, Università Cattolica del Sacro Cuore, Milano, Italy

³Institute of Health, School of Health Sciences, HES-SO Valais-Wallis, Sion, Switzerland

⁴Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

⁵Villa Beretta Rehabilitation Center, Ospedale Valduce, Costa Masnaga, Italy

⁶Ospedale Valduce, Como, Italy

Corresponding Author:

Sara Arlati, PhD

Institute of Intelligent Industrial Technologies and Systems for Advanced Manufacturing, National Research Council, via Previati 1/E, Lecco, Italy

Abstract

Background: Immersive virtual reality (VR) is promising in stroke rehabilitation; it is believed to promote motivation and intervention adherence among patients. However, existing work often lacks a structured assessment of user experience over a longer period of time.

Objective: This study aimed to assess the feasibility, user experience, and preliminary effectiveness of a VR-based rehabilitation program designed for patients with stroke to train upper limb and cognitive functions.

Methods: Thirty-two chronic (n=19, 59%) or postacute (n=13, 41%) patients with stroke (mean age 60, SD 11 years) were enrolled. All participants performed 4 weeks of training, performing exercises in the Virtual Supermarket for Stroke (VSS). The VSS is an ecological VR-based application allowing customization of difficulty to make the task of “doing the shopping” more challenging throughout the sessions. Subjective outcomes were assessed after the first and last sessions. Clinical scales were administered at baseline and at the end of the treatment.

Results: Of 32 participants, 31 (97%) completed the training. Flow (median 4.56, IQR 3.94-4.72; 5-point scale), sense of presence (Spatial Presence: median 3.44, IQR 12.85-3.85; Engagement: median 3.79, IQR 3.26-4.191; Naturalness: median 3.60, IQR 3.20-4.40; all 5-point scales), and affective state-related variables (Positive and Negative Affect Schedule; Positive Affect: median 4.60, IQR 4.00-5.00; Negative Affect: median 1, IQR 1.00-1.00) were satisfactory after the first session. Perceived ease of use was rated as very high (median 6.75, IQR 6.00-7.00; 7-point scale). No severe symptoms of cybersickness were recorded (Simulator Sickness Questionnaire [SSQ-TS]: median 11.22, IQR 0-20.57). At the end of the intervention, no significant differences were recorded in any subjective variable. Regarding clinical outcomes, significant improvements were recorded in balance (Berg Balance Scale pre: median 30, IQR 14.00-45.75; post: median 33.5, IQR 17.00-47.00; $P=.02$), upper limb motor functions (Motricity Index pre: median 45, IQR 15.25-69.00; post: median 46, IQR 32.00-77.00; $P=.02$; Box and Block pre: median 0, IQR 0-11.5; post: median 0, IQR 0-28; $P=.005$), and functional mobility (Time Up and Go pre: median 18, IQR 13.25-34.50; post: median 14, IQR 9.00-26.00; $P=.005$). No significant differences were recorded for general cognitive abilities (Mini-Mental State Examination pre: median 27, IQR 25-28; post: median 28, IQR 26-29), spasticity, and pain (visual analog scale pre: median 0, IQR 0-3.5; post: median 0, IQR 0-2).

Conclusions: The study showed the preliminary feasibility of a rehabilitation program using the VSS. It addressed the essential topic of assessing VR-based rehabilitation user experience throughout the entire training period, shedding light on the features that can contribute to an optimal psychological experience. Clinical outcomes suggested that the VSS promoted neuroplasticity and that the recorded improvements could translate into meaningful functional gains in daily activities. Further studies with larger samples and patients with more severe disabilities are needed to confirm these results.

(*JMIR Serious Games* 2026;14:e79584) doi:[10.2196/79584](https://doi.org/10.2196/79584)

KEYWORDS

motor training; user experience; subjective assessment; chronic patients; neuroplasticity

Introduction

Stroke is the second leading cause of death and the third leading cause of disability worldwide [1,2], and the number of people affected by stroke has doubled in the last 30 years [2]. The majority of patients with stroke survive after the acute event and live with deficits in the motor, sensory, cognitive, mood, and behavioral domains for many years [3,4]. The severity of such deficits and the rate of recovery may vary depending on the type, location, and extent of the brain lesion [5]. The burden of disability after stroke is very high due to economic consequences on health care, social care, and loss of productivity.

Therefore, research and therapies aimed at improving the quality of life of patients with stroke are crucial [6]. Even though most recovery seems to occur during the first few months after stroke, patients can obtain clinical improvements even several months after the event [7]. Current literature suggests that positive outcomes can be achieved through functional training and pharmacological treatments; however, contrasting results [8] and existing barriers to implementing evidence-based recommendations [9] underscore the need for a deeper understanding of stroke rehabilitative treatments and their effects [6,8]. In addition, an open point remains on the administration of the intervention: it is currently unclear what the optimal dosage, frequency, and duration of the proposed therapy are, especially considering the variability and severity of symptoms that patients with stroke may exhibit, as well as the time from the event [9].

Although precise guidelines have yet to be defined, it is currently agreed that adherence to the proposed intervention is essential. The importance of continuing to exercise has been emphasized in several studies, which have highlighted the potential of intensive, task-specific, and meaningful exercises to recover sensorimotor functions [10,11]. Nonetheless, the average adherence rate to therapists' recommendations after discharge ranges between 40% and 70% [12].

One of the solutions that researchers have explored to improve patients' adherence and effectiveness of nonpharmacological interventions is the employment of new technologies, among which virtual reality (VR) has become one of the most popular [10,11,13]. VR is a medium that immerses the user in computer-generated scenarios and enables interaction with the surroundings. It has several characteristics that make it suitable for implementing or supporting rehabilitation interventions [11]: it allows the simulation of ecologically valid tasks, facilitates the transfer of acquired abilities to real life, controlled repetitions, and easy customization of difficulty [14]. Furthermore, VR has been shown to promote neuroplasticity and brain reorganization, thus enhancing recovery [15]. Additionally, it can provide real-time feedback that helps patients regain awareness of their performance [14] and promote training in potentially unsafe situations (eg, road crossing) that would not be feasible in reality [16].

Recently, therapeutic VR-based applications have begun to be assessed in the context of the concept of "flow." Flow is a term used to describe a subjective psychological state that people experience when they are fully engaged in an activity to the point of losing track of time and their surroundings, except for the activity itself [17]. The concept of flow was first conceptualized by Csikszentmihalyi et al [17], which described it as an "optimal experience" characterized by 9 key features: *challenge-skill balance* (the activity difficulty matches the individual's ability); *action-awareness merging* (actions feel automatic and effortless); *clear goals*; *unambiguous feedback* (feedback guides performance); *concentration on the task*; *sense of control* (the user can influence the situation or outcomes); *loss of self-consciousness*; *transformation of time* (time may seem to speed up or slow down); and *autotelic experience* (the activity is intrinsically rewarding and enjoyable).

The experience of flow was first assessed in sport, educational, and work contexts, as it was noted that positive psychological flow states are correlated with better performance. Nonetheless, it has been acknowledged that the flow concept applies well to applications developed with therapeutic purposes [18]. VR-based applications for rehabilitation must have a clear goal, provide immediate feedback, and the level of difficulty can often be customized. Additionally, these applications often incorporate gamified elements to maintain high levels of attention and challenge.

The systematic assessment of flow in the rehabilitative context could represent an opportunity. First, it would enable the determination of whether a specific VR application can promote an optimal psychological experience, thereby indicating that it is usable, acceptable, and meaningful, ultimately leading to increased treatment compliance [19,20]. Second, it could suggest that it has the potential to enhance the therapeutic efficiency [18].

Although the evaluation of flow and user experience in general is often performed in sports [21,22] and work-related contexts [23], it is generally neglected or lacks analytical rigor in clinical studies [24,25]. A few studies observed a state of flow in their participants but did not assess it with appropriate tools [26]. A review conducted in 2021 analyzed the assessment of flow in VR applications developed for neurorehabilitation; the authors were able to trace only 10 studies [18]. More recently, Saric et al [27] and Pastore-Wapp et al [28] used flow to evaluate the user experience of individuals with Parkinson disease who underwent VR-based training for hand dexterity. Geiser and colleagues [29] assessed flow in patients with left neglect due to stroke who were trained with auditory motion stimulation for 3 weeks. In all cases, the positive experience was considered to be associated with the enjoyment of the experience and the motivation to continue using VR.

This work is positioned in this context by aiming to assess the feasibility, flow, and user experience of patients with stroke during a 4-week intervention with an immersive VR application specifically developed to support cognitive functions and upper

limb rehabilitation. The decision to develop an ad-hoc application was made to avoid the limitations commonly associated with commercial games, including inaccessible interactions for people with disabilities, a lack of ecological validity, and difficulty in customization [30]. The immersive VR application object of this work is derived from previous work carried out by our research group in previous years. In particular, we started with the development and validation of a virtual supermarket dedicated to the training of visuo-spatial abilities in which people could walk naturally; we assessed and demonstrated its usability and acceptance in a group of healthy young adults [31] and in older adults with mild cognitive decline or subjective cognitive decline [32]. Later, we started developing a second version with a simplified interaction, which was tested in the healthy population with positive results [33]. The application we used in this study, namely the Virtual Supermarket for Stroke (VSS), represents the continuation of this research line and, besides the simplified interaction, was specifically designed to reproduce an activity of daily living and to be customized to increase the level of difficulty throughout the 4 weeks of training.

We hypothesized that the chance of experiencing an activity of daily living (ADL) with a customized and adaptable level of difficulty could help maintain the flow levels of the exercise, as well as have other positive effects, and positively impact the sense of presence [22] and emotional state. Moreover, we expected the intervention to be feasible, with patients being able to interact with the application and not experiencing any adverse symptoms due to cybersickness.

Although this study was designed as a user-experience study, a secondary aim was to assess the effectiveness of the proposed VR-based intervention at the clinical level via standard clinical scales.

Methods

Ethical Considerations

The study obtained ethical clearance from the Insubria Ethical Committee (reference No. 70, December 1, 2020) and complies with the Declaration of Helsinki. All participants signed written informed consent before being enrolled; this consent also included permission to take pictures and record videos. To ensure the participants' privacy, data were collected in a pseudonymized form, and the keys were available only to the clinical personnel involved in the study. Also, any images included in the manuscript have been carefully reviewed to ensure that no individual can be identified. Participants did not receive any compensation for taking part in the study. This study adhered to the RATE-XR (Rationale, Accessibility, Training, and Effectiveness for Extended Reality) checklist [34] (Checklist 1).

Participants

Participants were recruited through convenience sampling among patients with stroke referred to the Villa Beretta

Rehabilitation Center (Costa Masnaga, LC, Italy) at the Rehabilitation Department of Valduce Hospital, where the study was carried out. Participants were either in the postacute (time from the stroke >15 d) or chronic phase (>6 mo). All met the following inclusion criteria: age 18 years or older; clinical stability; absence of pain, postural instability, muscle hyperactivity, and impairments that prevent the accomplishment of the reaching task; Mini-Mental State Examination (MMSE) 20 or an equivalent cognitive level for patients with aphasia.

Clinical stability, particularly for individuals with subacute poststroke, was defined as a period of 2 to 4 weeks poststroke during which patients completed acute hospitalization and were medically stable. This choice allowed for an optimal enrollment window of 1 to 3 months post stroke, where there is a balance between medical stability and the neuroplasticity window.

The cutoff set for MMSE score was selected based on several considerations. First, MMSE scores 20 or higher (≥ 20) indicate sufficient cognitive capacity to understand VR task instructions, safety protocols, and provide informed consent for participation in experimental technology [35,36]. Second, MMSE scores below 20 (< 20) have been associated with reduced rehabilitation potential and motor learning capacity in stroke populations, which could limit the ability to benefit from VR-based interventions [37,38].

Third, VR rehabilitation requires intact working memory, sustained attention, and executive function to engage with immersive environments and process multimodal feedback—cognitive domains that are relatively preserved in individuals with an MMSE score of 20 or higher [39,40]. Finally, this cutoff is consistent with previous VR stroke rehabilitation trials [41,42] and minimizes floor effects on secondary cognitive outcome measures.

Exclusion criteria were a history of seizure or motion sickness, severe visual deficits, and inability to provide informed consent. The sample size was calculated using the formula reported by Candel and van Breukelen [43] for single-arm pre-post studies, and considering flow (specifically Short Flow Scale; see Outcomes section) as the main outcome. Given previous experiences and available literature on the topic [44], the required sample was 23 participants, using $SD=1.2$, $\alpha=.05$, $power=80\%$, and expecting a pre-post variation equal to 0.65. The final sample size was increased to 30, considering the possibility of 20% dropouts and aligning the final sample to other studies assessing feasibility [45].

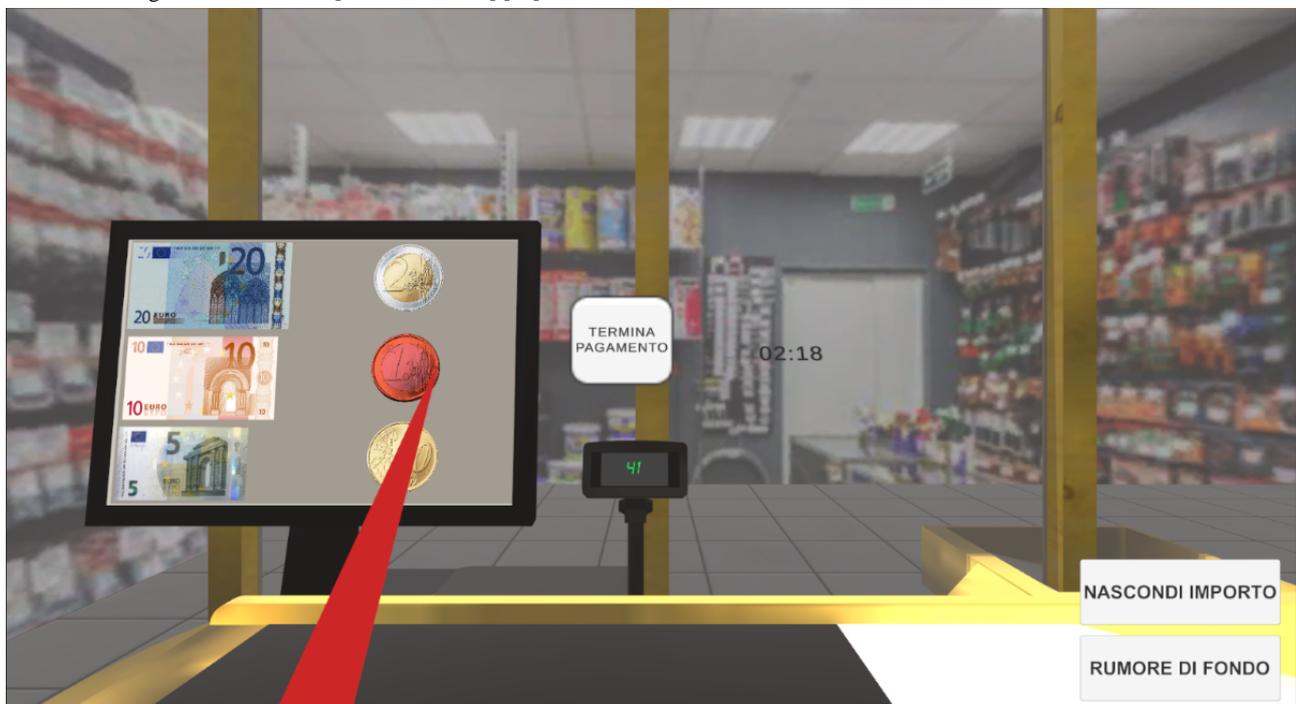
Equipment

The VSS is an immersive VR application dedicated to the rehabilitation of stroke patients. It was developed with Unity and deployed for Oculus Rift v2. The VR environment is constituted by 2 scenarios. The first is devoted to picking groceries from the shelf (Figure 1); the second is paying for such items (Figure 2). Both these tasks can be performed while staying seated to ensure all participants' safety and make the VSS accessible to patients in wheelchairs, too.

Figure 1. A screenshot of the aisle environment of the Virtual Supermarket for Stroke. To complete the shopping, the button “end shopping” [termina spesa] has to be clicked. Buttons at the bottom-right corner are available to the therapist to change the visibility of the list [visibilità lista], to turn the background noise on/off [rumore di fondo], and to replay the list items [leggi lista]. The basket could be placed on the right or the left side, depending on the upper limb to train [46].



Figure 2. A screenshot of the payment environment of the Virtual Supermarket for Stroke. The [termina pagamento] serves to end the selection of banknotes and coins. Buttons at the bottom right corner are available to the therapist to change the visibility of the amount to pay [visibilità importo] and to turn the background noise on/off [rumore di fondo] [46].



A series of features is available to therapists to customize the difficulty of training throughout the rehabilitative path. Before starting the exercises, the therapist can adjust the position (right/left) of the shopping basket, the number of items on the shelves (9 or 15), the number of target items (ie, items to be picked from the shelves) on the shopping list (from 1 to 8), and

select if grocery items are shown on the list with their names or as categories (eg, dairy products, dishware, vegetables); in this second case, the task is to pick all the items corresponding to such categories. Moreover, during the exercise, some buttons (visible only on the therapist’s PC screen) allowed the therapist to play a background noise, which served as a distracting sound,

and to hide or show the shopping list or the amount of money to be paid. When the shopping list was hidden, the therapist could play a clip (created via text-to-speech) indicating to the patient the grocery items to pick. All the therapists/psychologists involved in the study were trained by the VSS developer to set up the game area and use the application, which was developed by the researchers who created the VSS.

All the interactions within the VR environment occurred using the Oculus controller. The grocery items could be picked using a direct interaction, that is, reaching for the item and pressing the controller grip button to hold it and transport it around. All the other interactions (ie, coin selection and button press) occurred using a ray interactor (Figure 2) [47].

At the end of each task (item grabbing and payment), a panel was shown to the patients reporting their performance (ie, the time needed to complete the task, the errors, the omissions, or the amount of money still to pay).

Given its features, the VSS was expected to promote the retraining of motor and cognitive functionalities, in particular, the upper limb task-related reaching movement, a goal-directed, functional movement carried out in a natural environment [48], and visuospatial search, memory, and action planning.

Protocol

Each patient underwent 12 sessions of 20 minutes, in which the tasks of doing the shopping and paying at the cash register were repeated recursively. The therapist chose the level of difficulty, the presence of distractors, and other elements that increase the complexity of the task, both at the beginning and during the session. All sessions occurred in a quiet room, where only the patient and the therapist were present. Upon arrival, patients using a wheelchair were guided into the game area, and their wheelchair was braked. Patients who walked or used crutches were accommodated in a chair. All were then helped to wear and adjust the head-mounted display. Patients completed the task with the impaired arm whenever possible; if their motor functions were too low on the impaired side, they used the other upper limb.

During the first session, each participant was presented with the VSS and the tasks to perform, along with an explanation of the interaction modalities. All performed the first session at the lowest level of difficulty. After 20 minutes of exercise, the psychologist supervising the session administered paper-based questionnaires evaluating their user experience. The same questionnaires were administered again to all patients immediately after the last session. In the case of patients with aphasia, questionnaires were answered by pointing.

Patients were also administered the clinical scales presented in the Outcomes section at baseline and after completing the last exercise session.

Outcomes

The user experience was evaluated by administering a series of questionnaires at the end of the first and last sessions. Questionnaires were administered by a psychologist, and answers were collected on paper. The collected questionnaires were:

1. The Short Flow Scale (SFS) [49], which uses 9 items to measure *flow* (ie, a measure of absorption in what one is doing and how much the subjective experience is optimized);
2. The International Test Commission—Sense of Presence Inventory (ITC-SOPI) [50], which evaluates the Sense of Presence (ie, the sense of “being there” in a computer-generated environment along 4 subscales: spatial presence, engagement, naturalness, and side effects);
3. The Positive and Negative Affect Schedule—Short Form [51] for the assessment of positive and negative traits and states (ie, joy, high levels of energy, concentration, and distress, anger, contempt, or fear, respectively);
4. The Simulator Sickness Questionnaire (SSQ) [52] to assess cybersickness according to Nausea (SSQ-N), Oculomotor disturbance (SSQ-O), and Disorientation (SSQ-D), and general score (SSQ-TS); and
5. The Technology Acceptance Model (TAM3) subscale for the evaluation of perceived ease-of-use (PEOU) [53].

SFS, ITC-SOPI, and PANAS were assessed on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree), whereas PEOU was evaluated with a scale ranging from 1 to 7 (maximum perceived ease of use); within SSQ, the occurrence of each symptom was assessed from *not at all* to *severe* on a 4-point scale.

To evaluate whether the proposed intervention brought clinical improvements, the following clinical scales were administered before and after the training period:

1. MMSE [54] or an equivalent for patients with aphasia: MMSE is a brief 30-point questionnaire used to screen general cognitive impairments. It includes orientation, immediate memory, attention, calculation, recall, and language tasks; it represents a common instrument to assess cognitive changes.
2. Visual analog scale for pain [55]: the scale consists of a 10 cm horizontal line going from 0 (no pain) to 10 (worst imaginable pain). It represents a quick instrument to subjectively assess pain, which, in stroke, is mostly investigated from a clinical diagnosis perspective rather than from the patient’s perspective [56].
3. Motricity Index (MI) [57]: it is a clinical scale used to assess limb motor function, especially in individuals with stroke. It measures muscle strength in the upper limb (pinch grip, elbow flexion, and shoulder abduction) and lower limb (ankle dorsiflexion, knee extension, and hip flexion). Each movement is scored on a 6-point scale (0 - 33), based on the degree of movement and resistance against gravity.
4. Modified Ashworth Scale [58] for spasticity on arm abductors and pronator, elbow flexor and extensor, wrist flexor and extensor, and finger flexors (flexor digitorum superficialis and profundus, and flexor pollicis longus): an examiner performs it by moving a limb passively through its range of motion and grades the resistance felt in the muscle from 0 (no increase in muscle tone) to 4 (rigid flexion or extension).
5. Box and Block test (BnB) [59] with both arms (affected [A] and nonaffected [NA]): it constitutes a measure of manual dexterity, and it is computed as the number of

wooden blocks (2.5 cm) transferred from one compartment to another in 1 minute.

6. Berg Balance Scale [60]: it consists of 14 functional balance tasks that are rated by the examiner from 0 (unable to perform) to 4 (performs independently and safely). The maximum score is 56; scores below 45 indicate risk of falls.
7. Timed Up and Go test [61]: it is used to assess mobility, balance, and fall risk. The examinee starts seated, stands up, walks 3 m, turns around, and walks back to return seated. The time taken to perform the task constitutes its score.

Statistical Analysis

To assess the randomness of missing data, we used the Little's Missing Completely At Random (MCAR) test. Multiple imputation was performed as a sensitivity analysis to evaluate the robustness of the results with respect to missing data handling.

The reliability of the questionnaires' structure was calculated with the McDonald Ω index; a frequently cited acceptable range

of the Ω coefficient is a value of 0.70 or higher [62,63]. The normality of the data distribution was assessed using the Kolmogorov-Smirnov test. Descriptive statistics are used to describe the outcomes. In the case of nonnormal distribution, the 95% CIs estimated via bootstrapping (2000 resamples) pre-post differences were assessed via Wilcoxon signed rank test. In all analyses, statistical significance was set at P less than .05.

Results

Participants and Missing Data

Thirty-two patients with a mean age of 60 (SD 11) years were enrolled in the study. One participant experiencing the VSS is shown in [Figure 3](#). The median time from the stroke event was 7 (IQR 16, min: 16 d, max: 11 y) months, with 13 participants being in the subacute phase and 19 in the chronic stage. Twenty-three participants had right-brain damage. Demographic and baseline data of the study participants are presented in [Multimedia Appendix 1](#).

Figure 3. One of the participants during a Virtual Supermarket for Stroke session.



All patients except one completed the training; no adverse events were recorded. One participant left the study because they did not feel familiar with and did not feel at ease while interacting with technology. For this participant, only clinical scales were collected at T1, leading to user experience–related results being calculated on $n=31$. There were no other missing data.

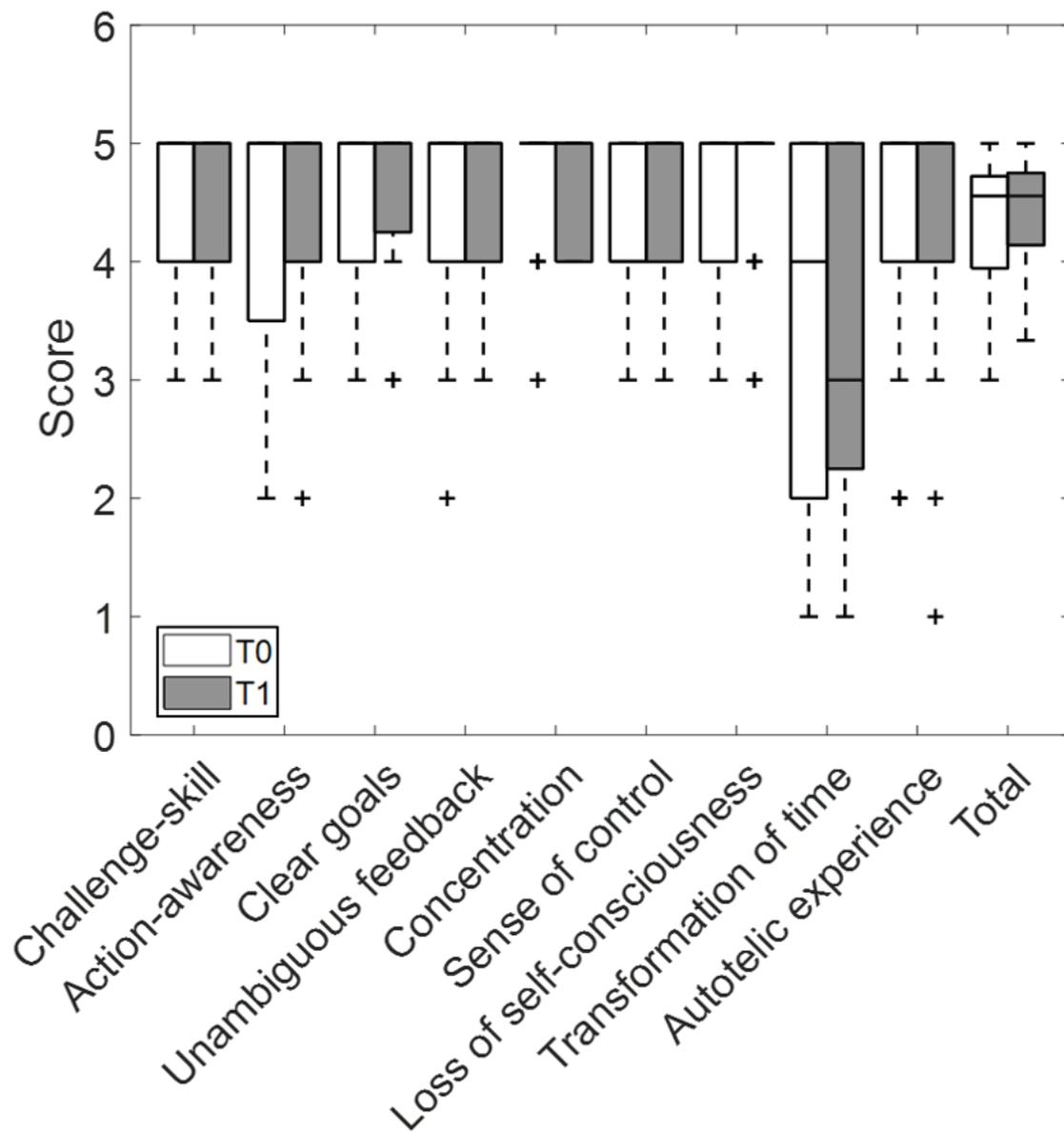
Missing data in the clinical scales were evaluated by recoding nonparticipation cases (eg, patients unable to walk could not perform the Time Up and Go test at any time point) as valid scores. Considering this, the overall proportion of missing data was approximately 4%, with up to 12% missingness in 5 clinical

scales. The Little's MCAR test was not significant ($\chi^2_{85}=104.308$; $P=.07$), indicating that other data were missing completely at random. Using multiple imputation as a sensitivity analysis, results were highly consistent across approaches, indicating that missing data did not materially influence the findings. Therefore, complete-case data were analyzed.

User Experience

The results obtained for all flow subscales are reported in [Figure 4](#); there are no statistically significant differences for any of them ($P>.05$). The increase in the total flow score was almost significant ($P=.05$).

Figure 4. Flow subscales collected from participants with stroke using the Virtual Supermarket for Stroke in the pretest-posttest study. Data are shown as boxplots, where whiskers extend to the most extreme values within 1.5 times the IQR and individual points denote outliers. Measurements were performed at T0, after the first session, and at T1. T1 was collected after 4 weeks of training, 3 times per week, at the end of the 12th session.



Descriptive statistics for the user experience questionnaires investigating presence, positive and negative affect, cybersickness, and perceived ease of use are reported in Table 1.

Table . Descriptive statistics of user experience-related variables collected from participants with stroke using the Virtual Supermarket for Stroke in the pretest-posttest study.^a

	Pre						Post							
	Median (IQR)	Min	Max	95% CI	Ω	Deleted items	K-S test (<i>P</i> values)	Median (IQR)	Min	Max	95% CI	Ω	Deleted items	K-S test (<i>P</i> values)
Flow	4.56 (0.64)	3	5	4.33-4.67	0.79	— ^b	.01 ^c	4.57 (1.00)	3	5	4.33-4.67	0.64	7, 8	.007 ^c
SP ^d	3.44 (1)	1.59	4.82	3.15-3.71	0.87	—	.01 ^c	3.71 (1.18)	1.41	4.88	3.16-3.80	0.89	—	.18
ENG ^e	3.79 (1)	2.17	4.92	3.45-3.90	0.74	5	.12	3.83 (1.00)	1.5	4.75	3.54-4.08	0.89	5	.048 ^c
NAT ^f	3.60 (1.60)	2	5	3.59-4.08	0.75	27	.16	4.00 (1.80)	1.2	5	3.20-4.20	0.82	—	.02 ^c
PA ^g	4.60 (1)	1.40	5	4-4.80	0.94	—	<.001 ^c	4.60 (0.80)	3.8	5	4.20-5	0.63	—	<.001 ^c
NA ^h	1 (0)	1	3	1-1	0.93	—	<.001 ^c	1 (0)	1	2	1-1	0.84	—	<.001 ^c
SSQ-N ⁱ	0 (4.77)	0	104.94	0-0	0.91	—	<.001 ^c	0 (7.16)	0	28.62	0-0	0.71	1	—
SSQ-O ^j	15.16 (22.74)	0	75.8	7.58-22.74	0.78	—	<.001 ^c	7.58 (15.16)	0	68.22	0-15.16	0.95	1	—
SSQ-D ^k	13.92 (27.84)	0	55.68	0-27.84	0.82	—	<.001 ^c	0 (27.84)	0	55.68	0-13.92	0.92	5, 8	<.001 ^c
SSQ-TS ^l	11.22 (20.57)	0	112.2	0-16.83	0.94	—	<.001 ^c	7.48 (14.96)	0	48.62	0-14.96	0.72	Nausea	<.001 ^c
PEOU ^m	6.75 (1)	3.75	7	6-7	0.94	—	<.001 ^c	6.75 (1)	3.8	5	6.25-7	0.85	—	<.001 ^c

^aData were collected at T0 (pre), after the first session, and at T1 (post). T1 was collected after 4 weeks of training, 3 times per week, at the end of the 12th session. Data is presented as median values with IQRs, 95% CIs, and minimum and maximum values. Results of the Ω test for reliability and the Kolmogorov-Smirnov test are also reported.

^bNot applicable.

^cIndicates statistical significance ($P < .05$).

^dSP: spatial presence.

^eENG: engagement.

^fNAT: naturalness.

^gPA: positive affect.

^hNA: negative affect.

ⁱSSQ-N: Simulator Sickness Questionnaire—Nausea.

^jSSQ-O: Simulator Sickness Questionnaire—Oculomotor Disturbance

^kSSQ-D: Simulator Sickness Questionnaire—Disorientation.

^lSSQ-TS: Simulator Sickness Questionnaire—General Score.

^mPEOU: perceived ease of use.

No significant difference was found between pre- and postassessment regarding negative effects, engagement, realism, spatial presence, ease of use, and positive and negative affects. No significant difference was found for nausea, oculomotor, and disorientation symptoms or the total score of the Simulator Sickness Questionnaire.

Clinical Outcomes

The measures of clinical outcomes before and after the 4 weeks of intervention are reported in [Table 2](#). Significant improvements were recorded for the affected arm in the Box and Block test, MI (elbow and total scores). Moreover, patients improved their balance and their performance in the Time Up and Go test. Significant differences are shown in [Figure 5](#).

Table . Descriptive statistics of the clinical scales collected from participants with stroke using the Virtual Supermarket for Stroke in the pretest-posttest study.^a

	Pre		Post		P value
	Median (IQR)	95% CI	Median (IQR)	95% CI	
MMSE ^b	27 (3)	25-28	28 (3)	27-29	.15
VAS ^c	0 (3.50)	0-2	0 (2)	0-1	.35
BBS ^d	30 (31.80)	20-37	33.50 (30)	23-44	.02 ^e
MAS ^f					
Arm abductor	0 (0)	0-0	0 (0)	0-0	.50
Arm internal rotation	0 (0.50)	0-0	0 (0.75)	0-0	.94
Arm pronator	0 (0.50)	0-0	0 (0.75)	0-0	>.99
Elbow extensor	0 (0)	0-0	0 (0)	0-0	.94
Elbow flexor	0 (1)	0-1	0 (1.38)	0-1	.50
Wrist extensor	0 (0)	0-0	0 (0)	0-0	>.99
Wrist flexor	0 (0.50)	0-0	0 (1)	0-0	>.99
Finger extensor	0 (0)	0-0	0 (0)	0-0	>.99
FDS ^g	0.50 (1.75)	0-1.25	1 (1.38)	0-1.25	.38
FDP ^h	0 (1)	0-0.50	0 (1)	0-0.50	>.99
FPL ⁱ	0 (1)	0-0	0 (1)	0-0	.28
BnB (A) ^j	0 (11.50)	0-29	0 (28)	0-41	.005 ^e
BnB (NA) ^j	44 (15.50)	31-45	50.50 (13.50)	29-51	.09
MI ^k (pinch)	11 (25)	0-19	19 (20.50)	11-26	.10
MI (elbow)	14 (14.80)	14-19	19 (15)	14-25	.02 ^e
MI (shoulder)	14 (14.50)	9-14	14 (16)	14-25	.19
MI tot	45 (53.75)	24-59	46 (45)	37.50-75	.02 ^e
TUG ^l (s)	18 (21.25)	14-34	14 (17)	10-23	.005 ^e

^aData were collected at T0 (PRE), after the first session, and at T1 (POST). T1 was collected after 4 weeks of training, 3 times per week, at the end of the 12th session. Data is presented as median values with interquartile ranges and 95% confidence intervals estimated via bootstrapping (2000 resamples).

^bMMSE: Mini-Mental State Examination.

^cVAS: visual analog scale.

^dBBS: Berg Balance Scale.

^eIndicates statistical significance ($P < .05$).

^fMAS: Modified Ashworth Scale.

^gFDS: flexor digitorum superficialis.

^hFDP: flexor digitorum profundus.

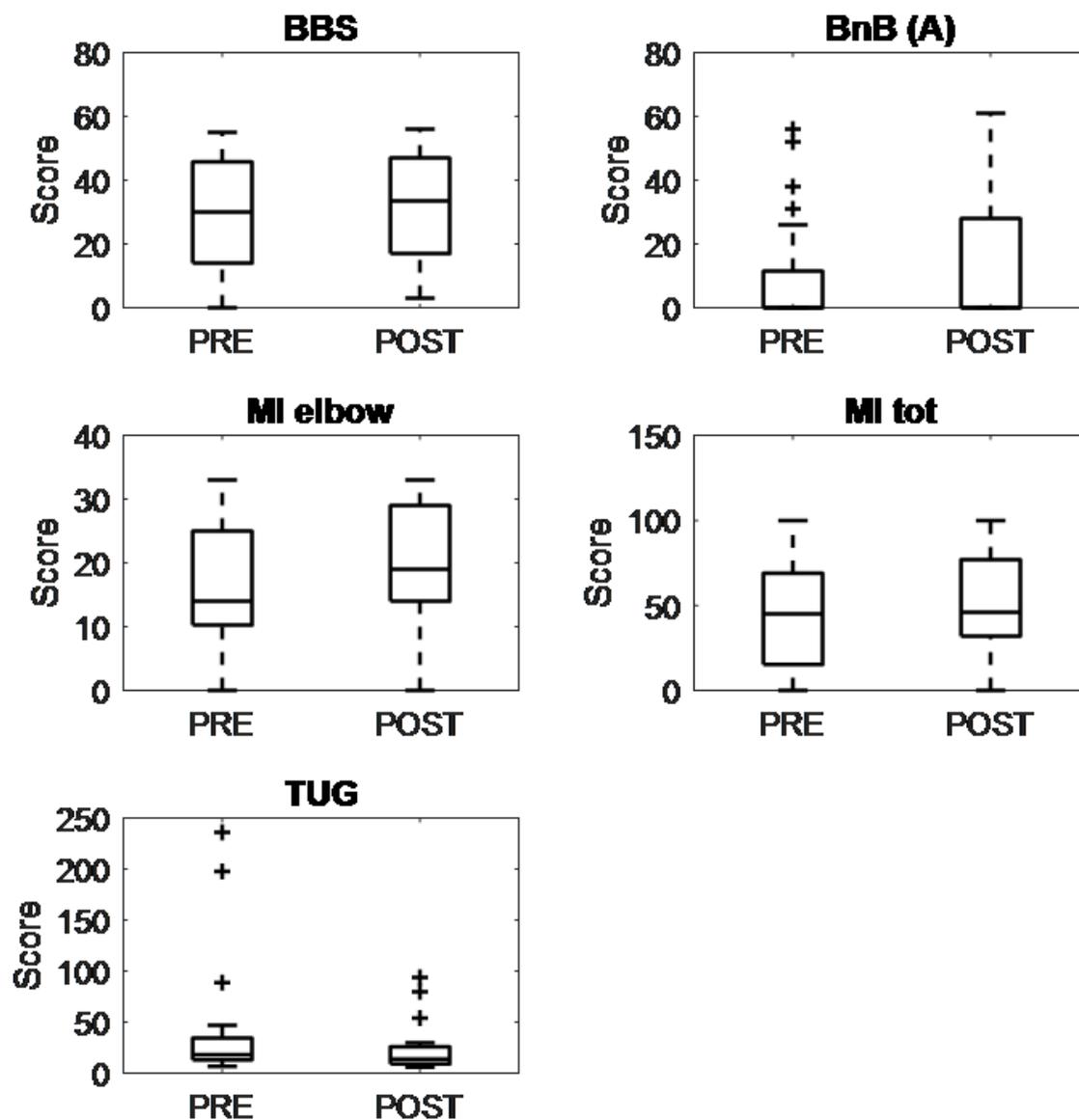
ⁱFPL: flexor pollicis longus.

^jBnB: box and block test with affected (A) and nonaffected (NA) arm.

^kMI: Motricity Index.

^lTUG: Timed Up and Go.

Figure 5. Boxplots of the clinical scales collected from participants with stroke using the Virtual Supermarket for Stroke (VSS) in the pretest-posttest study. Whiskers extend to the most extreme values within 1.5 times the interquartile range, and individual points denote outliers. The reported scales are the ones that showed a significant ($P<.05$) improvement postintervention (ie, after 4 wk of training, the VSS was used three times a wk). BBS: Berg Balance Scale; BnB (A): box and block (affected side); MI: Motricity Index; TUG: Time Up and Go.



Discussion

Main Findings

This study presents a feasibility assessment of a VR-based intervention for the rehabilitation of cognitive abilities and upper limb functions in patients with stroke, focusing on user experience and preliminary effectiveness. In particular, we assessed that the ecological environment and the features of the VSS application promoted an optimal psychological experience, enhancing flow and thereby increasing motivation and adherence throughout the entire duration of the training. We found that the general experience was positive, and participants welcomed the possibility of rehabilitation with the support of immersive

technologies, which also led to some improvements in functional outcomes. Also, we demonstrated that the designed intervention was feasible, with no side effects and excellent adherence among stroke survivors (97%).

Subjective Outcomes

The subjective outcomes collected after the first session showed satisfying levels, with flow and positive affect scores very close to the highest possible value. The same was recorded for PEOU, possibly indicating that the intuitive interaction with virtual objects in the VSS also mediated the positive experience. This finding aligns with a previous study that investigated the acceptance of immersive VR technologies among patients with stroke [30,64].

The choice of a simple interaction methodology was effective and made possible by the use of an ad-hoc, developed application. In fact, the application's features made the solution easily accessible, even for patients with motor limitations. On the contrary, the use of a commercially available application, which would have possibly provided more varied and engaging scenarios, would have required checking it for customization settings to enable effective interaction for patients with stroke [65].

Presence-related scales were all satisfactory; thus, patients felt present and recognized the environment as an ecological and familiar one. Conversely, negative affect was low, and median cybersickness scores were all within the acceptable range (≤ 20) [52]. This confirmed the outcomes obtained before with slightly different versions of the Virtual Supermarket and different populations (healthy volunteers [33] and older adults with cognitive decline [32,66]). In particular, in this case, we recorded almost no nausea-related symptoms; instead, the symptoms were mostly present in the oculomotor and disorientation domains, as more typically occurs in nonnavigational environments [67].

At the end of the period of trial, all the subjective outcomes preserved the same satisfactory and, in some cases, more than satisfactory trends. This demonstrated that the design choice we made for the application effectively addressed the maintenance of flow throughout the experience, that is, we could balance patients' skills and the challenge of the task, and provide an engaging scenario in which they could feel focused and immersed. Also, the obtained outcomes confirmed our hypothesis that introducing some challenging elements (ie, the possibility of customizing the levels of difficulty), addressing different cognitive abilities (eg, attention, memory), and the visualization of the performance could contribute to maintaining the experience engaging and able to induce positive feelings, even over a longer period of time.

These features represent essential points in the field of rehabilitation, in which prolonged, or even lifelong, interventions are often essential to counteract limb disuse or misuse, or cognitive decline [68]. As already mentioned, these variables and flow, in particular, have been explored in sport and other disciplines because of their positive association with performance, positive experience, motivation, and enjoyment [18]. Despite this, its assessment remains sparse in the health sector, which may negatively impact treatment compliance and the subsequent effectiveness of proposed VR-based interventions [65].

This has also been shown in a review performed in 2020, which focused on user experience in general and immersive VR; it showed that research in the field of human-immersive VR interaction still presents some methodological and technological gaps [69]. When examining the rehabilitation field, such gaps are even more evident: of the 65 articles included in the review, only four involved patients, and 35 included participants only in their 30 s.

Future studies should thus consider tackling the user experience assessment more broadly to evaluate patients' needs not only

from a clinical perspective but also from a subjective and intrinsic motivation point of view [65,70].

Clinical Outcomes

Regarding clinical outcomes, we observed positive results in various clinical scales. We recorded a significant improvement in the Berg Balance Scale and the Timed-Up-and-Go test, indicating increased trunk control in all patients. Moreover, considering the entire sample, we recorded an increase in the impaired limb scores, with significant changes in the Box-and-Block and oticity Index (total and elbow score) tests.

The observed improvements in upper limb dexterity (Box-and-Block Test) and motor strength (MI) suggest that the intervention may translate into meaningful functional gains in daily activities such as reaching, grasping, and manipulating objects [71]. The enhanced balance and mobility indicate potential benefits for overall independence and fall prevention [72].

These changes were recorded even if some patients performed the task with the nonimpaired arm, possibly suggesting that the improvements were mostly linked to better trunk control and action planning (which have recently been reported to be related [73]) rather than the arm motor function improvement per se. To this aim, the fact that the supermarket simulated an activity of daily living may have also contributed positively to the rehabilitation of executive functions [74]. Indeed, the ecological validity of shopping tasks has been previously verified in studies that utilize this scenario to assess and train executive functions and action planning, in particular [75].

The intervention proposed using the Virtual Supermarket frames within the task-oriented training. In fact, it emphasizes functional activities and patient involvement. Furthermore, it includes task repetition, active participation, and the modulation of the training intensity. From a clinical perspective, these interventions—even those supported by VR—have demonstrated effectiveness in enhancing limb function and balance, as well as promoting neuroplasticity [76,77].

Therefore, it is possible that, in our case as well, the modulating effects of neuroplasticity allowed us to observe clinical improvements even in chronic patients exercising with the less-affected limb [78,79]. Unfortunately, the small sample size, defined for the assessment of flow, did not allow us to explore subgroup differences further (ie, considering time from the stroke event, in the postacute or chronic phases, or the use of the affected or nonaffected side). However, the positive results should encourage the conduct of more structured, clinical-oriented trials to unveil potential indications for future intervention characterization. In fact, the recorded improvements across multiple functional domains support the potential clinical relevance of the intervention. The effects of the proposed intervention on both upper limb function and balance suggest that it could be integrated into standard rehabilitation protocols to promote more holistic recovery.

Limitations

All the discussed results must be treated cautiously. We are aware that the current study presented some limitations that

prevented the generalization of the results. First, the sample size was small and estimated based on the assessment of our main outcome (ie, flow); thus, the statistical power may not be sufficient to draw conclusions at the clinical level. Second, the population was heterogeneous, including both patients in the sub-acute and chronic phases who completed the exercise with either the impaired or less-impaired side. Finally, we did not have a control group. Nonetheless, the study allowed for assessing the user experience in a systematic and multidimensional way and highlighted the potential of a customizable, ecological, and immersive VR-based application for the rehabilitation of motor functions in patients with stroke.

Conclusions

This work addressed an important aspect in the context of VR-assisted rehabilitation, as user experience is generally evaluated on a single session or by focusing on a single aspect (eg, usability).

It contributed to the current body of knowledge on the use of immersive VR for stroke rehabilitation, providing insights into

the features that helped maintain high flow levels and highlighting the importance of exploring the subjective domain in clinical trials as well.

From a clinical perspective, the intervention appears feasible and well-tolerated. Its simple setup and ease of use make it potentially adaptable to a variety of clinical settings, and potentially for future use at home or in unsupervised contexts (eg, with a stand-alone head-mounted display and automatic progression of difficulty).

In the future, it would be valuable to increase the difficulty levels and include different interaction technologies (eg, hand tracking) to make the application more accessible to patients with limited arm motor function, thereby providing a longer training period.

Moreover, as already mentioned, it would be essential to enlarge the sample of participants, thus allowing for a more in-depth investigation of the optimal phase for administering the exercise and, in general, of the clinical effectiveness of the VSS application.

Funding

This work was supported by Fondazione Cariplo under Grant EMPATIA@Lecco—“EMpowerment del PAzienTe In cAsa.” The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The anonymized data collected during this study are available from the corresponding author upon a reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Demographic and baseline data of the patients enrolled in the study. BBS: Berg Balance Scale; flex: flexion; BnB: box and block test with the dominant (D) and the nondominant arm (ND); ext: extension; FDP: flexor digitorum profundus; FDS: flexor digitorum superficialis; FPS: flexor pollicis longus; IR: internal rotation; MAS: Modified Ashworth Scale; MI: Motricity Index; MMSE: Mini-Mental State Examination; VAS: visual analog scale; TUG: Time Up and Go.

[[DOCX File, 30 KB - games_v14i1e79584_app1.docx](#)]

Checklist 1

RATE-XR checklist.

[[PDF File, 196 KB - games_v14i1e79584_app2.pdf](#)]

References

1. Pu L, Wang L, Zhang R, Zhao T, Jiang Y, Han L. Projected global trends in ischemic stroke incidence, deaths and disability-adjusted life years from 2020 to 2030. *Stroke* 2023 May;54(5):1330-1339. [doi: [10.1161/STROKEAHA.122.040073](https://doi.org/10.1161/STROKEAHA.122.040073)] [Medline: [37094034](https://pubmed.ncbi.nlm.nih.gov/37094034/)]
2. Feigin VL, Owolabi MO, Feigin VL. Pragmatic solutions to reduce the global burden of stroke: a World Stroke Organization-Lancet Neurology Commission. *Lancet Neurol* 2023 Dec;22(12):1160-1206. [doi: [10.1016/S1474-4422\(23\)00277-6](https://doi.org/10.1016/S1474-4422(23)00277-6)] [Medline: [37827183](https://pubmed.ncbi.nlm.nih.gov/37827183/)]
3. Gittins M, Lugo-Palacios D, Vail A, et al. Stroke impairment categories: a new way to classify the effects of stroke based on stroke-related impairments. *Clin Rehabil* 2021 Mar;35(3):446-458. [doi: [10.1177/0269215520966473](https://doi.org/10.1177/0269215520966473)] [Medline: [33131321](https://pubmed.ncbi.nlm.nih.gov/33131321/)]
4. Carota A, Staub F, Bogousslavsky J. Emotions, behaviours and mood changes in stroke. *Curr Opin Neurol* 2002 Feb;15(1):57-69. [doi: [10.1097/00019052-200202000-00010](https://doi.org/10.1097/00019052-200202000-00010)] [Medline: [11796952](https://pubmed.ncbi.nlm.nih.gov/11796952/)]
5. de Haan RJ, Limburg M, Van der Meulen JH, Jacobs HM, Aaronson NK. Quality of life after stroke. Impact of stroke type and lesion location. *Stroke* 1995 Mar;26(3):402-408. [doi: [10.1161/01.str.26.3.402](https://doi.org/10.1161/01.str.26.3.402)] [Medline: [7886714](https://pubmed.ncbi.nlm.nih.gov/7886714/)]

6. Cramer SC, Wolf SL, Adams HP Jr, et al. Stroke recovery and rehabilitation research: issues, opportunities, and the National Institutes of Health StrokeNet. *Stroke* 2017 Mar;48(3):813-819. [doi: [10.1161/STROKEAHA.116.015501](https://doi.org/10.1161/STROKEAHA.116.015501)] [Medline: [28174324](https://pubmed.ncbi.nlm.nih.gov/28174324/)]
7. Saway BF, Palmer C, Hughes C, et al. The evolution of neuromodulation for chronic stroke: from neuroplasticity mechanisms to brain-computer interfaces. *Neurotherapeutics* 2024 Apr;21(3):e00337. [doi: [10.1016/j.neurot.2024.e00337](https://doi.org/10.1016/j.neurot.2024.e00337)] [Medline: [38377638](https://pubmed.ncbi.nlm.nih.gov/38377638/)]
8. Stinear CM, Lang CE, Zeiler S, Byblow WD. Advances and challenges in stroke rehabilitation. *Lancet Neurol* 2020 Apr;19(4):348-360. [doi: [10.1016/S1474-4422\(19\)30415-6](https://doi.org/10.1016/S1474-4422(19)30415-6)] [Medline: [32004440](https://pubmed.ncbi.nlm.nih.gov/32004440/)]
9. Hernandez A, Buby L, Archambault PS, Higgins J, Levin MF, Kairy D. Virtual reality-based rehabilitation as a feasible and engaging tool for the management of chronic poststroke upper-extremity function recovery: randomized controlled trial. *JMIR Serious Games* 2022 Sep 27;10(3):e37506. [doi: [10.2196/37506](https://doi.org/10.2196/37506)] [Medline: [36166289](https://pubmed.ncbi.nlm.nih.gov/36166289/)]
10. Laver KE, Lange B, George S, et al. Virtual reality for stroke rehabilitation. *Cochrane Database Syst Rev* 2025 Jun 20;6(6):CD008349. [doi: [10.1002/14651858.CD008349.pub5](https://doi.org/10.1002/14651858.CD008349.pub5)] [Medline: [40537150](https://pubmed.ncbi.nlm.nih.gov/40537150/)]
11. Masmoudi M, Zenati N, Izountar Y, et al. Assessing the effectiveness of virtual reality serious games in post-stroke rehabilitation: a novel evaluation method. *Multimed Tools Appl* 2024;83(12):36175-36202. [doi: [10.1007/s11042-023-17980-5](https://doi.org/10.1007/s11042-023-17980-5)]
12. Miller KK, Porter RE, DeBaun-Sprague E, Van Puymbroek M, Schmid AA. Exercise after stroke: patient adherence and beliefs after discharge from rehabilitation. *Top Stroke Rehabil* 2017 Mar;24(2):142-148. [doi: [10.1080/10749357.2016.1200292](https://doi.org/10.1080/10749357.2016.1200292)] [Medline: [27334684](https://pubmed.ncbi.nlm.nih.gov/27334684/)]
13. Lu W, Shi M, Liu L, et al. Effect of virtual reality-based therapies on lower limb functional recovery in stroke survivors: systematic review and meta-analysis. *J Med Internet Res* 2025 Jul 30;27:e72364. [doi: [10.2196/72364](https://doi.org/10.2196/72364)] [Medline: [40737641](https://pubmed.ncbi.nlm.nih.gov/40737641/)]
14. Villada Castillo JF, Montoya Vega MF, Muñoz Cardona JE, et al. Design of virtual reality exergames for upper limb stroke rehabilitation following iterative design methods: usability study. *JMIR Serious Games* 2024 Jan 11;12:e48900. [doi: [10.2196/48900](https://doi.org/10.2196/48900)] [Medline: [38206670](https://pubmed.ncbi.nlm.nih.gov/38206670/)]
15. Xu S, Xu Y, Wen R, Wang J, Qiu Y, Chan CC. Virtual reality enhanced exercise training in upper limb function of patients with stroke: meta-analytic study. *J Med Internet Res* 2025 Feb 19;27:e66802. [doi: [10.2196/66802](https://doi.org/10.2196/66802)] [Medline: [39969977](https://pubmed.ncbi.nlm.nih.gov/39969977/)]
16. Belger J, Wagner S, Gaebler M, et al. Application of immersive virtual reality for assessing chronic neglect in individuals with stroke: the immersive virtual road-crossing task. *J Clin Exp Neuropsychol* 2024 Apr;46(3):254-271. [doi: [10.1080/13803395.2024.2329380](https://doi.org/10.1080/13803395.2024.2329380)] [Medline: [38516790](https://pubmed.ncbi.nlm.nih.gov/38516790/)]
17. Abuhamdeh S, Nakamura J, Csikszentmihalyi M. Flow. In: Elliot AJ, Dweck CS, editors. *Handbook of Competence and Motivation*: The Guilford Press; 2005:598-608 URL: <https://philpapers.org/rec/ABUF> [accessed 2026-02-17]
18. Ottiger B, Van Wegen E, Keller K, et al. Getting into a “Flow” state: a systematic review of flow experience in neurological diseases. *J Neuroeng Rehabil* 2021 Apr 20;18(1):65. [doi: [10.1186/s12984-021-00864-w](https://doi.org/10.1186/s12984-021-00864-w)] [Medline: [33879182](https://pubmed.ncbi.nlm.nih.gov/33879182/)]
19. Aderinto N, Olatunji G, Abdulbasit MO, Edun M, Aboderin G, Egbunu E. Exploring the efficacy of virtual reality-based rehabilitation in stroke: a narrative review of current evidence. *Ann Med* 2023;55(2):2285907. [doi: [10.1080/07853890.2023.2285907](https://doi.org/10.1080/07853890.2023.2285907)] [Medline: [38010358](https://pubmed.ncbi.nlm.nih.gov/38010358/)]
20. Mondellini M, Rutkowski S, Colombo V. Cycling in immersive VR: motivation and affects in post-COVID patients. In: De Paolis LT, Arpaia P, Sacco M, editors. *Extended Reality: International Conference, XR Salento 2023, Lecce, Italy, September 6-9, 2023, Proceedings, Part II 2023*:353-366. [doi: [10.1007/978-3-031-43404-4_23](https://doi.org/10.1007/978-3-031-43404-4_23)]
21. Polechoński J, Szczechowicz B, Ryśnik J, Tomik R. Recreational cycling provides greater satisfaction and flow in an immersive virtual environment than in real life. *BMC Sports Sci Med Rehabil* 2024 Jan 30;16(1):31. [doi: [10.1186/s13102-024-00818-4](https://doi.org/10.1186/s13102-024-00818-4)] [Medline: [38291484](https://pubmed.ncbi.nlm.nih.gov/38291484/)]
22. Domínguez-González JA, Reigal RE, Morales-Sánchez V, Hernández-Mendo A. Analysis of the sports psychological profile, competitive anxiety, self-confidence and flow state in young football players. *Sports (Basel)* 2024 Jan 8;12(1):20. [doi: [10.3390/sports12010020](https://doi.org/10.3390/sports12010020)] [Medline: [38251294](https://pubmed.ncbi.nlm.nih.gov/38251294/)]
23. Bartholomeyczik K, Knierim MT, Weinhardt C, Oettingen G. Using mental contrasting to promote flow experiences at work: a just-in-time adaptive intervention. *Comput Hum Behav Rep* 2024 Dec;16:100488. [doi: [10.1016/j.chbr.2024.100488](https://doi.org/10.1016/j.chbr.2024.100488)]
24. Thomson K, Pollock A, Bugge C, Brady M. Commercial gaming devices for stroke upper limb rehabilitation: a systematic review. *Int J Stroke* 2014 Jun;9(4):479-488. [doi: [10.1111/ijvs.12263](https://doi.org/10.1111/ijvs.12263)] [Medline: [24661797](https://pubmed.ncbi.nlm.nih.gov/24661797/)]
25. Warland A, Paraskevopoulos I, Tsekleves E, et al. The feasibility, acceptability and preliminary efficacy of a low-cost, virtual-reality based, upper-limb stroke rehabilitation device: a mixed methods study. *Disabil Rehabil* 2019 Sep;41(18):2119-2134. [doi: [10.1080/09638288.2018.1459881](https://doi.org/10.1080/09638288.2018.1459881)] [Medline: [29644897](https://pubmed.ncbi.nlm.nih.gov/29644897/)]
26. Rätz R, Conti F, Thaler I, Müri RM, Marchal-Crespo L. Enhancing stroke rehabilitation with whole-hand haptic rendering: development and clinical usability evaluation of a novel upper-limb rehabilitation device. *J Neuroeng Rehabil* 2024 Sep 27;21(1):172. [doi: [10.1186/s12984-024-01439-1](https://doi.org/10.1186/s12984-024-01439-1)] [Medline: [39334423](https://pubmed.ncbi.nlm.nih.gov/39334423/)]
27. Saric L, Knobel SEJ, Pastore-Wapp M, Nef T, Mast FW, Vanbellinghen T. Usability of two new interactive game sensor-based hand training devices in Parkinson’s disease. *Sensors (Basel)* 2022 Aug 21;22(16):6278. [doi: [10.3390/s22166278](https://doi.org/10.3390/s22166278)] [Medline: [36016039](https://pubmed.ncbi.nlm.nih.gov/36016039/)]

28. Pastore-Wapp M, Kaufmann BC, Nyffeler T, Wapp S, Bohlhalter S, Vanbellingen T. Feasibility of a combined intermittent theta-burst stimulation and video game-based dexterity training in Parkinson's disease. *J Neuroeng Rehabil* 2023 Jan 12;20(1):2. [doi: [10.1186/s12984-023-01123-w](https://doi.org/10.1186/s12984-023-01123-w)] [Medline: [36635679](https://pubmed.ncbi.nlm.nih.gov/36635679/)]
29. Geiser N, Kaufmann BC, Cazzoli D, Nef T, Nyffeler T. Auditory motion stimulation as an add-on therapeutic approach is feasible, enjoyable, and associated with a significant improvement of neglect outcome: a non-randomized, controlled group-matched study. *Neuropsychol Rehabil* 2025 Dec;35(10):1903-1917. [doi: [10.1080/09602011.2025.2476068](https://doi.org/10.1080/09602011.2025.2476068)] [Medline: [40085736](https://pubmed.ncbi.nlm.nih.gov/40085736/)]
30. Specht J, Schroeder H, Krakow K, Meinhardt G, Stegmann B, Meinhardt-Injac B. Acceptance of immersive head-mounted display virtual reality in stroke patients. *Comput Hum Behav Rep* 2021;4:100141. [doi: [10.1016/j.chbr.2021.100141](https://doi.org/10.1016/j.chbr.2021.100141)]
31. Mondellini M, Arlati S, Pizzagalli S, et al. Assessment of the usability of an immersive virtual supermarket for the cognitive rehabilitation of elderly patients: a pilot study on young adults. Presented at: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH); May 16-18, 2018; Vienna, Austria p. 1-8. [doi: [10.1109/SeGAH.2018.8401313](https://doi.org/10.1109/SeGAH.2018.8401313)]
32. Arlati S, Di Santo SG, Franchini F, et al. Acceptance and usability of immersive virtual reality in older adults with objective and subjective cognitive decline. *J Alzheimers Dis* 2021;80(3):1025-1038. [doi: [10.3233/JAD-201431](https://doi.org/10.3233/JAD-201431)] [Medline: [33646164](https://pubmed.ncbi.nlm.nih.gov/33646164/)]
33. Mondellini M, Pizzagalli S, Greci L, Sacco M. Assessment of an immersive virtual supermarket to train post-stroke patients: a pilot study on healthy people. In: De Paolis LT, Arpaia P, Sacco M, editors. *Augmented Reality, Virtual Reality, and Computer Graphics: 6th International Conference, AVR 2019, Santa Maria al Bagno, Italy, June 24-27, 2019, Proceedings, Part I*:313-329. [doi: [10.1007/978-3-030-25965-5_23](https://doi.org/10.1007/978-3-030-25965-5_23)]
34. Vlakte JH, Drop DLQ, Van Bommel J, et al. Reporting guidelines for the early-phase clinical evaluation of applications using extended reality: RATE-XR qualitative study guideline. *J Med Internet Res* 2024 Nov 29;26:e56790. [doi: [10.2196/56790](https://doi.org/10.2196/56790)] [Medline: [39612482](https://pubmed.ncbi.nlm.nih.gov/39612482/)]
35. Kim SYH, Caine ED. Utility and limits of the mini mental state examination in evaluating consent capacity in Alzheimer's disease. *Psychiatr Serv* 2002 Oct;53(10):1322-1324. [doi: [10.1176/appi.ps.53.10.1322](https://doi.org/10.1176/appi.ps.53.10.1322)] [Medline: [12364686](https://pubmed.ncbi.nlm.nih.gov/12364686/)]
36. Marson DC, Hawkins L, McInturff B, Harrell LE. Cognitive models that predict physician judgments of capacity to consent in mild Alzheimer's disease. *J Am Geriatr Soc* 1997 Apr;45(4):458-464. [doi: [10.1111/j.1532-5415.1997.tb05171.x](https://doi.org/10.1111/j.1532-5415.1997.tb05171.x)] [Medline: [9100715](https://pubmed.ncbi.nlm.nih.gov/9100715/)]
37. Zinn S, Dudley TK, Bosworth HB, Hoenig HM, Duncan PW, Horner RD. The effect of poststroke cognitive impairment on rehabilitation process and functional outcome. *Arch Phys Med Rehabil* 2004 Jul;85(7):1084-1090. [doi: [10.1016/j.apmr.2003.10.022](https://doi.org/10.1016/j.apmr.2003.10.022)] [Medline: [15241754](https://pubmed.ncbi.nlm.nih.gov/15241754/)]
38. Heruti RJ, Lusky A, Dankner R, et al. Rehabilitation outcome of elderly patients after a first stroke: effect of cognitive status at admission on the functional outcome. *Arch Phys Med Rehabil* 2002 Jun;83(6):742-749. [doi: [10.1053/apmr.2002.32739](https://doi.org/10.1053/apmr.2002.32739)] [Medline: [12048650](https://pubmed.ncbi.nlm.nih.gov/12048650/)]
39. Burdea GC. Virtual rehabilitation--benefits and challenges. *Methods Inf Med* 2003;42(5):519-523. [doi: [10.1055/s-0038-1634378](https://doi.org/10.1055/s-0038-1634378)] [Medline: [14654886](https://pubmed.ncbi.nlm.nih.gov/14654886/)]
40. 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](https://doi.org/10.1111/j.1532-5415.1992.tb01992.x)] [Medline: [1512391](https://pubmed.ncbi.nlm.nih.gov/1512391/)]
41. Cannell J, Jovic E, Rathjen A, et al. The efficacy of interactive, motion capture-based rehabilitation on functional outcomes in an inpatient stroke population: a randomized controlled trial. *Clin Rehabil* 2018 Feb;32(2):191-200. [doi: [10.1177/0269215517720790](https://doi.org/10.1177/0269215517720790)] [Medline: [28719977](https://pubmed.ncbi.nlm.nih.gov/28719977/)]
42. Saposnik G, Cohen LG, Mamdani M, et al. Efficacy and safety of non-immersive virtual reality exercising in stroke rehabilitation (EVREST): a randomised, multicentre, single-blind, controlled trial. *Lancet Neurol* 2016 Sep;15(10):1019-1027. [doi: [10.1016/S1474-4422\(16\)30121-1](https://doi.org/10.1016/S1474-4422(16)30121-1)] [Medline: [27365261](https://pubmed.ncbi.nlm.nih.gov/27365261/)]
43. Candel M, van Breukelen GJP. Best (but oft forgotten) practices: efficient sample sizes for commonly used trial designs. *Am J Clin Nutr* 2023 Jun;117(6):1063-1085. [doi: [10.1016/j.ajcnut.2023.02.013](https://doi.org/10.1016/j.ajcnut.2023.02.013)] [Medline: [37270287](https://pubmed.ncbi.nlm.nih.gov/37270287/)]
44. Robinson J, Dixon J, Macsween A, van Schaik P, Martin D. The effects of exergaming on balance, gait, technology acceptance and flow experience in people with multiple sclerosis: a randomized controlled trial. *BMC Sports Sci Med Rehabil* 2015;7:8. [doi: [10.1186/s13102-015-0001-1](https://doi.org/10.1186/s13102-015-0001-1)] [Medline: [25969739](https://pubmed.ncbi.nlm.nih.gov/25969739/)]
45. Totton N, Lin J, Julious S, Chowdhury M, Brand A. A review of sample sizes for UK pilot and feasibility studies on the ISRCTN registry from 2013 to 2020. *Pilot Feasibility Stud* 2023 Nov 21;9(1):188. [doi: [10.1186/s40814-023-01416-w](https://doi.org/10.1186/s40814-023-01416-w)] [Medline: [37990337](https://pubmed.ncbi.nlm.nih.gov/37990337/)]
46. Arlati S, Mondellini M, Guanziroli E, Rossini M, Martinelli I, Molteni F. Virtual reality-based rehabilitation for patients with stroke: preliminary results on user experience. In: De Paolis LT, Arpaia P, Sacco M, editors. *Extended Reality: International Conference, XR Salento 2023, Lecce, Italy, September 6-9, 2023, Proceedings, Part II*:441-453. [doi: [10.1007/978-3-031-43404-4_30](https://doi.org/10.1007/978-3-031-43404-4_30)]
47. XR ray interactor. Unity. URL: <https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.0/manual/xr-ray-interactor.html> [accessed 2026-02-25]

48. Thielman GT, Dean CM, Gentile AM. Rehabilitation of reaching after stroke: task-related training versus progressive resistive exercise. *Arch Phys Med Rehabil* 2004 Oct;85(10):1613-1618. [doi: [10.1016/j.apmr.2004.01.028](https://doi.org/10.1016/j.apmr.2004.01.028)] [Medline: [15468020](https://pubmed.ncbi.nlm.nih.gov/15468020/)]
49. Jackson SA, Martin AJ, Eklund RC. Long and short measures of flow: the construct validity of the FSS-2, DFS-2, and new brief counterparts. *J Sport Exerc Psychol* 2008 Oct;30(5):561-587. [doi: [10.1123/jsep.30.5.561](https://doi.org/10.1123/jsep.30.5.561)] [Medline: [18971512](https://pubmed.ncbi.nlm.nih.gov/18971512/)]
50. Lessiter J, Freeman J, Keogh E, Davidoff J. A cross-media presence questionnaire: the ITC-sense of presence inventory. *Presence (Camb)* 2001 Jun;10(3):282-297. [doi: [10.1162/105474601300343612](https://doi.org/10.1162/105474601300343612)]
51. Thompson ER. Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *J Cross Cult Psychol* 2007 Mar;38(2):227-242. [doi: [10.1177/0022022106297301](https://doi.org/10.1177/0022022106297301)]
52. 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](https://doi.org/10.1207/s15327108ijap0303_3)]
53. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. *Decis Sci* 2008 May;39(2):273-315. [doi: [10.1111/j.1540-5915.2008.00192.x](https://doi.org/10.1111/j.1540-5915.2008.00192.x)]
54. Magni E, Binetti G, Bianchetti A, Rozzini R, Trabucchi M. Mini-mental state examination: a normative study in Italian elderly population. *Eur J Neurol* 1996 May;3(3):198-202. [doi: [10.1111/j.1468-1331.1996.tb00423.x](https://doi.org/10.1111/j.1468-1331.1996.tb00423.x)] [Medline: [21284770](https://pubmed.ncbi.nlm.nih.gov/21284770/)]
55. Langley GB, Sheppard H. The visual analogue scale: its use in pain measurement. *Rheumatol Int* 1985;5(4):145-148. [doi: [10.1007/BF00541514](https://doi.org/10.1007/BF00541514)] [Medline: [4048757](https://pubmed.ncbi.nlm.nih.gov/4048757/)]
56. Jönsson AC, Lindgren I, Hallström B, Norrving B, Lindgren A. Prevalence and intensity of pain after stroke: a population based study focusing on patients' perspectives. *J Neurol Neurosurg Psychiatry* 2006 May;77(5):590-595. [doi: [10.1136/jnnp.2005.079145](https://doi.org/10.1136/jnnp.2005.079145)] [Medline: [16354737](https://pubmed.ncbi.nlm.nih.gov/16354737/)]
57. Bohannon RW. Motricity index scores are valid indicators of paretic upper extremity strength following stroke. *J Phys Ther Sci* 1999;11(2):59-61. [doi: [10.1589/jpts.11.59](https://doi.org/10.1589/jpts.11.59)]
58. Gregson JM, Leathley M, Moore AP, Sharma AK, Smith TL, Watkins CL. Reliability of the Tone Assessment Scale and the modified Ashworth scale as clinical tools for assessing poststroke spasticity. *Arch Phys Med Rehabil* 1999 Sep;80(9):1013-1016. [doi: [10.1016/s0003-9993\(99\)90053-9](https://doi.org/10.1016/s0003-9993(99)90053-9)] [Medline: [10489001](https://pubmed.ncbi.nlm.nih.gov/10489001/)]
59. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. *Am J Occup Ther* 1985 Jun;39(6):386-391. [doi: [10.5014/ajot.39.6.386](https://doi.org/10.5014/ajot.39.6.386)] [Medline: [3160243](https://pubmed.ncbi.nlm.nih.gov/3160243/)]
60. Berg K, Wood-Dauphine S, Williams JL, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can* 1989 Nov;41(6):304-311. [doi: [10.3138/ptc.41.6.304](https://doi.org/10.3138/ptc.41.6.304)]
61. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991 Feb;39(2):142-148. [doi: [10.1111/j.1532-5415.1991.tb01616.x](https://doi.org/10.1111/j.1532-5415.1991.tb01616.x)] [Medline: [1991946](https://pubmed.ncbi.nlm.nih.gov/1991946/)]
62. Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 2009 Mar;74(1):107-120. [doi: [10.1007/s11336-008-9101-0](https://doi.org/10.1007/s11336-008-9101-0)] [Medline: [20037639](https://pubmed.ncbi.nlm.nih.gov/20037639/)]
63. Dunn TJ, Baguley T, Brunsden V. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *Br J Psychol* 2014 Aug;105(3):399-412. [doi: [10.1111/bjop.12046](https://doi.org/10.1111/bjop.12046)] [Medline: [24844115](https://pubmed.ncbi.nlm.nih.gov/24844115/)]
64. Chen J, Or CK, Li Z, Yeung EHK, Chen T. Perceptions of patients with stroke regarding an immersive virtual reality-based exercise system for upper limb rehabilitation: questionnaire and interview study. *JMIR Serious Games* 2025 Jan 1;13:e49847. [doi: [10.2196/49847](https://doi.org/10.2196/49847)] [Medline: [39742513](https://pubmed.ncbi.nlm.nih.gov/39742513/)]
65. Gustavsson M, Kjörk EK, Erhardsson M, Alt Murphy M. Virtual reality gaming in rehabilitation after stroke - user experiences and perceptions. *Disabil Rehabil* 2022 Nov;44(22):6759-6765. [doi: [10.1080/09638288.2021.1972351](https://doi.org/10.1080/09638288.2021.1972351)] [Medline: [34465269](https://pubmed.ncbi.nlm.nih.gov/34465269/)]
66. Mondellini M, Arlati S, Gapeyeva H, et al. User experience during an immersive virtual reality-based cognitive task: a comparison between Estonian and Italian older adults with MCI. *Sensors (Basel)* 2022 Oct 27;22(21):8249. [doi: [10.3390/s22218249](https://doi.org/10.3390/s22218249)] [Medline: [36365947](https://pubmed.ncbi.nlm.nih.gov/36365947/)]
67. Stanney KM, Kennedy RS, Drexler JM. Cybersickness is not simulator sickness. *Proc Hum Factors Ergonomics Soc Annu Meeting* 1997 Oct;41(2):1138-1142. [doi: [10.1177/107118139704100292](https://doi.org/10.1177/107118139704100292)]
68. Gangwani R, Cain A, Collins A, Cassidy JM. Leveraging factors of self-efficacy and motivation to optimize stroke recovery. *Front Neurol* 2022;13:823202. [doi: [10.3389/fneur.2022.823202](https://doi.org/10.3389/fneur.2022.823202)] [Medline: [35280288](https://pubmed.ncbi.nlm.nih.gov/35280288/)]
69. Kim YM, Rhiu I, Yun MH. A systematic review of a virtual reality system from the perspective of user experience. *Int J Hum Comput Interact* 2020 Jun 14;36(10):893-910. [doi: [10.1080/10447318.2019.1699746](https://doi.org/10.1080/10447318.2019.1699746)]
70. Mondellini M, Colombo V, Arlati S, Lawson G, Cobb S. Human factors and ergonomics. In: Alcañiz M, Sacco M, Tromp JG, editors. *Roadmapping Extended Reality: Fundamentals and Applications*: Wiley Online Library; 2022:229-256. [doi: [10.1002/9781119865810](https://doi.org/10.1002/9781119865810)]
71. Papagiannis G, Triantafyllou A, Yiannopoulou KG, et al. Hand dexterities assessment in stroke patients based on augmented reality and machine learning through a box and block test. *Sci Rep* 2024 May 8;14(1):10598. [doi: [10.1038/s41598-024-61070-x](https://doi.org/10.1038/s41598-024-61070-x)] [Medline: [38719940](https://pubmed.ncbi.nlm.nih.gov/38719940/)]
72. Darcy B, Rashford L, Huizenga D, Reed KB, Bamberg SJM. Balance improvement and fall risk reduction in stroke survivors after treatment with a wearable home-use gait device: single-arm longitudinal study with 1-year follow-up. *JMIR Form Res* 2025 Aug 15;9:e67297. [doi: [10.2196/67297](https://doi.org/10.2196/67297)] [Medline: [40815786](https://pubmed.ncbi.nlm.nih.gov/40815786/)]

73. Sakai K, Hosoi Y, Harada Y, Kato Y. Association between the executive dysfunction and balance function in patients with stroke. *Brain Behav* 2024 May;14(5):e3542. [doi: [10.1002/brb3.3542](https://doi.org/10.1002/brb3.3542)] [Medline: [38783598](https://pubmed.ncbi.nlm.nih.gov/38783598/)]
74. Rose Sin Yi L, Jing Jing S, Hammada AO, Jonathan B, Ladislav B, Jing Q. Effects of virtual reality-based cognitive interventions on cognitive function and activity of daily living among stroke patients: systematic review and meta-analysis. *J Clin Nurs* 2024 Mar;33(3):1169-1184. [doi: [10.1111/jocn.16986](https://doi.org/10.1111/jocn.16986)] [Medline: [38234275](https://pubmed.ncbi.nlm.nih.gov/38234275/)]
75. Josman N, Kizony R, Hof E, Goldenberg K, Weiss PL, Klinger E. Using the virtual action planning-supermarket for evaluating executive functions in people with stroke. *J Stroke Cerebrovasc Dis* 2014;23(5):879-887. [doi: [10.1016/j.jstrokecerebrovasdis.2013.07.013](https://doi.org/10.1016/j.jstrokecerebrovasdis.2013.07.013)] [Medline: [24008131](https://pubmed.ncbi.nlm.nih.gov/24008131/)]
76. Halford E, Jakubiszak S, Krug K, Umphress A. What is task-oriented training? A scoping review. *Student J Occup Ther* 2024;4(1):1-23. [doi: [10.46409/001.DPYW4980](https://doi.org/10.46409/001.DPYW4980)]
77. Yan W, Lin Y, Chen YF, Wang Y, Wang J, Zhang M. Enhancing neuroplasticity for post-stroke motor recovery: mechanisms, models, and neurotechnology. *IEEE Trans Neural Syst Rehabil Eng* 2025;33:1156-1168. [doi: [10.1109/TNSRE.2025.3551753](https://doi.org/10.1109/TNSRE.2025.3551753)] [Medline: [40100694](https://pubmed.ncbi.nlm.nih.gov/40100694/)]
78. Smyth C, Broderick P, Lynch P, Clark H, Monaghan K. To assess the effects of cross-education on strength and motor function in post stroke rehabilitation: a systematic literature review and meta-analysis. *Physiotherapy* 2023 Jun;119:80-88. [doi: [10.1016/j.physio.2023.02.001](https://doi.org/10.1016/j.physio.2023.02.001)] [Medline: [36940490](https://pubmed.ncbi.nlm.nih.gov/36940490/)]
79. Ausenda C, Carnovali M. Transfer of motor skill learning from the healthy hand to the paretic hand in stroke patients: a randomized controlled trial. *Eur J Phys Rehabil Med* 2011 Sep;47(3):417-425. [Medline: [21555982](https://pubmed.ncbi.nlm.nih.gov/21555982/)]

Abbreviations

ITC-SOPI: International Test Commission—Sense of Presence Inventory

MI: Motricity Index

MMSE: Mini-Mental State Examination

PANAS: Positive and Negative Affect Schedule

PEOU: Perceived Ease of Use

RATE-XR: Rationale, Accessibility, Training, and Effectiveness for Extended Reality

SFS: Short Flow Scale

SP: spatial presence

SSQ: Simulator Sickness Questionnaire

TAM3: Technology Acceptance Model

TS: total score

TUG: Timed Up and Go

VR: virtual reality

VSS: Virtual Supermarket for Stroke

Edited by S Brini; submitted 24.Jun.2025; peer-reviewed by A Higaki, C Abbatantuono; accepted 29.Dec.2025; published 11.Mar.2026.

Please cite as:

Arlati S, Mondellini M, Martinelli I, Guanziroli E, Rossini M, Sacco M, Molteni F

Feasibility and User Experience of Immersive Virtual Reality-Based Rehabilitation in Patients With Stroke: Single-Arm Pretest-Posttest Pilot Study

JMIR Serious Games 2026;14:e79584

URL: <https://games.jmir.org/2026/1/e79584>

doi: [10.2196/79584](https://doi.org/10.2196/79584)

© Sara Arlati, Marta Mondellini, Isabella Martinelli, Eleonora Guanziroli, Mauro Rossini, Marco Sacco, Franco Molteni. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 11.Mar.2026. 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.

Home-Based Immersive Virtual Reality to Improve Motor Performance in Children and Adolescents With Developmental Coordination Disorder: Crossover Study

Mohammed Alharbi^{1,2}, PT, MSc; David Harris¹, PhD; Helen Dodd¹, PhD; Greg Wood³, PhD; Gavin Buckingham¹, PhD

¹Department of Public Health and Sport Sciences, University of Exeter Medical School, University of Exeter, St. Luke's Campus, Heavitree Road, Exeter, United Kingdom

²Department of Physical Therapy, Faculty of Applied Medical Sciences, University of Hail, Hail, Saudi Arabia

³Department of Sport and Exercise Sciences, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, United Kingdom

Corresponding Author:

Mohammed Alharbi, PT, MSc

Department of Public Health and Sport Sciences, University of Exeter Medical School, University of Exeter, St. Luke's Campus, Heavitree Road, Exeter, United Kingdom

Abstract

Background: Children with developmental coordination disorder (DCD) experience motor difficulties that limit daily activities and reduce physical activity enjoyment. Immersive virtual reality (VR) offers the potential for feedback-rich movement practice, but evidence for these effects in DCD remains limited.

Objective: This study aims to investigate the effects of an immersive VR rhythm game compared to tablet-based gameplay within a home-based setting on motor performance, enjoyment, and motivation in children and adolescents with DCD.

Methods: This crossover study included 27 participants (21 boys and 6 girls) aged 10 to 16 years with DCD who completed 2 home-based interventions, each delivered over 5 consecutive days: VR gameplay using Beat Saber (Beat Games) and tablet-based gameplay using Cut the Rope (ZeptoLab). Participants were recruited in England using convenience sampling via social media; eligible participants were aged 10 to 16 years, met research criteria for DCD based on a Developmental Coordination Disorder Questionnaire screening, and had no alternative neurological or musculoskeletal diagnosis. Each condition required at least 30 minutes of daily gameplay and was separated by a 2-week or more washout period. Motor performance was assessed pre- and postintervention using the *Movement Assessment Battery for Children, Third Edition* (MABC-3) and the Box and Block Test. Enjoyment was measured pre- and postintervention using the Physical Activity Enjoyment Scale. Participants also rated their motivation and feelings during each gameplay session. Repeated-measures ANOVAs and paired-samples 2-tailed *t* tests ($\alpha=.05$) were used to examine the data.

Results: For MABC-3 domains, condition \times time interactions were nonsignificant, although exploratory within-condition analyses showed pre-post improvements in the VR condition. For the Box and Block Test, condition \times time interactions were significant for both hands, with a greater degree of pre-post improvement in VR than tablet-based gameplay. In the VR condition, mean block transfer increased for the dominant hand (ΔM 5.93, 95% CI 3.49 - 8.36; $t_{26}=-4.99$; $P<.001$; Cohen $d=0.96$) and the nondominant hand (ΔM 5.11, 95% CI 2.65 - 7.57; $t_{26}=-4.27$; $P<.001$; Cohen $d=0.82$), whereas no significant changes were observed in the tablet condition (all $P\geq.36$). VR gameplay also yielded higher enjoyment, and children reported higher motivation and feeling ratings across VR sessions than tablet sessions.

Conclusions: This study, in contrast to most of the existing literature on nonimmersive technologies, examined the effect of a home-based *immersive* VR rhythm game intervention for children and adolescents with DCD. This study provides early evidence that this VR rhythm game was engaging and may support greater short-term improvements in motor performance than tablet gameplay. These findings suggest home-based immersive VR rhythm games could be a practical adjunct to pediatric rehabilitation to increase movement practice and motivation, although larger and longer trials are needed to confirm clinical impact and identify which game features drive this benefit.

(*JMIR Serious Games* 2026;14:e84995) doi:[10.2196/84995](https://doi.org/10.2196/84995)

KEYWORDS

developmental coordination disorder; DCD; game; Movement Assessment Battery for Children, Third Edition; MABC-3; pediatric; virtual reality; VR; extended reality; XR

Introduction

Background

Developmental coordination disorder (DCD), or dyspraxia, is a neurodevelopmental condition characterized by marked impairment in the acquisition and execution of coordinated motor skills that interferes with daily living and academic achievement [1,2]. Motor coordination describes a range of fundamental movement skills such as object manipulation (eg, grasping, kicking, and throwing) and locomotor activities (eg, walking, running, and jumping) [1,2]. It is an essential component of human development, supporting everyday self-care, communication, and active engagement in learning environments [1,2]. These motor skills support physical play and enable independent completion of activities of daily living [2]. For school-aged children, motor tasks represent a substantial portion of everyday academic activities [3]. Approximately 46% of children's time spent at school involves participation in tasks requiring motor coordination, such as handwriting, coloring, cutting with scissors, and using tools [3]. Adequate motor coordination supports physical and cognitive development, psychosocial health, and overall well-being [4]. Deficits in motor skills could restrict participation, reduce quality of life, impair confidence, and adversely affect developmental outcomes [1].

Difficulties in motor coordination often appear in early childhood, typically around preschool age [5]. These difficulties become more evident in primary school as motor demands increase in daily routines [6]. DCD affects about 5% to 6% of school-aged children who often struggle with everyday tasks needed for home, school, and play [2,7]. Common difficulties include handwriting, dressing, tying shoelaces, brushing teeth, and taking part in sports and leisure activities [8]. As a result, participation is often less frequent, and children with DCD engage in a narrower range of activities than their typically developing peers [9]. These constraints affect current motor performance and broader psychological, emotional, and social development [1]. Early identification and effective intervention are therefore needed [10].

In recent years, growing research interest has emerged regarding innovative and engaging approaches aimed at enhancing motor performance and physical activities among children [11]. Immersive Virtual Reality (VR), characterized by an interactive computer-generated environment typically explored through full-body movements, has rapidly gained prominence within the consumer entertainment sector [12]. In this paper, we use the term immersive virtual reality (VR) to refer to computer-generated, 3D environments that surround the user and are typically accessed via head-mounted displays or comparable 3D displays, together with motion- or spatial-tracking technologies (eg, handheld controllers, optical tracking of hand and body parts, or other sensor-based interfaces) that enable active interaction with virtual objects and spaces [12]. In rehabilitation, VR offers a motivating, realistic setting with immediate and often intrinsic feedback [13,14], potentially increasing adherence to treatment programs [15] and facilitating the repetition of targeted motor tasks

[16,17]. For children, motivation and enjoyment are key to sustaining activity [18]. VR-based games could stimulate intrinsic motivation and enjoyment, making rehabilitation feel playful and rewarding, encouraging persistent practice [19]. Studies report benefits of VR for skill acquisition in typically developing children [20] and in children with acquired brain injury [21], autism spectrum disorder [22], and cerebral palsy [17], but evidence on immersive VR for children and adolescents with DCD is scarce. A recent scoping review by Alharbi et al [23] found no studies using fully immersive VR systems in this population, indicating a clear research gap.

Exercise-based VR rhythm games, such as Beat Saber, involve physically interactive gameplay synchronized with rhythmic auditory and visual stimuli [24]. Beat Saber has a growing evidence base in motor control research, demonstrating improvements in reaction time and hand-eye coordination in typically developing youth populations [25-27]. In a recent Patient and Public Involvement (PPI) activity, we conducted a focus group where children and adolescents with DCD experienced various VR sports and rhythm games, including Beat Saber. Contributors in this activity identified Beat Saber as especially engaging because of its music movement coupling and simple controls. This rhythm game requires bimanual, goal-directed arm movements, precise timing, and rapid visuomotor responses, with graded difficulty that enables high-dose practice of coordinated actions. These features suggest that VR-based rhythm games may benefit motor skills in children and adolescents with DCD. Accordingly, the primary aim of this study was to investigate the impact of a home-based immersive VR rhythm game on motor performance, motivation, and enjoyment in children and adolescents diagnosed with or suspected of having DCD.

PPI Activity

This study was informed by a recent PPI activity that explored the perspectives of children and adolescents with DCD on sports participation and immersive VR games [28]. Five participants aged 13 to 16 years who self-reported as having DCD took part in this study.

The PPI activity had 2 visits. The first visit involved a focus group where contributors discussed experiences and challenges in relation to traditional sports and manual tasks and commented on promotional videos of various VR sports and rhythm-based games. In the second visit, contributors took part in a hands-on VR session, followed by brief semistructured interviews about usability, enjoyment, perceived difficulty, any adverse symptoms (eg, motion sickness), content preferences, and perceived feasibility of VR for home use and rehabilitation. Contributors played VR sports titles (eg, Rezzil Index football drills, table tennis, tennis, racket sports, volleyball, and golf) and the rhythm game Beat Saber using a Valve Index VR headset.

Contributors shared a variety of opinions about the immersive VR games. Some described the games as enjoyable and accessible; others found them less engaging because of their lack of dynamic interactions and player movement. These responses suggested that engagement with immersive VR games may vary depending on individual preference and perceived challenge. Beat Saber, a rhythm-based game included in the

session, was described as especially engaging and motivating. Contributors highlighted that they liked the combination of music, movement, and straightforward controls, which made the activity physically demanding yet manageable. The clear rhythmic structure and lack of competitive pressure were described as supporting enjoyment and confidence.

Insights from this PPI activity directly informed the study design, guiding the selection of this rhythm-based game as the primary VR task and supporting a home-based protocol with graded difficulty and repeated short sessions.

Methods

Participants

We recruited for a fixed time period of 6 months from September 2024 to March 2025; within the available recruitment period, 30 children and adolescents, aged 10 to 16 years (23 male and 7 female participants), with confirmed or suspected DCD participated in the study. With this sample size, based on an a priori power analysis conducted in G*Power (version 3.1.9.6; Heinrich Heine University Düsseldorf), using a 2-dependent-means 2-tailed *t* test with a desired power of 0.80 ($\alpha=.05$), we would be able to detect a medium effect size ($d_z=.53$). We selected the age range of 10 to 16 years on developmental, clinical, and practical grounds. Ten years was chosen as the lower bound because by around the age of 10 years, most children are in middle childhood, with sufficiently developed concrete reasoning, sustained attention, and independent reading skills to participate in structured, school-like tasks [29]. So, children of this age typically have sufficient cognitive, attentional, and reading abilities to follow a home-based protocol, complete self-report measures, and use an immersive VR headset safely. From a safety and feasibility perspective, we did not include children younger than 10 years because Tychsen and Foeller [30] found that those aged 4 to 10 years showed increased general discomfort, head-neck discomfort, fatigue, and visually induced motion sickness on the child-friendly Simulator Sickness Questionnaire measures after VR exposure compared with their pre-VR ratings. An age of 16 years was chosen as the upper bound to capture mid- to late-adolescence while still focusing on school-aged young people who are typically managed within pediatric and educational services for DCD.

Participants were recruited from England, United Kingdom, using convenience sampling via social media platforms including X (formerly known as Twitter), Facebook, and Instagram. Because recruitment occurred via open social-media platforms, the number of individuals reaching or viewing the advert is unknown, and a response rate cannot be calculated. Eligibility screening was performed using the Developmental Coordination Disorder Questionnaire (DCDQ) [31], which was completed by participants' parents or guardians. The DCDQ is a validated screening instrument designed to assess children's motor coordination capabilities in those aged approximately 5 to 15 years [32]. Inclusion required DCDQ scores of 57 or less on the 15-item DCDQ (total 15 - 75; lower scores indicate greater difficulty); 58 to 75 is interpreted as not indicative of DCD. For adolescents aged 16 years, DCDQ scores were used

as supportive screening information alongside baseline motor assessment rather than as a diagnostic indicator to determine participation in our study, and we recognize that using the DCDQ slightly beyond its validated age range may reduce screening precision for the oldest participants. Exclusion criteria were children with co-occurring neurological or musculoskeletal conditions affecting motor or cognitive function (eg, muscular dystrophy and cerebral palsy), or children who had undergone recent surgery or suffered a severe physical injury. Participants who already owned a home VR headset were excluded to reduce variability due to differences in prior exposure, unequal device familiarity, and uncontrolled use outside the prescribed days (during 2 wk or washout period of the tablet). Families meeting the inclusion criteria were invited to the research laboratory for participation.

Three participants withdrew during the interventions (2 male and 1 female participants): one during the VR condition and two during the tablet condition. Reasons for withdrawal included difficulties managing study participation alongside school commitments, disinterest (in tablet-based activities), or a lack of motivation to continue participation. Analyses were conducted on participants who completed both conditions ($n=27$; mean 12.7, SD 2.07; 21 male and 6 female participants). Data from withdrawals were excluded. Among these 27 participants, the baseline of the *Movement Assessment Battery for Children, Third Edition* (MABC-3) data indicated that 25 scored at ≤ 16 th percentile on the total motor score and, together with DCDQ ≤ 57 and the exclusion of alternative neurological or musculoskeletal conditions, were classified as meeting research criteria consistent with the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5) for DCD. The remaining 2 participants had the MABC-3 scores in the 16th to 37th percentile range, but DCDQ < 57 scores and functional motor difficulties and are therefore considered as having probable DCD. A 16th percentile or less was used as the cutoff for the total motor score in line with international clinical practice recommendations for DCD when using the MABC-3 (or equivalent objective motor measures) [33]. For brevity, we refer to this combined group as "children and adolescents with DCD" throughout the paper.

Ethical Considerations

Ethical approval was obtained from the local research ethics board at the University of Exeter (reference: 5759511). Written informed consent was obtained from the participants' parents or guardians, and all participating children provided their assent prior to the initiation of the study procedures. The consent or assent process made clear that the participation was voluntary and that families could withdraw at any time without giving a reason. To protect privacy and confidentiality, each participant was assigned a unique study ID, and all questionnaires and assessments were stored in deidentified form. Identifiable information (eg, names and contact details) was stored separately from research data on secure, password-protected university servers with access restricted to authorized members of the research team. Only anonymized data are reported herein. Participants received financial compensation of £50 (US \$67.66) as a fixed reimbursement for their time and any travel costs. No

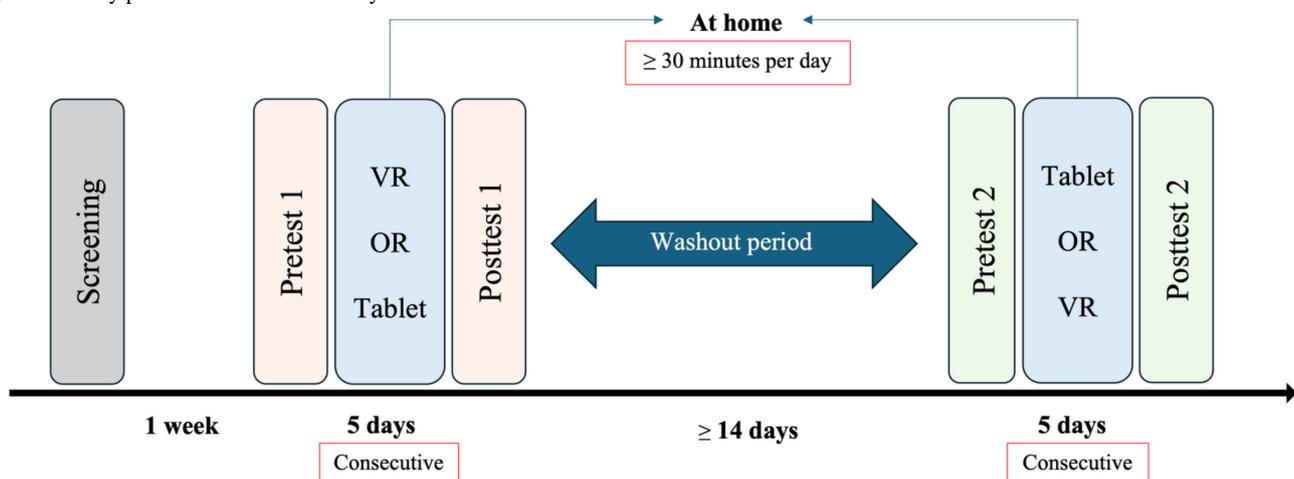
images in the study or supplementary materials include identifiable participants.

Study Design and Experimental Procedure

This study used a crossover design. Each participant completed two 1-week interventions: a tablet-based game (tablet condition) and Beat Saber on a VR headset (VR condition). A 5-day intervention period was chosen primarily for comparability with previous immersive VR studies using Beat Saber, which has demonstrated meaningful changes over a similar short, intensive timeframe. In particular, Grosprêtre et al [26] and Rutkowski et al [25] used Beat Saber-based protocols of less than 1 week and reported measurable improvements, suggesting that this duration is sufficient to elicit short-term effects while remaining acceptable and feasible for children and families. The order of conditions was counterbalanced by simple alternation based on

enrollment: the first enrolled participant was assigned to the tablet condition first, the second to VR first, and so on. The interventions were separated by a washout period of at least 14 days. Participants were instructed to play a minimum of 30 minutes per day for 5 consecutive days at home (Figure 1). A minimum of 30 minutes per day for 5 consecutive days was chosen to provide a dose that clearly exceeds the minimum effective exposure reported in previous Beat Saber VR studies [25,26], where participants played for around 15 minutes per day over 5 days and still showed improvements. By doubling the daily playtime, we aimed to increase the volume of motor practice and the likelihood of detecting short-term effects. Adherence was monitored with a daily log completed after each session; entries recorded session completion and approximate duration. No device logs were used, so adherence reflects self-report.

Figure 1. Study protocol. VR: virtual reality.



Assessments were conducted both before (pre) and after (post) each intervention using the MABC-3, the Box and Block Test (BBT), and the Physical Activity Enjoyment Scale (PACES). After each gameplay session, participants were asked to report their motivation and feeling ratings in the daily log. After the postassessment, the researcher collected the device, erased participant data, and reset the game application for the next participant. Blinding was not possible because participants and caregivers were aware of the intervention being used (VR vs tablet); outcome assessments were conducted by the same assessor who was not blinded to the condition. All outcome assessments were administered using standardized instructions by a trained assessor, using the same equipment and testing environment at each visit to reduce measurement variability. This study is reported in accordance with the APA JARS-Quant (Journal Article Reporting Standards for Quantitative Research) guidelines (Checklist 1) [34].

Interventions

VR Gameplay Condition

The VR condition used Meta Quest 2 headsets (Meta Platforms, Inc; 1832×1920 pixels per eye; 120 Hz) with 2 handheld controllers. Participants played Beat Saber (version 1.40.7_7060; Beat Games), a rhythm-based VR game in which colored cubes approach in sync with music. Players use 2 virtual sabers, one

per hand, to slice cubes in the specified direction with the correct hand and in time to the beat, and must avoid obstacles. Scoring is based on the number of accurately sliced cubes, the precision of the slicing motion, and alignment with the rhythm of the soundtrack. Participants were instructed to begin playing the game on the default (“normal”) difficulty setting. They were further informed that if the “normal” difficulty was too challenging, they could switch to the “easy” mode. Conversely, participants who found the “normal” mode insufficiently challenging were permitted to progress to the “hard” mode. The difficulty level was not systematically tracked during the intervention. This game was considered to be a promising intervention due to its emphasis on whole-body motor coordination, bilateral upper-limb control, visuomotor integration, and rhythmic timing [25,26], skills that are commonly affected in children with DCD [7]. Prior studies have highlighted this game as an engaging platform for motor training, offering real-time feedback and adaptive difficulty levels [35,36].

Participants completed the VR gameplay at home under caregiver supervision, in accordance with the safety instructions provided. Each caregiver received a detailed guidance sheet outlining safety protocols, including monitoring for motion sickness, ensuring a clear play area, and being present during all gameplay sessions (VR instructions can be located online)

[37]. Prior to the VR intervention, each participant received a brief orientation on how to use the VR headset and interact with the game. Adherence was monitored through self-report logs, in which participants or caregivers recorded daily gameplay duration and any issues encountered. Any adverse effects, such as motion sickness or fatigue, were also noted. No such incidents were reported.

Tablet Gameplay Condition

The tablet condition used a Lenovo Tab M8 (fourth generation) Android tablet with an 8-inch LCD (1280×800) display and dimensions of 197.97×119.82×8.95 mm. Participants played *Cut the Rope* (version 3.75.0; Zeptolab), a puzzle game that requires cutting virtual ropes in a planned sequence to deliver candy to the character Om Nom. The game was administered using the default settings. In this mode, levels are presented sequentially, with successful completion of a level automatically unlocking the next one. Puzzle complexity increases gradually (eg, greater numbers of ropes, obstacles, and timing constraints), but there is no separate user-adjustable difficulty setting or fixed level cap relevant to the amount of play time in this study. Players can repeat the same level as many times as needed until it is solved. Therefore, if they become stuck, progression may slow, and they may remain on earlier levels for longer. We selected this game as an active control that emphasizes problem-solving with minimal motor demands and has been used as a control in DCD research [38]. The game is self-paced and does not include rhythm- or music-based timing mechanics, and actions are not performed in synchrony with a beat or external tempo.

As with the VR condition, gameplay for the tablet condition occurred at home, and participants were instructed to play for a minimum of 30 minutes daily across 5 consecutive days. Prior to the tablet intervention, each participant received a brief orientation on how to navigate and interact with the game. Caregivers were not asked to supervise sessions. Participants were given the tablet directly and expected to engage with the game independently. Engagement and adherence were recorded in daily logs. The application was reset after each participant to protect data privacy.

In both conditions, children and parents were asked to “Play at least 30 minutes per day,” but were explicitly told there were no “right” or “wrong” amounts of play, that missing or shortening sessions were acceptable if they were tired, busy, or unwell, and that honest recording was more important than appearing highly adherent. They were also reminded that logs were for research only, anonymized, and not shared with anyone else.

Outcome Measurements

Movement Assessment (MABC-3)

The primary outcome was performance on the MABC-3, a standardized, norm-referenced assessment of motor coordination for ages 3 to 25 years [39]. The MABC is widely used in clinical and research settings to identify motor difficulties and monitor development [40]. The MABC-3 includes 10 subtests across 3 domains: manual dexterity, aiming and catching, and balance and locomotion. Tasks were administered by age band per the

manual (7 - 11 y=band 2; 12 - 25 y=band 3). Trained assessors followed the standardized protocol. Testing took place in a quiet laboratory room using standard MABC-3 equipment as specified in the manual. Raw scores from each item were converted to standard scores using age-referenced normative data. A total motor score was computed by summing the standard scores from each domain, providing an overall estimate of motor proficiency. The MABC-3 scores are interpreted using percentile ranks. Consistent with published guidelines, scores above the 16th percentile are considered indicative of typical motor performance; scores between the 6th and 15th percentile suggest a risk of motor coordination difficulties, and scores at or below the 5th percentile are classified as indicative of significant motor coordination impairment. Although MABC-3 is a newer version of the widely validated *Movement Assessment Battery for Children, Second Edition* (MABC-2), both editions have demonstrated strong psychometric properties, including high test-retest reliability and construct validity for use in pediatric populations [39]. Its application in this study was chosen due to its relevance for identifying motor coordination challenges commonly observed in children with DCD.

Upper-Limb Manual Dexterity Assessment (BBT)

The secondary outcome was performance on the BBT, a standardized measure of unilateral gross manual dexterity [41]. The apparatus is a wooden box with 2 equal compartments (53.7×25.4×8.5 cm), a central partition, and 150 wooden blocks with sides measuring 2.5 cm [42]. Participants are asked to transfer as many blocks as possible, one at a time, from one compartment to the other within 60 seconds using a single hand.

Administration followed the standardized guidelines described by Mathiowetz et al [42]. Each hand received a 15-second practice trial, followed by one 60-second timed trial. Testing was completed for the dominant hand first, then the nondominant hand, at each pre- and postsession. Hand dominance was determined by self-report. Scores were the number of blocks transferred per hand; if multiple blocks were moved at once, only one was counted. The BBT shows high test-retest reliability in pediatric samples [43,44]. The BBT was administered before and after each intervention (VR and tablet), and dominant and nondominant hand scores were analyzed separately. The BBT was selected as a primary measure of unimanual gross manual dexterity because fine and gross upper-limb coordination are commonly impaired in children with DCD. The BBT was selected as a measure of unimanual gross manual dexterity because fine and gross upper-limb coordination are commonly impaired in children with DCD and are central to our motor performance objectives. This test has been used in other studies quantifying motor performance in this population [45].

Enjoyment of Physical Activity (PACES)

PACES was used to assess enjoyment during each intervention [46]. A child-appropriate 16-item version validated for pediatric physical activity research was administered [47-49]. At baseline, PACES was administered with a generic physical-activity stem, with items phrased as “when I am doing physical activity...” (eg, “I enjoy it” and “It’s very exciting”). After each intervention block, PACES was then readministered with condition-specific

wording (“when I am playing the VR game...” / “when I am playing the tablet game...”). The questionnaire included both positively and negatively worded items (eg, “It makes me feel depressed” and “It gives me energy”), with reverse scoring applied to negatively worded items before analysis. Children were instructed to reflect on their experience during the specific condition and respond accordingly. Higher scores indicated greater enjoyment. The questionnaire was administered electronically via Qualtrics (Qualtrics, LLC) [50]. A template of the PACES questionnaire can be located online [51].

The PACES questionnaire was completed before (pre) and after (post) each condition (VR and tablet) to account for baseline expectations and effects that can distort post scores. While enjoyment is commonly assessed postintervention, relying solely on postintervention scores can obscure the effect of the activity itself, particularly if participants hold strong initial expectations or attitudes. This approach was informed by methodological concerns raised in related domains, such as tolerability studies using the Simulator Sickness Questionnaire, where assuming a symptom-free baseline has led to inaccurate conclusions [52]. It is worth noting that children may have varying levels of enjoyment or anticipation before gameplay, meaning that not all postintervention responses can be attributed solely to the gameplay experience. Capturing enjoyment both before and after each condition allowed us to account for individual baseline affect and more accurately assess the impact of the intervention. If a child requested help, the assessor or a parent read items aloud and clarified word meanings using neutral definitions, and the child selected their own response.

Motivation and Feeling

Daily motivation and feeling during gameplay were recorded using a child-friendly log adapted from Scott et al [38] for home-based DCD interventions. The adapted “Motivation and Feeling Log” (a template can be located online) [53] was aligned to the 2 conditions (VR and tablet). Each participant completed a 5-day log for each condition immediately after each session. Motivation was rated on a 5-point Likert scale (1=“very low” to 5=“very high”), with an option to add brief comments. Feeling was captured by choosing one of a series of emoji-style faces (very unhappy to very happy) and by responding to short open-ended prompts (eg, “Tell us how you feel it went” and “Did you enjoy the tasks?”).

Statistical Analyses

All analyses were conducted in Jamovi (version 2.6.44; The Jamovi project) and R (version 4.2.2; R Foundation for Statistical Computing; R.app v. 1.79, macOS Big Sur ARM). The per-protocol sample (n=27) was used for all inferential tests. There were no missing outcome data among participants who completed both conditions; therefore, no imputation was performed. For MABC-3 and BBT outcomes, 2×2 within-subjects ANOVAs tested effects of time (pre and post) and condition (VR, tablet). In addition to the primary ANOVAs, exploratory within-condition paired-samples 2-tailed *t* tests were conducted to examine pre-post changes in each outcome. Exploratory within-condition tests were not adjusted for multiple comparisons and are interpreted cautiously. For motivation and feeling across days, a 2×5 repeated-measures ANOVA tested condition (VR and tablet) by day (1-5). Assumptions of normality (Shapiro-Wilk) and homogeneity of variance (Levene test) were tested in Jamovi. Effect sizes were reported as Cohen *d* for paired comparisons and partial η^2 for ANOVA terms. Statistical significance was set at $P<.05$ for all analyses.

Results

A total of 30 participants enrolled in this study; 3 withdrew during the intervention period (one during VR and two during tablet), leaving 27 participants who completed both conditions and were included in the per-protocol analyses. Repeated-measures ANOVAs examined condition (VR, tablet)×time (pre and post) effects on motor performance outcomes. For brevity, here we report on the interactions—the reporting of the main effects can be found in a study by Alharbi et al [54]. Interactions were nonsignificant for MABC-3 domains: manual dexterity ($F_{1,25}=0.81$; $P=.38$; $\eta^2_p=0.03$; Figure 2A), aiming and catching ($F_{1,25}=0.83$; $P=.37$; $\eta^2_p=0.03$; Figure 2B), balance and locomotion ($F_{1,25}=2.42$; $P=.13$; $\eta^2_p=0.09$; Figure 2C), and total motor score ($F_{1,25}=3.36$; $P=.08$; $\eta^2_p=0.19$; Figure 2D). Significant interactions were observed for the BBT dominant hand ($F_{1,25}=7.47$; $P=.01$; $\eta^2_p=0.23$; Figure 2E) and BBT nondominant hand ($F_{1,25}=11.94$; $P=.002$; $\eta^2_p=0.32$; Figure 2F). Estimated marginal means (95% CIs) for each outcome by condition (tablet vs VR) and time point (pre vs post) are summarized in Table 1.

Figure 2. Within-participant pre-post motor performance across tablet and virtual reality (VR) conditions: (A) manual dexterity—*Movement Assessment Battery for Children, Third Edition (MABC-3)*, (B) aiming and catching—*MABC-3*, (C) balance and locomotion—*MABC-3*, (D) Total Motor Score—*MABC-3*, (E) dominant hand—*Box and Block Test (BBT)*, and (F) nondominant hand—*BBT*. Asterisks indicate statistically significant differences ($P < .05$).

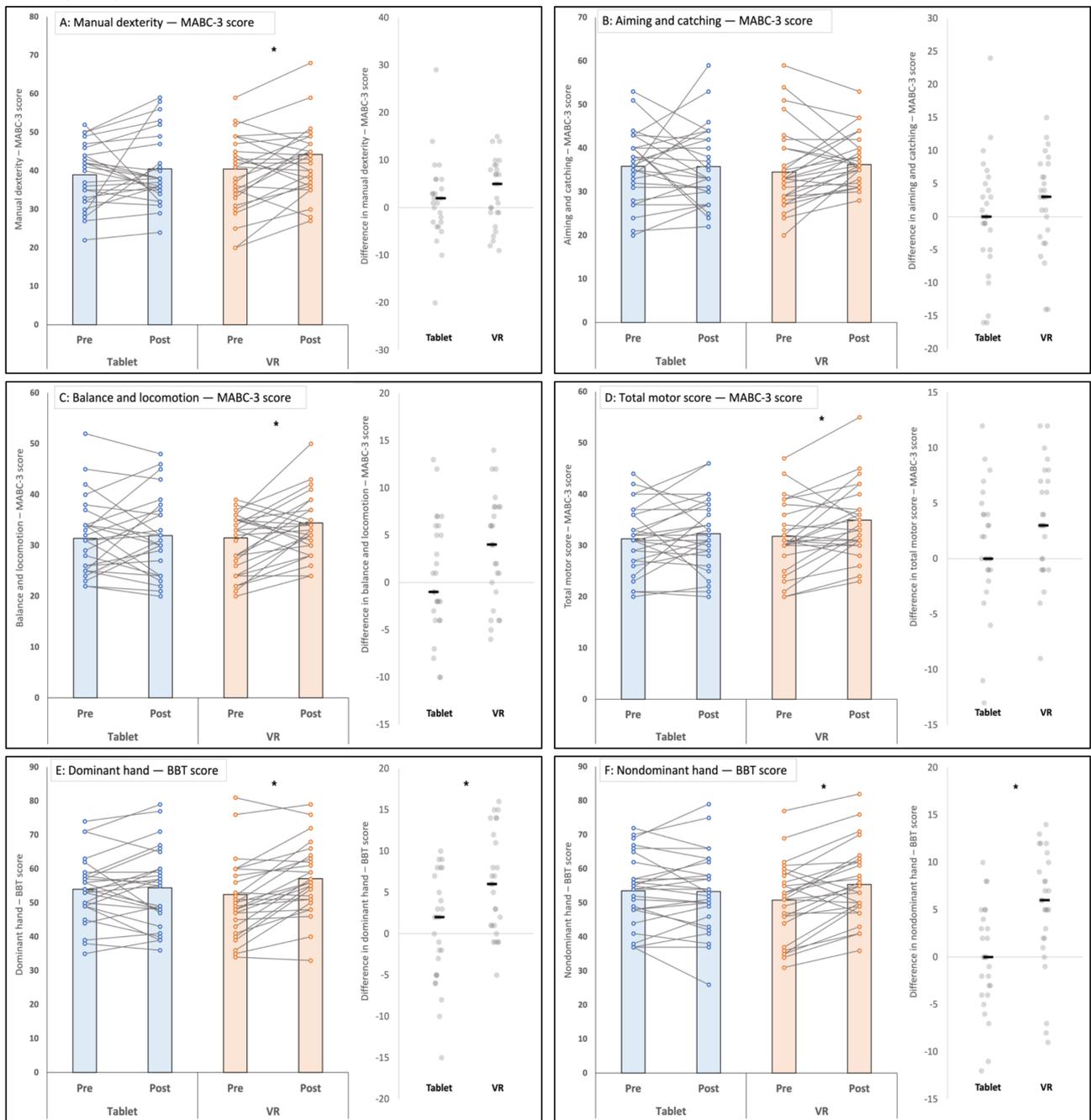


Table . Estimated marginal means (95% CIs) for outcomes by condition (tablet vs virtual reality [VR]) and time point (pre vs post).

Outcome	Tablet		VR	
	Pre, mean (95% CI)	Post, mean (95% CI)	Pre, mean (95% CI)	Post, mean (95% CI)
MABC-3^a				
Manual dexterity	39.0 (35.8 - 42.2)	40.5 (36.8 - 44.2)	39.4 (35.5 - 43.2)	42.6 (39.1 - 46.2)
Aiming and catching	35.8 (32.6 - 39.1)	35.7 (32.2 - 39.2)	35.2 (31.4 - 39.1)	37.3 (35.0 - 39.7)
Balance and locomotion	31.3 (28.4 - 34.2)	31.9 (28.8 - 35.0)	30.2 (27.9 - 32.6)	33.7 (31.3 - 36.2)
Total motor score	31.3 (28.7 - 33.9)	32.3 (29.5 - 35.1)	31.5 (28.8 - 34.2)	34.8 (32.0 - 37.5)
BBT^b				
Dominant hand	54.0 (50.2 - 57.8)	54.4 (49.8 - 59.1)	50.8 (46.5 - 55.1)	56.6 (52.5 - 60.8)
Nondominant hand	53.5 (49.3 - 57.8)	53.3 (48.6 - 58.0)	50.5 (46.1 - 55.0)	55.7 (51.6 - 59.9)
PACES^c				
Enjoyment	62.8 (59.5 - 66.0)	53.6 (48.3 - 58.9)	62.1 (58.9 - 65.3)	63.2 (59.1 - 67.4)

^aMABC-3: *Movement Assessment Battery for Children, Third Edition*.

^bBBT: Box and Block Test.

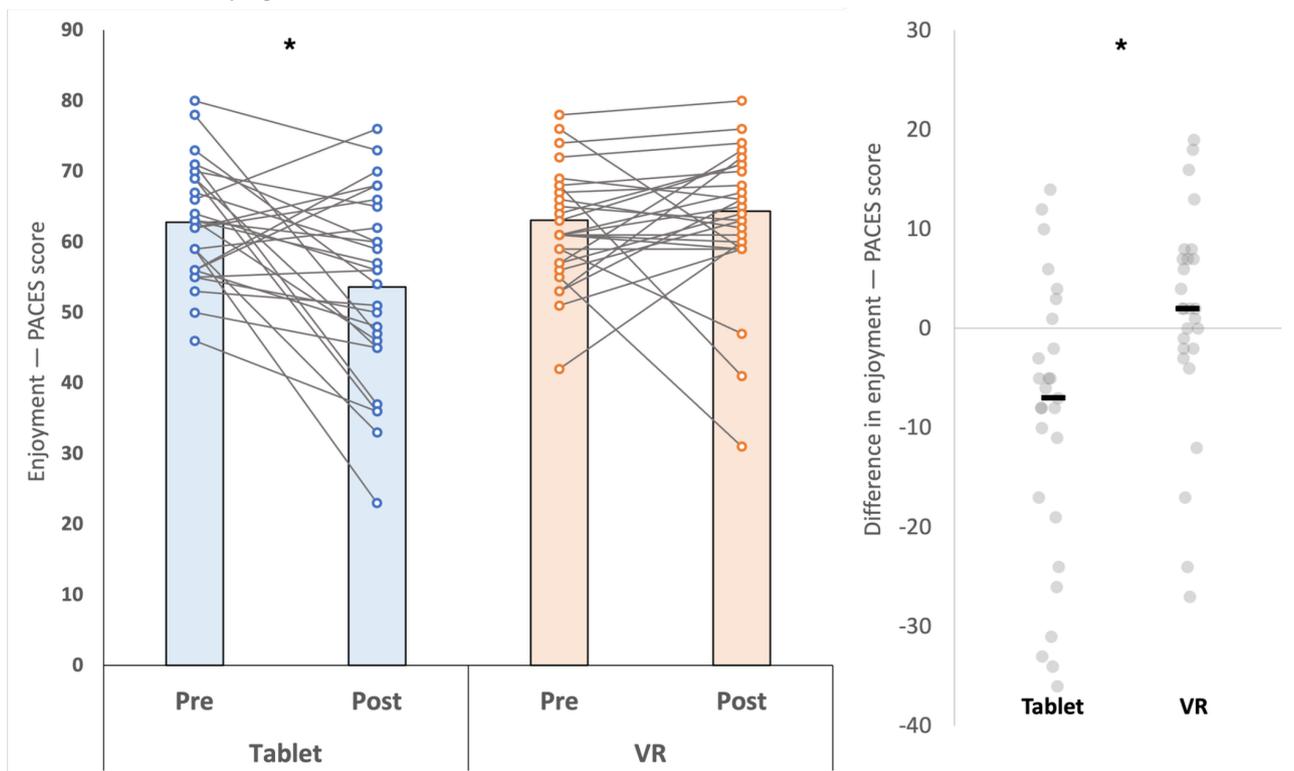
^cPACES: Physical Activity Enjoyment Scale.

Although no significant interactions were observed in the ANOVA with MABC-3 domains as the dependent variable, exploratory post hoc within-condition (paired-samples) *t* tests were conducted to explicitly examine performance improvements in the VR and tablet conditions. Paired-samples 2-tailed *t* tests revealed significant pre-post improvements in the VR condition across the MABC-3 domains of manual dexterity (ΔM 3.3, 95% CI 0.51 - 6.08; $t_{26}=-2.43$; $P=.02$; Cohen $d=0.47$; [Figure 2A](#)); balance and locomotion (ΔM 3.52, 95% CI 1.3 - 5.74; $t_{26}=-3.25$; $P=.003$; Cohen $d=0.63$; [Figure 2C](#)), and total motor score (ΔM 3.33, 95% CI 1.34 - 5.32; $t_{26}=-3.44$; $P=.002$; Cohen $d=0.66$; [Figure 2D](#)). No significant change was found for the aiming and catching domain (ΔM 2.15, 95% CI -0.72 to 5.01; $t_{26}=-1.54$; $P=.14$; Cohen $d=0.30$; [Figure 2B](#)).

For the BBT in the VR condition, significant improvements were found for both the dominant hand (ΔM 5.93, 95% CI 3.49 - 8.36; $t_{26}=-4.99$; $P<.001$; Cohen $d=0.96$; [Figure 2E](#)) and the nondominant hand (ΔM 5.11, 95% CI 2.65 - 7.57; $t_{26}=-4.27$; $P<.001$; Cohen $d=0.82$; [Figure 2F](#)). By contrast, there were no significant pre-post changes observed across any MABC-3 domains or BBT performance in the tablet condition (all $P\geq.36$).

For enjoyment (PACES), a condition \times time interaction was detected ($F_{1,26}=11.67$; $P=.002$; $\eta^2_p=0.31$; [Figure 3](#)). Enjoyment decreased from pre to post in the tablet condition (ΔM -9.19, 95% CI -14.79 to -3.58; $t_{26}=3.37$; $P=.002$; Cohen $d=0.65$) and did not change in the VR condition (ΔM 1.11, 95% CI -3.27 to 5.49; $t_{26}=-0.52$; $P=.61$; Cohen $d=0.10$).

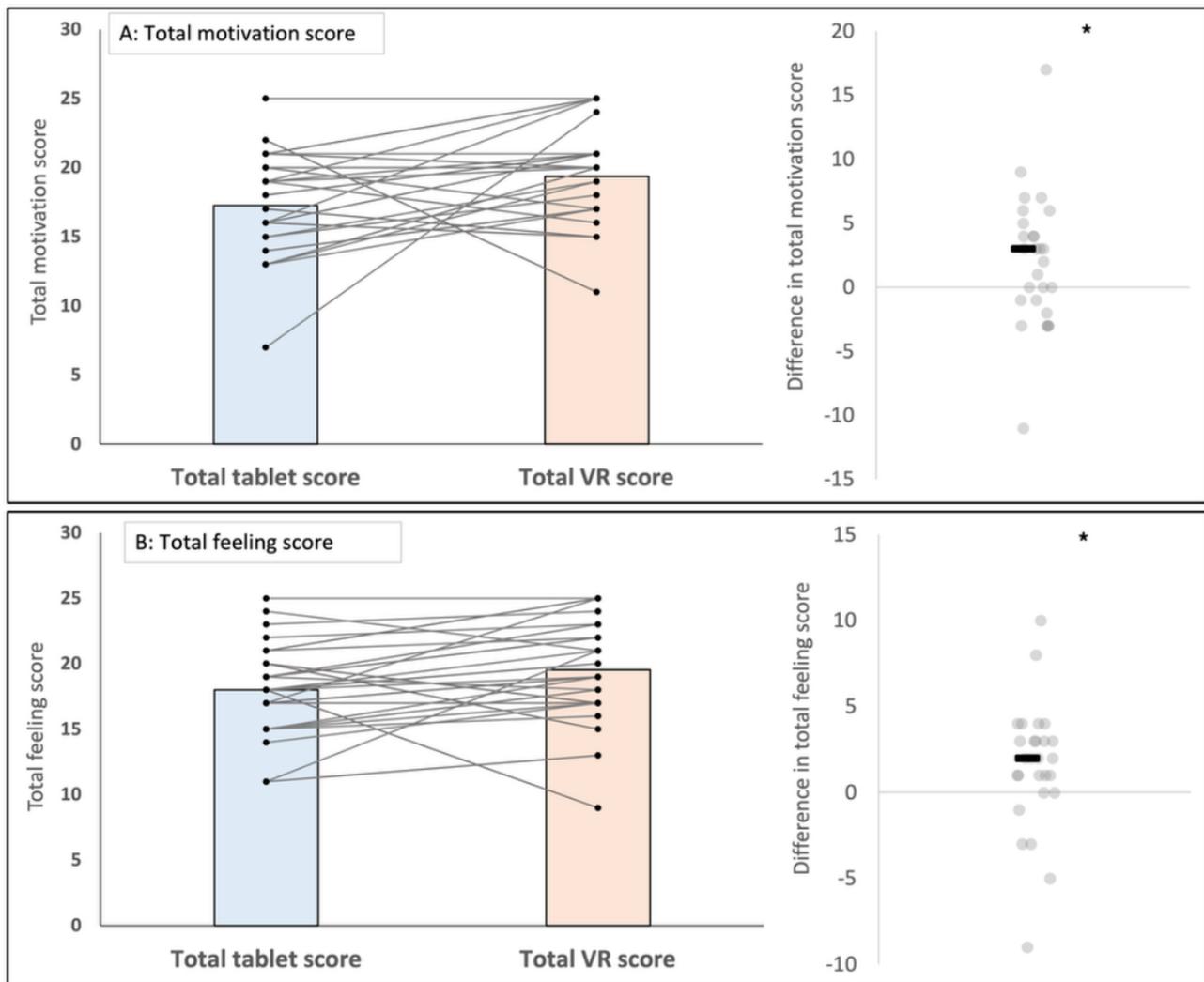
Figure 3. Within-participant pre-post enjoyment score (Physical Activity Enjoyment Scale [PACES]) across tablet and virtual reality (VR) conditions. Asterisks indicate statistically significant differences ($P < .05$).



Across the 5 days, there was no condition \times day interaction for motivation ($F_{4,104}=0.54$; $P=.71$; $\eta^2_p=0.02$), or feeling ($F_{4,104}=0.26$; $P=.91$; $\eta^2_p=0.01$). Main effects of condition were present, suggesting that motivation was higher in VR than tablet ($F_{1,26}=4.49$; $P=.04$; $\eta^2_p=0.15$); estimated marginal means on

the 1 to 5 scale were VR 3.87 (95% CI 3.60 - 4.15) versus tablet 3.45 (95% CI 3.15 - 3.75). Levels of positive feeling were also higher in VR than tablet ($F_{1,26}=4.54$; $P=.04$; $\eta^2_p=0.15$); estimated marginal means were VR 3.90 (95% CI 3.60 - 4.21) versus tablet 3.60 (95% CI 3.32 - 3.88; Figure 4A-B).

Figure 4. Effect of condition (virtual reality [VR] vs tablet) on self-reported outcomes: (A) motivation and (B) feeling. Asterisks indicate statistically significant differences between the VR and tablet conditions ($P < .05$).



There was no evidence that the order of conditions, nor the duration of the washout period following the VR task, affected performance on the second session [54].

The raw data, full analysis figures, and optional session comments from the daily logs are available on Open Science Framework [55,56,57].

Discussion

Principal Findings

This study investigated the effects of an immersive VR rhythm game on motor performance, motivation, and enjoyment in children and adolescents with DCD. The findings highlight immersive VR rhythm games as a promising tool for movement skill development, addressing a documented gap in the literature regarding immersive VR interventions for children and adolescents with DCD [23]. Our findings tentatively suggest that immersive VR rhythm games could be associated with improvements in various aspects of motor performance and may substantially enhance intrinsic motivation and emotional engagement in children with DCD, offering a supportive approach to traditional therapeutic approaches.

The key pattern in the data provides preliminary evidence of improvement in several domains of motor performance in children and adolescents with DCD following a short-term, home-based immersive VR rhythm game. Exploratory within-condition analyses indicated that participants in the VR condition showed pre-post improvements in manual dexterity, balance and locomotion, and total motor score on the MABC-3; however, the condition \times time interactions were non-significant. Consistent with this, the MABC-3 estimated marginal means showed wide and overlapping 95% CIs (Table 1), indicating limited precision for between-condition differences over this short period. Therefore, any differential effects of VR versus the tablet condition on these outcomes should be interpreted cautiously. These patterns were more clearly supported for upper-limb motor skills, where significant condition \times time interactions and within-condition changes in BBT scores for both the dominant and nondominant hands suggested greater gains following VR than tablet play. The improvements in manual dexterity and bilateral upper-limb control, as evidenced by the MABC-3 and BBT results, are particularly noteworthy given that these are common areas of difficulty for children with DCD [2,7]. Taken together, these findings tentatively support the practical utility of immersive VR as a potentially

effective tool for improving functional motor skills required for daily living and academic activities in children and adolescents with DCD [1,3,58].

Descriptively, the tablet condition failed to elicit any significant pre-to-post changes across any MABC-3 domains or BBT performance (all $P > .35$). This divergence in outcomes is consistent with, although does not definitively establish, the idea that the observed motor improvements are not merely attributable to general practice effects, maturation, or nonspecific engagement with a digital device. Instead, the pattern of results tentatively suggests that features of the immersive VR game may have contributed to the observed changes in motor performance. However, the present study cannot determine whether these effects are driven primarily by the immersive VR context itself, the rhythmic, music-synchronized gameplay, the whole-body movement demands, or a combination of these elements. The overall pattern is broadly consistent with prior evidence indicating that VR can support motor skill development across other pediatric conditions, including cerebral palsy and autism spectrum disorder [17,21,22]. Moreover, this study extends these benefits specifically to children and adolescents with DCD, where empirical research on immersive VR has been limited [23]. The nonsignificant condition \times time interactions for the MABC-3 domains mean that the between-condition differences in pre-post change on these measures should be interpreted cautiously. Several factors may explain the lack of interaction, including the modest sample ($n=27$), the relatively short exposure, variable levels of motivation and fatigue of participants with assessments often conducted after school, and reliance on self-reported adherence to the VR and tablet interventions. However, the significant condition \times time interactions identified for BBT scores for both the dominant hand and the nondominant hand provide evidence that, for this outcome, the immersive VR rhythm game improved participants' motor performance more than the tablet intervention. This finding might show that Beat Saber specifically targets the rapid whole-arm coordination that underpins the BBT.

Beyond motor performance, this study also highlights the potential impact of immersive VR rhythm games, like Beat Saber, on the subjective experiences of motivation and enjoyment in children and adolescents with DCD. Participants in the VR condition consistently reported higher motivation scores and more positive emotional experiences when compared to the tablet condition. Enjoyment is closely linked to persistence, as children are more likely to continue with physical activity and learning tasks when they experience them as fun rather than purely effortful or externally pressured [18,19]. The capacity of immersive VR rhythm games to transform these experiences into intrinsically motivating activities may therefore offer a significant clinical advantage. Further emphasizing this distinction, the PACES scores revealed a significant decrease in enjoyment after the tablet condition, whereas no significant change in enjoyment was observed in the VR condition. This suggests that the tablet game's capacity to sustain enjoyment over the intervention period was limited. Moreover, the tablet game lacked the physical interactivity, multisensory stimulation, and immersive qualities that characterize the VR experience.

This absence of immersive presence and direct physical engagement likely contributed to the decline in enjoyment, highlighting a fundamental difference in the experiential quality offered by the 2 modalities. This is further illustrated by a parent's observation:

This is interesting... [their child] played with the tablet on the way home [two days ago] but yesterday she didn't ask for it even though it was clearly in sight on the kitchen table whereas when she had the VR she was counting down the hours when she could go on it.

This observation supports the idea that immersive VR rhythm games led to more consistent and higher levels of engagement. These findings are also consistent with early feedback from the preliminary PPI activity, where participants expressed a preference for VR-based activities. However, a parent reported that their child used the tablet "on the way home," indicating that, in at least some cases, the tablet was used outside the intended "home-based" context. This raises the possibility that either the home-use instructions were not sufficiently clear or that adherence varied between participants. Importantly, this context difference could have influenced both the amount and quality of engagement in a direction that may confound comparisons between conditions. For example, tablet use during travel may have occurred in shorter, fragmented sessions and/or alongside other activities (eg, conversations, distractions, and fatigue), which could reduce sustained attention and engagement quality despite increasing opportunities to play. Conversely, the greater portability of the tablet could increase total exposure time relative to VR by enabling additional ad hoc play. Either pattern could bias engagement-related outcomes and any dose-response relationship independent of the game content itself. Future studies should therefore strengthen monitoring and standardization of context (eg, clearer instructions specifying location and timing, brief daily logs of where or when sessions occurred, or objective in-app time-stamps) and, where feasible, match interventions on portability or include context-of-use as a planned covariate.

Several features of the immersive VR rhythm task plausibly underlie both the motor improvements and the higher motivation and enjoyment observed. The Beat Saber game demands precise, bilateral upper-limb coordination, rapid visuomotor integration, and accurate rhythmic timing, skills commonly affected in DCD [24,25]. High-dose, task-specific practice with immediate visual, auditory, and haptic feedback supports error detection and correction, consistent with motor-learning principles [16,17,35,36]. Moreover, the adaptive difficulty levels within the game, allowing participants to progress from "normal" to "hard" or regress to "easy" mode, align with principles of optimal challenge, ensuring that the task remains sufficiently demanding yet achievable, thereby maximizing learning [59], but we did not formally titrate or verify that the task was optimally matched to each child's capacity. It is therefore unclear to what extent the overall difficulty was fine-tuned for individual participants with DCD. The immersive nature of VR creates a powerful sense of "being there" (presence), which can significantly enhance engagement and reduce the perception of effort associated with physical activity [12,19]. This

transformation of therapeutic exercise into a playful and rewarding experience is particularly potent for children, who are naturally drawn to engaging and interactive activities [60]. The coupling of synchronized music, rhythmic movement, and immediate, clear feedback provides an intrinsically rewarding loop that fosters competence and self-efficacy [23-26,61]. Together, these properties can improve upper-limb coordination and related motor performance, and also sustain motivation and positive affect across sessions [25-27,60]. A participant's remark illustrates this alignment:

Felt confident on harder levels. Coordination (hand-eye) is getting better. I found the tasks [VR's one] pleasurable and exciting, up to my growth.

This suggests that the level progression matched their developing skill level and maintained engagement. Thus, this adaptability may have supported adherence and sustained participation, which are common barriers in traditional rehabilitation programs [19,24,27]. By contrast, although the tablet game (Cut the Rope) also includes level progression, increases in difficulty primarily reflect greater puzzle complexity rather than changes in movement demands, and the game lacks the physical interactivity, multisensory feedback, and immersive presence of VR. These differences in the type of challenge and the absence of immersive, movement-based engagement are consistent with the tablet game being experienced as lower in direct physical demand and less engaging overall. Future work could examine how in-game performance relates to improvements in movement assessment scores to shed light on the mechanisms that underpin these effects.

At the same time, participants' comments also highlighted important drawbacks and limits to the acceptability of immersive VR. Not all children enjoyed or tolerated the VR condition as some described the headset as heavy or uncomfortable (eg, "I don't like the game and the headset [was] heavy"), and one participant characterized the experience as "too stressful." Another child reported that their "head just felt like it was going to explode even though Mum helped clean the VR lens and adjust the straps" and that they did not feel they had improved. Other comments referred to a loss of enthusiasm and concentration towards the end of the week, or a dislike of the repetitiveness of the game and its music, even when parents perceived some improvement. Together, these accounts indicate that immersive VR is not universally acceptable or tolerable for all children with DCD, and that issues such as physical discomfort, perceived stress, visual or cognitive overload, and content-specific preferences (eg, music and repetition) can limit engagement [62]. These quotes warrant more cautious conclusions about immersive VR and highlight the importance of flexible VR content that can accommodate differing preferences and tolerance levels [63].

Limitations and Future Directions

Despite the promising findings, this study is not without limitations that warrant consideration for future research. First, the relatively small sample size might constrain the generalizability of the findings to the broader population of children and adolescents with DCD. The sample included more boys than girls, which is common in DCD cohorts but may also

limit the generalizability of this intervention [64]. Although the within-subject crossover design mitigates the risk that this imbalance biased the comparison between VR and tablet conditions [65], the study was not powered to examine gender as a moderator, and we cannot determine whether interest, engagement, or motor outcomes differed meaningfully between boys and girls. Future work should purposively recruit more girls and formally test gender-related differences in response to immersive VR interventions. Second, the intervention duration of five consecutive days, while demonstrating acute effects, may not be sufficient to capture long-term motor learning, skill retention, or sustained behavioral changes. Some gains, particularly in motivation and enjoyment, may reflect novelty, since VR is new to most children compared with tablets. We do not know whether this motivation difference would persist once the novelty of the intervention declines [66]. Longitudinal studies with extended intervention periods and follow-up assessments are crucial to ascertain the durability of the observed motor improvements and the persistence of enhanced motivation and enjoyment. Third, the home-based nature of the intervention posed challenges for direct supervision and objective compliance monitoring. Although caregiver and self-report logs were used, these are susceptible to social desirability bias [67]. Future studies should integrate objective measures of engagement and adherence, such as in-game metrics, wearable sensor data, or remote video monitoring, to provide more accurate and verifiable data on intervention fidelity [68]. Fourth, only one immersive VR rhythm game (Beat Saber) was used, so the findings may not generalize to all immersive VR content; whether other immersive VR games without rhythm-based mechanics yield similar benefits remains to be tested. This study, therefore, cannot determine whether the observed effects are driven primarily by immersion, rhythmic or music-synchronized gameplay, whole-body movement demands, or their combination. Future research could address this by using multiple comparison conditions (eg, nonimmersive rhythm games, immersive nonrhythm VR games, or rhythm-based games with different movement demands) to help disentangle the contribution of immersion, rhythm, and whole-body engagement. Fifth, the tablet game (Cut the Rope) was intentionally chosen to minimize motor demands, but it may be more passive than the VR rhythm task. As a result, between-condition differences may simply reflect a greater 'dose' of movement, rather than effects specific to immersive VR. Future trials could include an active, movement-based nonimmersive control (eg, a screen-based rhythm game) matched on session length and movement intensity, verified with telemetry or accelerometry or heart-rate monitoring, to better isolate VR-specific contributions. Sixth, assessments frequently occurred after school, introducing variability in fatigue and attention. Future work, therefore, should standardize testing times or record time-of-day and brief fatigue or attention ratings so these factors can be controlled [69]. Finally, because multiple exploratory within-condition tests were conducted in MABC-3 domains after a nonsignificant interaction, the risk of type I error is increased, and these findings should be interpreted with caution [70].

Conclusions

This study is the first to evaluate a home-based immersive VR rhythm game for children and adolescents with DCD in comparison to a tablet-based control condition with a crossover design. Prior literature in this area has focused on nonimmersive VR and provides limited direct evidence comparing immersive VR rhythm gameplay against nonimmersive gameplay with a similar practice dose. Our findings add early evidence that

immersive VR rhythm gameplay can be highly engaging and may support short-term gains in manual dexterity relative to tablet gameplay. In real-world terms, these games could be a scalable home-based supplement to conventional therapy to increase practice opportunities and enjoyment; future trials should use larger samples, longer follow-up, objective usage or context tracking, and multiple comparison conditions to disentangle the effects of immersion, rhythm, and whole-body movement.

Acknowledgments

The authors would like to thank all the participants who contributed to this study. No generative artificial intelligence was used in any part of the manuscript.

Funding

This study was funded by University of Hail and Ministry of Education of Saudi Arabia. The funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The deidentified datasets generated and analyzed in this study are openly available on the Open Science Framework repository [56]. No additional restrictions apply to data access. The study was not preregistered, and no registered analysis plan exists.

Authors' Contributions

Concept/research design: MA, GB, DH, GW, HD

Data collection: MA

Project management: MA, GB

Consultation (including review of manuscript before submitting): MA, GB, DH, GW, HD

Conflicts of Interest

None declared.

Checklist 1

JARS-Quant checklist.

[[PDF File, 1373 KB](#) - [games_v14ile84995_app1.pdf](#)]

References

1. Summers J, Larkin D, Dewey D. Activities of daily living in children with developmental coordination disorder: dressing, personal hygiene, and eating skills. *Hum Mov Sci* 2008 Apr;27(2):215-229. [doi: [10.1016/j.humov.2008.02.002](https://doi.org/10.1016/j.humov.2008.02.002)] [Medline: [18348898](https://pubmed.ncbi.nlm.nih.gov/18348898/)]
2. Homolová M. Dyspraxia: limitations in coordination and locomotion in a group of preschool children. Presented at: 2024 International Masaryk Conference for PhD Students and Young Researchers [Conference in Czech]; Dec 16-18, 2024; Hradec Králové, The Czech Republic p. 458-466 URL: <https://sekarl.euba.sk/arl-eu/en/csg/?repo=eurepo&key=8806214757> [accessed 2026-01-30]
3. Axford C, Joosten AV, Harris C. iPad applications that required a range of motor skills promoted motor coordination in children commencing primary school. *Aust Occup Ther J* 2018 Apr;65(2):146-155. [doi: [10.1111/1440-1630.12450](https://doi.org/10.1111/1440-1630.12450)] [Medline: [29355982](https://pubmed.ncbi.nlm.nih.gov/29355982/)]
4. Haga M. The relationship between physical fitness and motor competence in children. *Child Care Health Dev* 2008 May;34(3):329-334. [doi: [10.1111/j.1365-2214.2008.00814.x](https://doi.org/10.1111/j.1365-2214.2008.00814.x)] [Medline: [18410639](https://pubmed.ncbi.nlm.nih.gov/18410639/)]
5. Logan SW, Robinson LE, Wilson AE, Lucas WA. Getting the fundamentals of movement: a meta-analysis of the effectiveness of motor skill interventions in children. *Child Care Health Dev* 2012 May;38(3):305-315. [doi: [10.1111/j.1365-2214.2011.01307.x](https://doi.org/10.1111/j.1365-2214.2011.01307.x)] [Medline: [21880055](https://pubmed.ncbi.nlm.nih.gov/21880055/)]
6. Lopes L, Santos R, Pereira B, Lopes VP. Associations between sedentary behavior and motor coordination in children. *Am J Hum Biol* 2012;24(6):746-752. [doi: [10.1002/ajhb.22310](https://doi.org/10.1002/ajhb.22310)] [Medline: [22893386](https://pubmed.ncbi.nlm.nih.gov/22893386/)]
7. Farmer M, Echenne B, Bentourkia M. Study of clinical characteristics in young subjects with developmental coordination disorder. *Brain Dev* 2016 Jun;38(6):538-547. [doi: [10.1016/j.braindev.2015.12.010](https://doi.org/10.1016/j.braindev.2015.12.010)] [Medline: [26763621](https://pubmed.ncbi.nlm.nih.gov/26763621/)]

8. Tsai CL, Wilson PH, Wu SK. Role of visual-perceptual skills (non-motor) in children with developmental coordination disorder. *Hum Mov Sci* 2008 Aug;27(4):649-664. [doi: [10.1016/j.humov.2007.10.002](https://doi.org/10.1016/j.humov.2007.10.002)] [Medline: [18374437](https://pubmed.ncbi.nlm.nih.gov/18374437/)]
9. Fong SSM, Lee VYL, Chan NNC, Chan RSH, Chak WK, Pang MYC. Motor ability and weight status are determinants of out-of-school activity participation for children with developmental coordination disorder. *Res Dev Disabil* 2011;32(6):2614-2623. [doi: [10.1016/j.ridd.2011.06.013](https://doi.org/10.1016/j.ridd.2011.06.013)] [Medline: [21767931](https://pubmed.ncbi.nlm.nih.gov/21767931/)]
10. Gao J, Song W, Zhong Y, et al. Children with developmental coordination disorders: a review of approaches to assessment and intervention. *Front Neurol* 2024 May;15:1359955. [doi: [10.3389/fneur.2024.1359955](https://doi.org/10.3389/fneur.2024.1359955)] [Medline: [38846037](https://pubmed.ncbi.nlm.nih.gov/38846037/)]
11. Moon J, Webster CA, Stodden DF, et al. Systematic review and meta-analysis of physical activity interventions to increase elementary children's motor competence: a comprehensive school physical activity program perspective. *BMC Public Health* 2024 Mar 15;24(1):826. [doi: [10.1186/s12889-024-18145-1](https://doi.org/10.1186/s12889-024-18145-1)] [Medline: [38491432](https://pubmed.ncbi.nlm.nih.gov/38491432/)]
12. Hamad A, Jia B. How virtual reality technology has changed our lives: an overview of the current and potential applications and limitations. *Int J Environ Res Public Health* 2022 Sep 8;19(18):11278. [doi: [10.3390/ijerph191811278](https://doi.org/10.3390/ijerph191811278)] [Medline: [36141551](https://pubmed.ncbi.nlm.nih.gov/36141551/)]
13. Li S, Li B, Zhang S, et al. A markerless visual-motor tracking system for behavior monitoring in DCD assessment. Presented at: 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC); Dec 12-15, 2017. [doi: [10.1109/APSIPA.2017.8282139](https://doi.org/10.1109/APSIPA.2017.8282139)]
14. Phelan I, Furness PJ, Dunn HD, et al. Immersive virtual reality in children with upper limb injuries: findings from a feasibility study. *J Pediatr Rehabil Med* 2021;14(3):401-414. [doi: [10.3233/PRM-190635](https://doi.org/10.3233/PRM-190635)] [Medline: [34151871](https://pubmed.ncbi.nlm.nih.gov/34151871/)]
15. Meyns P, Pans L, Plasmans K, Heyrman L, Desloovere K, Molenaers G. The effect of additional virtual reality training on balance in children with cerebral palsy after lower limb surgery: a feasibility study. *Games Health J* 2017 Feb;6(1):39-48. [doi: [10.1089/g4h.2016.0069](https://doi.org/10.1089/g4h.2016.0069)] [Medline: [28051880](https://pubmed.ncbi.nlm.nih.gov/28051880/)]
16. Holden MK. Virtual environments for motor rehabilitation: review. *Cyberpsychol Behav* 2005 Jun;8(3):187-211. [doi: [10.1089/cpb.2005.8.187](https://doi.org/10.1089/cpb.2005.8.187)] [Medline: [15971970](https://pubmed.ncbi.nlm.nih.gov/15971970/)]
17. Gagliardi C, Turconi AC, Biffi E, et al. Immersive virtual reality to improve walking abilities in cerebral palsy: a pilot study. *Ann Biomed Eng* 2018 Sep;46(9):1376-1384. [doi: [10.1007/s10439-018-2039-1](https://doi.org/10.1007/s10439-018-2039-1)] [Medline: [29704186](https://pubmed.ncbi.nlm.nih.gov/29704186/)]
18. Atkins MR, Johnson DM, Force EC, Petrie TA. Peers, parents, and coaches, oh my! the relation of the motivational climate to boys' intention to continue in sport. *Psychol Sport Exerc* 2015 Mar;16(3):170-180. [doi: [10.1016/j.psychsport.2014.10.008](https://doi.org/10.1016/j.psychsport.2014.10.008)]
19. Iten N, Petko D. Learning with serious games: is fun playing the game a predictor of learning success? *Br J Educ Technol* 2016 Jan;47(1):151-163. [doi: [10.1111/bjet.12226](https://doi.org/10.1111/bjet.12226)]
20. Günar BB, Bavlı Ö. Effectiveness of virtual reality games on the specific sport skills of adolescents. *BMC Sports Sci Med Rehabil* 2025 Jul 10;17(1):194. [doi: [10.1186/s13102-025-01243-x](https://doi.org/10.1186/s13102-025-01243-x)] [Medline: [40640974](https://pubmed.ncbi.nlm.nih.gov/40640974/)]
21. Biffi E, Beretta E, Cesareo A, et al. An immersive virtual reality platform to enhance walking ability of children with acquired brain injuries. *Methods Inf Med* 2017 Mar 23;56(2):119-126. [doi: [10.3414/ME16-02-0020](https://doi.org/10.3414/ME16-02-0020)] [Medline: [28116417](https://pubmed.ncbi.nlm.nih.gov/28116417/)]
22. Dixon DR, Miyake CJ, Nohelty K, Novack MN, Granpeesheh D. Evaluation of an immersive virtual reality safety training used to teach pedestrian skills to children with autism spectrum disorder. *Behav Anal Pract* 2019;13(3):631-640. [doi: [10.1007/s40617-019-00401-1](https://doi.org/10.1007/s40617-019-00401-1)] [Medline: [32953391](https://pubmed.ncbi.nlm.nih.gov/32953391/)]
23. Alharbi M, Du H, Harris D, Wood G, Dodd H, Buckingham G. Evaluating the impact of virtual reality game training on upper limb motor performance in children and adolescents with developmental coordination disorder: a scoping review using the ICF framework. *J Neuroeng Rehabil* 2024 Jun 5;21(1):95. [doi: [10.1186/s12984-024-01393-y](https://doi.org/10.1186/s12984-024-01393-y)] [Medline: [38840217](https://pubmed.ncbi.nlm.nih.gov/38840217/)]
24. Albert I, Burkard N, Queck D, Herrlich M. The effect of auditory-motor synchronization in exergames on the example of the VR rhythm game Beat Saber. *Proc ACM Hum Comput Interact* 2022 Oct 25;6(CHI PLAY):1-26. [doi: [10.1145/3549516](https://doi.org/10.1145/3549516)]
25. Rutkowski S, Jakóbczyk A, Abrahamek K, et al. Training using a commercial immersive virtual reality system on hand-eye coordination and reaction time in students: a randomized controlled trial. *Virtual Real* 2024 Mar;28(1):7. [doi: [10.1007/s10055-023-00898-6](https://doi.org/10.1007/s10055-023-00898-6)]
26. Grosprêtre S, Marcel-Millet P, Eon P, Wollesen B. How exergaming with virtual reality enhances specific cognitive and visuo-motor abilities: an explorative study. *Cogn Sci* 2023 Apr;47(4):e13278. [doi: [10.1111/cogs.13278](https://doi.org/10.1111/cogs.13278)] [Medline: [37029516](https://pubmed.ncbi.nlm.nih.gov/37029516/)]
27. Boots B, Berg D, Hewitt E, Naugle K, Naugle K. Physical activity and enjoyment in active virtual reality games in youth: comparative analysis of Gorilla Tag and Beat Saber. *JMIR Serious Games* 2025 Apr 1;13(1):e66593. [doi: [10.2196/66593](https://doi.org/10.2196/66593)] [Medline: [40168642](https://pubmed.ncbi.nlm.nih.gov/40168642/)]
28. Alharbi M, Harris D, Dodd H, Wood G, Buckingham G. PPI clean copy. Open Science Framework. 2024 Sep 4. URL: <https://osf.io/ewjcy/files/efgm4> [accessed 2026-01-30]
29. Betts J, McKay J, Maruff P, Anderson V. The development of sustained attention in children: the effect of age and task load. *Child Neuropsychol* 2006 Jun;12(3):205-221. [doi: [10.1080/09297040500488522](https://doi.org/10.1080/09297040500488522)] [Medline: [16837396](https://pubmed.ncbi.nlm.nih.gov/16837396/)]
30. Tychsen L, Foeller P. Effects of immersive virtual reality headset viewing on young children: visuomotor function, postural stability, and motion sickness. *Am J Ophthalmol* 2020 Jan;209:151-159. [doi: [10.1016/j.ajo.2019.07.020](https://doi.org/10.1016/j.ajo.2019.07.020)] [Medline: [31377280](https://pubmed.ncbi.nlm.nih.gov/31377280/)]
31. Wilson B, Crawford S. The Developmental Coordination Disorder Questionnaire 2007 (DCDQ'07): administration manual for the DCDQ'07 with psychometric properties. : DCDQ.ca; 2012 Mar URL: <https://www.dcdq.ca/uploads/pdf/DCDQ'07%20Manual%20Feb%2020th%202012.pdf> [accessed 2026-01-30]

32. Ku B. Validity of the developmental coordination disorder questionnaire' 07 and its association with physical activity in young children with developmental disabilities. *Asian J Kinesiol* 2021 Jan;23(1):10-19. [doi: [10.15758/ajk.2021.23.1.10](https://doi.org/10.15758/ajk.2021.23.1.10)]
33. Blank R, Barnett AL, Cairney J, et al. International clinical practice recommendations on the definition, diagnosis, assessment, intervention, and psychosocial aspects of developmental coordination disorder. *Dev Med Child Neurol* 2019 Mar;61(3):242-285. [doi: [10.1111/dmcn.14132](https://doi.org/10.1111/dmcn.14132)] [Medline: [30671947](https://pubmed.ncbi.nlm.nih.gov/30671947/)]
34. Appelbaum M, Cooper H, Kline RB, Mayo-Wilson E, Nezu AM, Rao SM. Journal Article Reporting Standards for Quantitative Research in psychology: the APA Publications and Communications Board task force report. *Am Psychol* 2018 Jan;73(1):3-25. [doi: [10.1037/amp0000191](https://doi.org/10.1037/amp0000191)] [Medline: [29345484](https://pubmed.ncbi.nlm.nih.gov/29345484/)]
35. Dongas R, Grace K. Designing to leverage presence in VR rhythm games. *Multimodal Technol Interact* 2023;7(2):18. [doi: [10.3390/mti7020018](https://doi.org/10.3390/mti7020018)]
36. Pesek M, Hirci N, Žnidaršič K, Marolt M. Enhancing music rhythmic perception and performance with a VR game. *Virtual Real* 2024;28(2):118. [doi: [10.1007/s10055-024-01014-y](https://doi.org/10.1007/s10055-024-01014-y)]
37. Alharbi M, Harris D, Dodd H, Wood G, Buckingham G. VR instructions. Open Science Framework. 2024 Sep 11. URL: <https://osf.io/87qk3/files/u6wya> [accessed 2026-01-30]
38. Scott MW, Wood G, Holmes PS, Marshall B, Williams J, Wright DJ. Combined action observation and motor imagery improves learning of activities of daily living in children with developmental coordination disorder. *PLoS One* 2023;18(5):e0284086. [doi: [10.1371/journal.pone.0284086](https://doi.org/10.1371/journal.pone.0284086)] [Medline: [37220154](https://pubmed.ncbi.nlm.nih.gov/37220154/)]
39. Wu OK, Brown T, Yu ML, Joshua N, Wilson CJ, Campbell HC. Test-retest reliability and convergent validity of the Movement Assessment Battery for Children-Third Edition with Australian 3-6-year-olds and their parents. *J Occup Ther Sch Early Interv* 2025 Jan 2;18(1):1-21. [doi: [10.1080/19411243.2023.2271480](https://doi.org/10.1080/19411243.2023.2271480)]
40. Brown T, Lalor A. The Movement Assessment Battery for Children--Second Edition (MABC-2): a review and critique. *Phys Occup Ther Pediatr* 2009;29(1):86-103. [doi: [10.1080/01942630802574908](https://doi.org/10.1080/01942630802574908)] [Medline: [19197761](https://pubmed.ncbi.nlm.nih.gov/19197761/)]
41. Desrosiers J, Bravo G, Hébert R, Dutil E, Mercier L. Validation of the Box and Block Test as a measure of dexterity of elderly people: reliability, validity, and norms studies. *Arch Phys Med Rehabil* 1994 Jul;75(7):751-755. [doi: [10.1016/0003-9993\(94\)90130-9](https://doi.org/10.1016/0003-9993(94)90130-9)] [Medline: [8024419](https://pubmed.ncbi.nlm.nih.gov/8024419/)]
42. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. *Am J Occup Ther* 1985 Jun;39(6):386-391. [doi: [10.5014/ajot.39.6.386](https://doi.org/10.5014/ajot.39.6.386)] [Medline: [3160243](https://pubmed.ncbi.nlm.nih.gov/3160243/)]
43. Connell LA, Tyson SF. Clinical reality of measuring upper-limb ability in neurologic conditions: a systematic review. *Arch Phys Med Rehabil* 2012 Feb;93(2):221-228. [doi: [10.1016/j.apmr.2011.09.015](https://doi.org/10.1016/j.apmr.2011.09.015)] [Medline: [22289230](https://pubmed.ncbi.nlm.nih.gov/22289230/)]
44. Liang KJ, Chen HL, Shieh JY, Wang TN. Measurement properties of the box and block test in children with unilateral cerebral palsy. *Sci Rep* 2021 Oct 25;11(1):20955. [doi: [10.1038/s41598-021-00379-3](https://doi.org/10.1038/s41598-021-00379-3)] [Medline: [34697312](https://pubmed.ncbi.nlm.nih.gov/34697312/)]
45. Krajenbrink H, Lust JM, Wilmut K, Steenbergen B. Motor and cognitive dual-task performance under low and high task complexity in children with and without developmental coordination disorder. *Res Dev Disabil* 2023 Apr;135:104453. [doi: [10.1016/j.ridd.2023.104453](https://doi.org/10.1016/j.ridd.2023.104453)] [Medline: [36773489](https://pubmed.ncbi.nlm.nih.gov/36773489/)]
46. Kendzierski D, DeCarlo KJ. Physical Activity Enjoyment Scale: two validation studies. *J Sport Exerc Psychol* 1991;13(1):50-64. [doi: [10.1123/jsep.13.1.50](https://doi.org/10.1123/jsep.13.1.50)]
47. Moore JB, Yin Z, Hanes J, et al. Measuring enjoyment of physical activity in children: validation of the physical activity enjoyment scale. *J Appl Sport Psychol* 2009 Jan 1;21(S1):S116-S129. [doi: [10.1080/10413200802593612](https://doi.org/10.1080/10413200802593612)] [Medline: [20209028](https://pubmed.ncbi.nlm.nih.gov/20209028/)]
48. Trecroci A, Cavaggioni L, Lastella M, et al. Effects of traditional balance and slackline training on physical performance and perceived enjoyment in young soccer players. *Res Sports Med* 2018;26(4):450-461. [doi: [10.1080/15438627.2018.1492392](https://doi.org/10.1080/15438627.2018.1492392)] [Medline: [29963921](https://pubmed.ncbi.nlm.nih.gov/29963921/)]
49. Carraro A, Young MC, Robazza C. A contribution to the validation of the Physical Activity Enjoyment Scale in an Italian sample. *Soc Behav Pers* 2008 Jan 1;36(7):911-918. [doi: [10.2224/sbp.2008.36.7.911](https://doi.org/10.2224/sbp.2008.36.7.911)]
50. Qualtrics. URL: <https://www.qualtrics.com> [accessed 2026-01-30]
51. Alharbi M, Harris D, Dodd H, Wood G, Buckingham G. PACES questionnaire. Open Science Framework. URL: <https://osf.io/87qk3/files/5nfmz8> [accessed 2026-01-30]
52. Brown P, Spronck P, Powell W. The Simulator Sickness Questionnaire, and the erroneous zero baseline assumption. *Front Virtual Real* 2022;3:945800. [doi: [10.3389/frvir.2022.945800](https://doi.org/10.3389/frvir.2022.945800)]
53. Alharbi M, Harris D, Dodd H, Wood G, Buckingham G. Motivation and feeling logbook. Open Science Framework. 2024 Sep 11. URL: <https://osf.io/87qk3/files/wuz2m> [accessed 2026-01-30]
54. Alharbi M, Harris D, Dodd H, Wood G, Buckingham G. All data statistics. Open Science Framework. 2024 Sep 11. URL: <https://osf.io/87qk3/files/8xzqj> [accessed 2026-01-30]
55. Alharbi M, Harris D, Dodd H, Wood G, Buckingham G. BeatSaber raw data. Open Science Framework. 2024 Sep 11. URL: <https://osf.io/87qk3/files/7byd2> [accessed 2026-01-30]
56. Alharbi M, Harris D, Dodd H, Wood G, Buckingham G. Files. Open Science Framework. URL: <https://osf.io/87qk3/files/osfstorage> [accessed 2026-01-30]
57. Alharbi M, Harris D, Dodd H, Wood G, Buckingham G. Participants comments on VR and tablet conditions. Open Science Framework. URL: <https://osf.io/87qk3/files/njyrw> [accessed 2026-01-30]

58. Van der Linde BW, van Netten JJ, Otten B, Postema K, Geuze RH, Schoemaker MM. Activities of daily living in children with developmental coordination disorder: performance, learning, and participation. *Phys Ther* 2015 Nov;95(11):1496-1506. [doi: [10.2522/ptj.20140211](https://doi.org/10.2522/ptj.20140211)] [Medline: [26045605](https://pubmed.ncbi.nlm.nih.gov/26045605/)]
59. Yildirim O, Surer E. Developing adaptive serious games for children with specific learning difficulties: a two-phase usability and technology acceptance study. *JMIR Serious Games* 2021 May 31;9(2):e25997. [doi: [10.2196/25997](https://doi.org/10.2196/25997)] [Medline: [34057415](https://pubmed.ncbi.nlm.nih.gov/34057415/)]
60. Jones T, Baque E, O'Grady KA, et al. Experiences of children with bronchiectasis and their parents in a novel play-based therapeutic exercise programme: a qualitative analysis. *BMJ Open* 2024 Jul 31;14(7):e078994. [doi: [10.1136/bmjopen-2023-078994](https://doi.org/10.1136/bmjopen-2023-078994)] [Medline: [39089712](https://pubmed.ncbi.nlm.nih.gov/39089712/)]
61. Santayana de Souza M, Nobre GC, Valentini NC. Effect of a motor skill-based intervention in the relationship of individual and contextual factors in children with and without Developmental Coordination Disorder from low-income families. *Psychol Sport Exerc* 2023 Jul;67:102406. [doi: [10.1016/j.psychsport.2023.102406](https://doi.org/10.1016/j.psychsport.2023.102406)] [Medline: [37665867](https://pubmed.ncbi.nlm.nih.gov/37665867/)]
62. Simón-Vicente L, Rodríguez-Cano S, Delgado-Benito V, Ausín-Villaverde V, Cubo Delgado E. Cybersickness. A systematic literature review of adverse effects related to virtual reality. *Neurologia (Engl Ed)* 2024 Oct;39(8):701-709. [doi: [10.1016/j.nrleng.2022.04.007](https://doi.org/10.1016/j.nrleng.2022.04.007)] [Medline: [39396266](https://pubmed.ncbi.nlm.nih.gov/39396266/)]
63. 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. [doi: [10.2196/19840](https://doi.org/10.2196/19840)] [Medline: [33095182](https://pubmed.ncbi.nlm.nih.gov/33095182/)]
64. Li H, Ke X, Huang D, et al. The prevalence of developmental coordination disorder in children: a systematic review and meta-analysis. *Front Pediatr* 2024;12:1387406. [doi: [10.3389/fped.2024.1387406](https://doi.org/10.3389/fped.2024.1387406)] [Medline: [39391054](https://pubmed.ncbi.nlm.nih.gov/39391054/)]
65. Dwan K, Li T, Altman DG, Elbourne D. CONSORT 2010 statement: extension to randomised crossover trials. *BMJ* 2019 Jul 31;366:14378. [doi: [10.1136/bmj.14378](https://doi.org/10.1136/bmj.14378)] [Medline: [31366597](https://pubmed.ncbi.nlm.nih.gov/31366597/)]
66. Sun H. Impact of exergames on physical activity and motivation in elementary school students: a follow-up study. *J Sport Health Sci* 2013 Sep;2(3):138-145. [doi: [10.1016/j.jshs.2013.02.003](https://doi.org/10.1016/j.jshs.2013.02.003)]
67. Latkin CA, Edwards C, Davey-Rothwell MA, Tobin KE. The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in Baltimore, Maryland. *Addict Behav* 2017 Oct;73:133-136. [doi: [10.1016/j.addbeh.2017.05.005](https://doi.org/10.1016/j.addbeh.2017.05.005)] [Medline: [28511097](https://pubmed.ncbi.nlm.nih.gov/28511097/)]
68. Mason M, Cho Y, Rayo J, Gong Y, Harris M, Jiang Y. Technologies for medication adherence monitoring and technology assessment criteria: narrative review. *JMIR mHealth uHealth* 2022 Mar 10;10(3):e35157. [doi: [10.2196/35157](https://doi.org/10.2196/35157)] [Medline: [35266873](https://pubmed.ncbi.nlm.nih.gov/35266873/)]
69. Li YC, Graham JD, Chirico D, Cairney J. Time-of-day effect on motor coordination in youth. *Chronobiol Int* 2022 Jun;39(6):761-768. [doi: [10.1080/07420528.2022.2033761](https://doi.org/10.1080/07420528.2022.2033761)] [Medline: [35189761](https://pubmed.ncbi.nlm.nih.gov/35189761/)]
70. Parker RA, Weir CJ. Multiple secondary outcome analyses: precise interpretation is important. *Trials* 2022 Jan 10;23(1):27. [doi: [10.1186/s13063-021-05975-2](https://doi.org/10.1186/s13063-021-05975-2)] [Medline: [35012627](https://pubmed.ncbi.nlm.nih.gov/35012627/)]

Abbreviations

BBT: Box and Block Test

DCD: developmental coordination disorder

DCDQ: Developmental Coordination Disorder Questionnaire

DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*

JARS-Quant: Journal Article Reporting Standards for Quantitative Research

MABC-2: *Movement Assessment Battery for Children, Second Edition*

MABC-3: *Movement Assessment Battery for Children, Third Edition*

PACES: Physical Activity Enjoyment Scale

PPI: Patient and Public Involvement

VR: virtual reality

Edited by S Brini; submitted 29.Sep.2025; peer-reviewed by M Lemay, OCY Ho; accepted 21.Jan.2026; published 10.Mar.2026.

Please cite as:

Alharbi M, Harris D, Dodd H, Wood G, Buckingham G

Home-Based Immersive Virtual Reality to Improve Motor Performance in Children and Adolescents With Developmental Coordination Disorder: Crossover Study

JMIR Serious Games 2026;14:e84995

URL: <https://games.jmir.org/2026/1/e84995>

doi: [10.2196/84995](https://doi.org/10.2196/84995)

© Mohammed Alharbi, David Harris, Helen Dodd, Greg Wood, Gavin Buckingham. Originally published in JMIR Serious Games (<https://games.jmir.org>), 10.Mar.2026. 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.

Original Paper

User Experience of a Virtual Reality–Based Treadmill for Children With a Chronic Disease Affecting Physical Health: Cross-Sectional Feasibility Study

Capucine Hennequin^{1,2}, MD; Lena Carcreff^{3,4,5}, PhD; Adélie Christiaens^{3,5}, MD; Mickaël Dinomais^{3,5}, MD, PhD; Josselin Demas^{5,6}, PhD

¹Faculty of Medicine, University of Rouen Normandy, Rouen, France

²Rouen University Hospital, Rouen, France

³Department of Physical and Rehabilitation Medicine, Angers University Hospital, Angers, France

⁴Angers Living Lab in Hospital Geriatrics, Department of Geriatric Medicine and Rehabilitation, Angers University Hospital, Angers, France

⁵Angevin Research Laboratory for Systems Engineering, EA7315, University of Angers, Angers, France

⁶Training Institutes, Laval Hospital Center, Laval, France

Corresponding Author:

Lena Carcreff, PhD

Department of Physical and Rehabilitation Medicine

Angers University Hospital

4 rue Larrey

Angers, 49100

France

Phone: 33 2 41 35 36 37

Email: lena.carcreff@chu-angers.fr

Abstract

Background: For children with chronic conditions affecting physical health and who require long-term care, the use of a connected treadmill for gait training as part of a home program can be a way to promote motivation in rehabilitation. Furthermore, the device must be evaluated by all user groups to ensure that its development best meets the rehabilitation needs of children.

Objective: The study aimed to assess the user experience of a connected treadmill called *Amy*—with both immersive and nonimmersive virtual reality—among children with a chronic disease impacting physical health, as well as their parents and therapists, to explore the feasibility and potential of such a device for home-based rehabilitation in this population.

Methods: Children with cerebral palsy, neuromuscular diseases, or obesity, along with one of their parents and rehabilitation therapists, were recruited. The study involved evaluating preexisting *Amy* solutions and collecting user experience feedback from participants with questionnaires. *Amy* solutions consisted of immersive virtual reality (using a virtual reality headset) and nonimmersive (tablet-based) games, both controlled through body movements on a treadmill conceived to train walking and balance. Questionnaires were the short version of the User Experience Questionnaire; the Usability Metric for User Experience; the Virtual Reality Sickness Questionnaire; a customized questionnaire evaluating comfort, fun, sense of presence and immersion; and a customized questionnaire evaluating parent's perception.

Results: Twenty-eight children, 28 parents, and 18 therapists participated in the study. Compared with User Experience Questionnaire benchmark data, the overall results with immersive and nonimmersive virtual reality in all participants were in the range of 10% best results or in the “excellent” category. The mean (95% CI) scores for each group of participants, with nonimmersive and immersive virtual reality, were as follows: 1.9 (1.6-2.2) and 2.1 (1.6-2.5) for children, 2.0 (1.7-2.2) and 2.3 (2.0-2.5) for parents, and 1.4 (1.1-1.7) and 1.4 (1.1-1.7) for therapists, respectively. User experience was significantly better for children and parents than for therapists (P adjusted .001). From the Usability Metric for User Experience, participants rated the *Amy* treadmill's usability as “good to excellent” on the System Usability Scale, regardless of whether immersive virtual reality was used. Immersive virtual reality was well tolerated by children. Children experienced immersive virtual reality positively in terms of comfort, immersion, presence, and fun. Parents' acceptability of the connected treadmill was positively assessed.

Conclusions: This study is, to our knowledge, the first to assess the user experience of a playful treadmill-based virtual environment controller in children with chronic conditions affecting physical health in a user-centered and multidisciplinary

team-based approach. This initial test demonstrates promising potential for using the connected treadmill as a rehabilitation tool. Therapists may need improvements to better meet their expectations, highlighting the importance of further iterations to align technological features and practical clinical context.

(*JMIR Serious Games* 2026;14:e82953) doi:[10.2196/82953](https://doi.org/10.2196/82953)

KEYWORDS

cerebral palsy; neuromuscular diseases; obesity; pediatrics; walking; rehabilitation; virtual reality; user-centered design

Introduction

Chronic diseases in children are a major public health issue. Over the past few years, their prevalence has increased, with 10%-30% of children affected by a chronic disease [1]. Cerebral palsy, neuromuscular diseases, and obesity are chronic conditions characterized by physical impairment, particularly affecting the musculoskeletal and cardiorespiratory systems [2-4]. This physical impairment affects the gross motor function and walking ability of these children. Indeed, the walking endurance is reduced in children with cerebral palsy compared with their typically developing peers, even among those with the highest functional abilities [5]. The walking endurance is also reduced in children with obesity or neuromuscular diseases [6,7]. Combined with personal and environmental factors, this physical impairment can therefore lead to activity limitations and participation restrictions, affecting the child's quality of life [2,4,6,8].

Functional gait training, including treadmill training, consists in endurance, balance, and strength training exercises. In children with cerebral palsy, neuromuscular diseases, or obesity, it can improve gross motor function, balance, and walking ability, as well as metabolic, cardiovascular, and respiratory parameters [9-14]. Gait training including treadmill training has shown promising results for physical health and functional capacities in children with chronic diseases [13-15]. Although there is no consensus due to the great heterogeneity in terms of duration, frequency, and intensity, it appears that longer programs, conducted over several weeks and with progressive intensity, are preferred [9,10,12-14,16]. Future research is required to recommend the best programs in terms of intensity and duration [11].

Rehabilitation is a huge part of clinical management in those 3 chronic diseases. Thus, maintaining motivation over a long time is a major objective. New technologies and rehabilitation methods, such as the combination of a treadmill with virtual reality, included in a personalized home-based program, could be an ideal way to promote motivation in rehabilitation [10,11]. Home-based programs are considered a useful and feasible complement to therapies in rehabilitation centers, which are sometimes located far from home [17,18]. They offer an opportunity to increase the intensity and repetition of the tasks being practiced, integrating them into children's familiar environment, while still providing a structured framework under the supervision of a therapist with regular feedback and have shown positive effects in improving motor function, cardiometabolic health, and self-care [11,12,19,20]. Virtual reality offers a real-time multisensory experience, promotes motor learning through task repetition and feedback, and enables

accurate and personalized performance monitoring, even remotely, thanks to digital devices. Also, virtual reality may enhance enjoyment for children, increase their engagement in physical activity, and improve adherence to rehabilitation programs [9,10,19,21-23]. Nevertheless, the wide variability in exercise modalities, technological supports, and modes of patient participation means that no clear guidelines are currently available to inform the design of new programs [18].

New technologies such as treadmills and virtual reality were initially developed for adults and must be adapted for children, especially for children with disabilities. *Amy* is a device created by the company EzyGain for adults but that can also be adjusted for children. *Amy* is a connected treadmill, meaning that the treadmill itself, through its use, enables interaction with a virtual environment to which it is linked. It was developed as a rehabilitation tool to train walking and balance at home, in an entertaining way, within a telerehabilitation model that includes regular monitoring by a therapist.

Before trying a treadmill equipped with virtual reality and integrating it in a home-based program, we must identify key priorities to assess the device to the target population. The user experience (UX) refers to a person's interaction with a system. UX is subjective and encompasses 2 main categories [24]. The first category focuses on the pragmatic and functional aspects of the system. The second category focuses on the emotional, hedonic, and aesthetic aspects of the system. These aspects will influence the user's emotional reactions, resulting in acceptance, positive or negative evaluation of the system, and the intention to use it. A positive UX is therefore essential for prolonged use, especially in future projects involving a home-based rehabilitation program lasting several weeks. UX involves not only the person actively using the system, here the child using the connected treadmill, but also anyone passively observing the user [25]. A multidisciplinary, team-based approach that involves parents and rehabilitation therapists, as well as children, provides the necessary perspectives for user-centered design [20,26]. Parental opinion is a crucial element in the success of an intervention, especially in a home-based program where they are directly involved [20,27]. The therapists' opinion is also essential. They will need to integrate the device into their therapeutic arsenal and clinical practice, conduct the remote rehabilitation program, personalize it for each child, and educate and coach parents [15,16,22]. Involving all end users from the earliest stages of development allows to work toward finding the best fit between a technological system—in this case, the connected treadmill—and the expectations of children, parents, and therapists, ultimately supporting optimal rehabilitation outcomes [28].

In the field of pediatric rehabilitation, existing literature mainly includes papers focusing on the UX of devices designed for cognitive [29-33] or upper limb rehabilitation [34,35]. To our knowledge, and to date, few technological devices targeting gait training in children have been codeveloped using a UX-driven approach. Ammann-Reiffer et al [23] studied the UX of a virtual reality headset as a rehabilitation tool for everyday walking activities. No study has explored the UX of a treadmill, combined with VR—that is, designed as a game controller—in children with chronic diseases affecting physical health in a user-centered design and a multidisciplinary, team-based approach.

The primary objective of this study was to investigate the UX of a treadmill connected to immersive and nonimmersive virtual reality among children with a chronic disease affecting physical health, as well as their parents and therapists. The secondary objectives were to (1) evaluate the usability of a connected treadmill among children, parents, and therapists; (2) evaluate cybersickness induced by immersive virtual reality in children; (3) evaluate the comfort, fun, sense of presence, and immersion experienced with immersive virtual reality in children; (4) evaluate the parent's perception of a connected treadmill; and (5) compare UX across the different groups as well as between immersive and nonimmersive virtual reality. We hypothesized that UX would be generally positive among the sample participants, with usability expected as an area for improvement. We also expected virtual reality to be well tolerated, and that comfort, fun, presence, and immersion would receive positive evaluations in immersive mode. Additionally, we anticipated parental heterogeneous acceptance of the system. Finally, we hypothesized that UX would be similar across the 3 groups but superior in immersive mode.

Methods

Study Design

A cross-sectional feasibility UX study was conducted at the Capucins Rehabilitation Centre and the University Hospital in Angers, France, in October and November 2024. During this initial user test, quantitative data were collected via questionnaires and qualitative data via focus groups. The subject of this paper focuses on quantitative data. The study adheres to the RATE-XR guideline recommended in early-stage evaluations of extended reality applications [36].

Ethical Considerations

The study was approved by the Angers University Hospital Ethics Committee (approval no. 2024-148). It was also declared to the National Commission for Information Technology and Civil Liberties (registration no. 349000031/559947). An information letter and an opposition form were provided for each participant either in person or via email. The information letters were tailored to specific age groups (those aged 8-10, 11-14, and 15-17 years and adults). Each child gave verbal assent to participate. Informed consent was obtained from parents on behalf of their child and for their own participation,

and from all participating therapists. Participation was voluntary without financial compensation, and individuals could withdraw at any time without consequences. All collected data were anonymized before analysis. All necessary measures have been taken to remove any identifying features from the images included in the manuscript. Written consent regarding the evaluator depicted in one of the figures has been obtained and provided to the journal.

Recruitment

Children with cerebral palsy, neuromuscular diseases, and overweight or obesity, along with one of their parents and therapists, were invited to participate in this study. Recruitment was based on children and parents meeting eligibility criteria and receiving medical follow-up from the pediatric rehabilitation departments of the 2 centers. Therapists in the relevant departments were invited to participate. Three groups were formed based on disease type: the "CP" group (children with cerebral palsy), the "OB" group (children with overweight or obesity), and the "NM" group (children with neuromuscular disease). Children aged 8-18 years who could walk on a treadmill for 5 minutes, with or without support or mobility aids, were included. Parents were included if their child met the study eligibility criteria. Therapists were included if they were accustomed to working with at least one of the study-eligible populations. All participants were required to (1) speak French fluently enough to complete questionnaires, and (2) provide informed consent. Children with uncontrolled epilepsy were excluded. Adults were excluded if deprived of liberty by judicial or administrative decision or if subject to legal protection measures.

Description of the Virtual Reality–Based Treadmill

Amy is a CE-marked (European Conformity) medical device designed for rehabilitation originally created for adults. It consists of a compact treadmill (65×130 cm) with a robust frame, height-adjustable support bars, an antifall harness for safety, and an integrated force plate for capturing weight shifts and controlling games through body movements. The treadmill supports walking speeds ranging from 0.2 to 3 km/h and accommodates users weighing up to 100 kg. A touch-sensitive tablet connects to the treadmill via Bluetooth Low Energy, enabling interaction games and facilitating remote communication with a therapist. Additionally, the *Amy* system is compatible with the Pico 4 virtual reality headset, allowing integration of immersive rehabilitation modules. The height of the support bars, the harness, and the tablet device stand can be adjusted to suit the child's size. The preexisting *Amy* solutions consisted of exergames aimed at training balance and gait for short periods. Exergames were composed of a sequence of mini games, each played one after the other. The progress and completion of each virtual reality game were made possible by the child's use of the treadmill. See Table 1 for description of the virtual reality games. Games were launched in an application called *EzyGain* app (version 4.7.3) available on the Apple store at the time of user testing.

Table 1. Description of the virtual reality games (nonexhaustive list).

Screenshot ^a	Motor task	Level of immersion	Game objective	Control modality	Game environment	Progression mechanics
	Gaits	Nonimmersive VR ^b	Mediolateral weight transfer while walking	Force plate	Snake-themed	Number of targets and obstacles
	Balance	Nonimmersive VR	Lateral weight shifting	Force plate	Pirate-themed	Number, speed, and size of targets
	Balance	Nonimmersive VR	Anteroposterior weight shifting	Force plate	Pong-inspired	Target speed
	Balance	Nonimmersive VR	Multidirectional weight shifting	Force plate	Memory game	Number and size of cards
	Squat	Nonimmersive VR	Lower limb muscle strengthening	Force plate	Frog-jumping	Height of stages
	Gait	Immersive VR	Visuomotor target aiming while walking	Speed of the treadmill and VR headset	Forest	Fixed task difficulty
	Gait	Immersive VR	Visual exploration while walking	VR headset	Real-world environment (eg, Paris)	Fixed task difficulty

^aScreenshots provided by EzyGain.

^bVR: virtual reality.

Procedure

The study involved evaluating preexisting *Amy* solutions and collecting UX feedback from participants (Figure 1). Each child participated in 2 consecutive 20-minute sessions the same day, accompanied by one of their parents and a therapist. Each child was first assisted in putting on the safety harness to ensure secure use of the treadmill. Once properly equipped, they stepped onto the treadmill and tried out different games. Between games, the treadmill was stopped so that the child could rest, if necessary. During the test sessions, parents and therapists were purely observers (Figure 2). An investigator was present in the room to supervise, in particular, to help the child

settle in and start playing games. In the first session, children used the *Amy* treadmill combined with the tablet-based interface, experiencing 4 nonimmersive (2D) virtual reality games (see examples in Table 1). After completing this session, all participants filled in a set of UX questionnaires. In the second session, children used the *Amy* treadmill with the virtual reality headset, experiencing 2 immersive (3D) virtual reality games (Table 1). Following this session, participants again completed a set of UX questionnaires. While the specific games differed between participants, all children were exposed to a range of game types addressing various motor objectives, as summarized in Table 1.

Figure 1. Study participation timeline: testing of a connected treadmill with nonimmersive and immersive virtual reality by children with chronic conditions affecting their physical health and gathering of their user experience as well as that of their parents and therapists. UEQ: User Experience Questionnaire; UMUX: Usability Metric for User Experience; VRSQ: Virtual Reality Sickness Questionnaire.

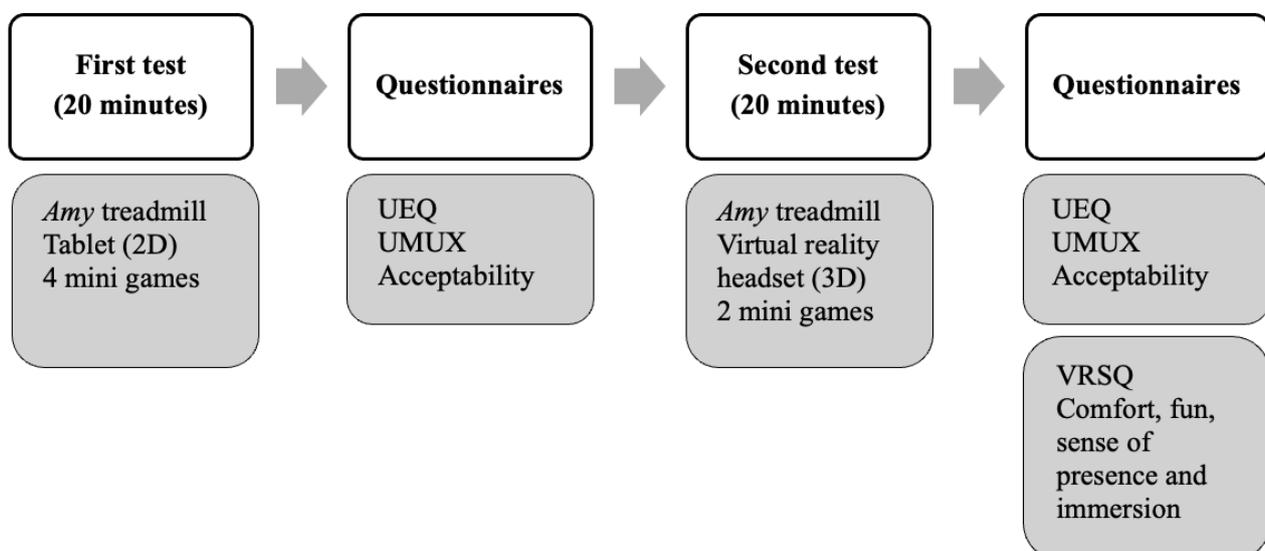


Figure 2. Test session set up. The Amy-connected treadmill being tested by a child with obesity, observed by his parent and a therapist. An investigator is present alongside the child to supervise the session.



To minimize the Hawthorne effect and social desirability bias, participants were explicitly informed that the investigators' role was limited to data collection and that responses should reflect their personal perceptions and experiences. An investigator was present to support the children during questionnaire completion, providing clarification when necessary, given that the questionnaires were not yet validated for this young population. Throughout the entire process, participants provided informal verbal feedback and comments, which contributed to enriching the discussion of this study.

UX Assessment

Ease with technology was evaluated by a single question: "On a scale of 1 to 10, how comfortable would you say you are with technology?" This single question provided a rough estimate of participants' level of acceptance of technologies of all kinds.

UX was evaluated using the French version of the User Experience Questionnaire (UEQ) [37]. The UEQ allows for a quick assessment of UX. The short version of the UEQ was used, which consists of 8 items grouped into 2 subscales: pragmatic quality and hedonic quality, each with 4 items. Pragmatic quality assesses efficiency, perspicuity, and dependability; hedonic quality assesses stimulation and originality. Each item is formatted as a Likert scale semantic differential that represents opposites (eg, "simple-complicated") [37,38]. The range rates between -3 and 3 . All participants completed the UEQ's short version. The short version of the UEQ has been scientifically validated for adults in English (Cronbach α values of 0.85 for the pragmatic quality and 0.81 for the hedonic quality) [37].

Usability was evaluated using the French version of the Usability Metric for User Experience (UMUX). The UMUX is a

questionnaire of 4 items evaluating usability components: effectiveness, satisfaction, overall, and efficiency. Each item is a 7-point Likert scale (from 1 "strongly disagree" to 7 "strongly agree"). After being collected, data were converted in a range of 0-100 UMUX score following the author's guidelines [39]. The advantage of the UMUX is that it is highly correlated with the System Usability Scale (SUS), which is the most widely used usability questionnaire. The UMUX is shorter than the SUS and aligns more closely with the ISO (International Organization for Standardization) 9241-11 (1998) definition of usability and evaluates more specifically the usefulness of a system in addition to usability, unlike the SUS and the UEQ-S [39,40]. All participants completed the UMUX. The UMUX has been scientifically validated for adults in English (Cronbach α value of 0.94) [39].

Cybersickness was evaluated using the Virtual Reality Sickness Questionnaire (VRSQ). It measures motion sickness-like sensations resulting from the use of virtual reality headsets. Only the children completed the VRSQ. They assigned a severity score from 0 to 3 (none, mild, moderate, and severe) to each of the 9 VRSQ items. The 9 items are grouped into 2 subscales: oculomotor discomfort and disorientation. The total score reflects the average of the oculomotor discomfort and disorientation scores out of 100 [41]. The VRSQ has been scientifically validated for adults in English (Cronbach α values of 0.847 for the oculomotor component and 0.886 for the disorientation component) [41].

A customized questionnaire covering the aspects of comfort, fun, sense of presence, and immersion in the virtual environment with the virtual reality headset was proposed. This questionnaire was based on the study by Ammann-Reiffer et al [23]. Only the children completed the questionnaire. The questions were as

follows: “How comfortable did you feel during the game?” “How much did you enjoy performing the task?” “How present did you feel in the virtual environment?” “How much of a problem was it for you not to be able to see the real environment and your body?” Children answered on a visual analogue scale from 1 to 10.

Acceptability from the parent’s point of view was assessed by a customized questionnaire also inspired by the study by Ammann-Reiffer et al [23]. The items were as follows: “The child’s movement on the *Amy* treadmill was not impaired compared with his usual walking,” “The child required no more support than usual,” “The child required no external motivation to walk,” “The child’s involvement was greater than usual (at home/in the usual physiotherapy session),” “The child remained interested and did not seem bored,” and “The use of the treadmill could be a good complement to conventional physiotherapy approaches for this child.” Parents answered on a 7-point Likert scale (from 1 “strongly disagree” to 7 “strongly agree”).

Outcomes

The primary outcome was the mean scores on the UEQ-S overall and in the pragmatic and hedonic scales. The secondary outcomes were (1) the UMUX mean scores; (2) the VRSQ median scores; (3) responses to the customized questionnaire evaluating sensations of comfort, fun, sense of presence, and immersion; (4) responses to the customized questionnaire evaluating parent’s perception; and (5) comparisons between the different groups of patients and between nonimmersive and immersive virtual reality.

Sample Size

The sampling strategy was nonprobabilistic and purposive [42]; that is, participants were recruited according to specific criteria, using maximum variation sampling [43] to ensure diversity among participants and, consequently, a wide range of perspectives within the group. The following variability criteria were considered: (1) age of the children and therapists, (2) sex of the children (including both girls and boys in each group), and (3) variety of pathological conditions. The median number of participants in UX studies reported in the literature is 20 [44]. Given the diversity among participants, we expected to include 30 children and consequently 30 parents. We also aimed to include 15 therapists for a total of 75 participants.

Statistical Analysis

Scores of the standardized scales (UEQ, UMUX, and VRSQ) were calculated as recommended by the reference methods

[37,39,41,45]. Descriptive statistics reporting the mean and 95% CI were used to present UEQ-S and UMUX results. Values obtained on the UEQ below -0.8 correspond to a negative assessment, between -0.8 and 0.8 to a neutral assessment, and above 0.8 to a positive assessment [46]. Mean UEQ scores were compared with published reference values using a benchmark dataset created by the author, which contains evaluations of 452 products’ evaluation and 20191 responses [46]. This enabled the evaluated product to be placed in one of 5 categories: “excellent” (in the range of the 10% best results), “good” (10% of the results in the benchmark dataset are better and 75% of the results are worse), “above average” (25% of the results in the benchmark are better than the result for the evaluated product and 50% of the results are worse), “below average” (50% of the results in the benchmark are better than the result for the evaluated product and 25% of the results are worse), and “bad” (in the range of the 25% worst results) [45]. UMUX scores were used to position the usability of the solution on the SUS interpretation scale (“worst imaginable,” “poor,” “acceptable,” “good,” “excellent,” or “best imaginable”) [47]. Descriptive statistics reporting the median and quartiles were used to present results from the VRSQ and customized questionnaires (comfort with technology; comfort, fun, sense of presence, and immersion; and parents’ perception). These scores were calculated for each version of the product tested, that is, with nonimmersive (2D) and immersive (3D) virtual reality.

To compare between-group differences in mean or median scores, we used nonparametric statistical tests, given the low number of participants and the data distribution. Wilcoxon test was used for paired samples. Kruskal-Wallis test was used to compare more than 2 samples. *P* value of less than .05 was considered statistically significant. Bonferroni correction was used to adjust the *P* value in the case of multiple statistical tests.

Results

Demographics

Twenty-eight children (13 females; 15 males; mean age 11.9, SD 2.7 years), 28 parents, and 18 therapists participated in the study. Eleven children had cerebral palsy, 9 had obesity, and 8 had neuromuscular disease. Missing data were observed for secondary descriptive variables only (mobility aid and school level). No formal test for missing completely at random was performed. Demographics for the participants are shown in Table 2.

Table 2. Descriptive characteristics of the study participants (N=74) presented to provide context for the user experience data collected from children (n=28, part A) with chronic conditions trying a connected treadmill, as well as their parents (n=28, part B) and therapists (n=18, part C).

Characteristics	Participants
Part A: children (n=28)	
Diagnosis, n (%)	
Cerebral palsy	11 (39)
Obesity	9 (32)
Neuromuscular diseases	8 (29)
Age (years), mean (SD)	
Overall	11.9 (2.7)
Cerebral palsy	11.3 (2.9)
Obesity	13 (2.7)
Neuromuscular diseases	11.5 (2.1)
Gender, n (%)	
Women	13 (46)
Men	15 (54)
School level, n (%)	
Elementary school	8 (28)
Secondary school	12 (43)
High school	3 (11)
Medicoeducational institute	1 (3.5)
Vocational training	1 (3.5)
Missing values	3 (11)
Mobility aid on short distances, n (%)	
None	20 (71)
Ankle-foot orthosis	5 (18)
Crutches	1 (4)
Walker	2 (7)
Missing values	1 (4)
Ease with technology on a scale of 1-10, median (Q1^a-Q3^b)	
Overall	8 (6-9)
Part B: parents (n=28)	
Age (years), mean (SD)	
Overall	44.3 (7.5)
Gender, n (%)	
Women	21 (75)
Men	7 (25)
Ease with technology on a scale of 1-10, median (Q1-Q3)	
Overall	7 (7-8)
Part C: therapists (n=18)	
Age (years), mean (SD)	
Overall	31.6 (11.9)
Gender, n (%)	
Women	15 (83)

Characteristics	Participants
Men	3 (17)
Profession, n (%)	
Physiotherapist	5 (28)
Adapted physical activity teacher	4 (22)
Occupational therapist	8 (44)
Medical doctor	1 (6)
Ease with technology on a scale of 1-10, median (Q1-Q3)	
Overall	7 (5-7)

^aQ1: first quartile.

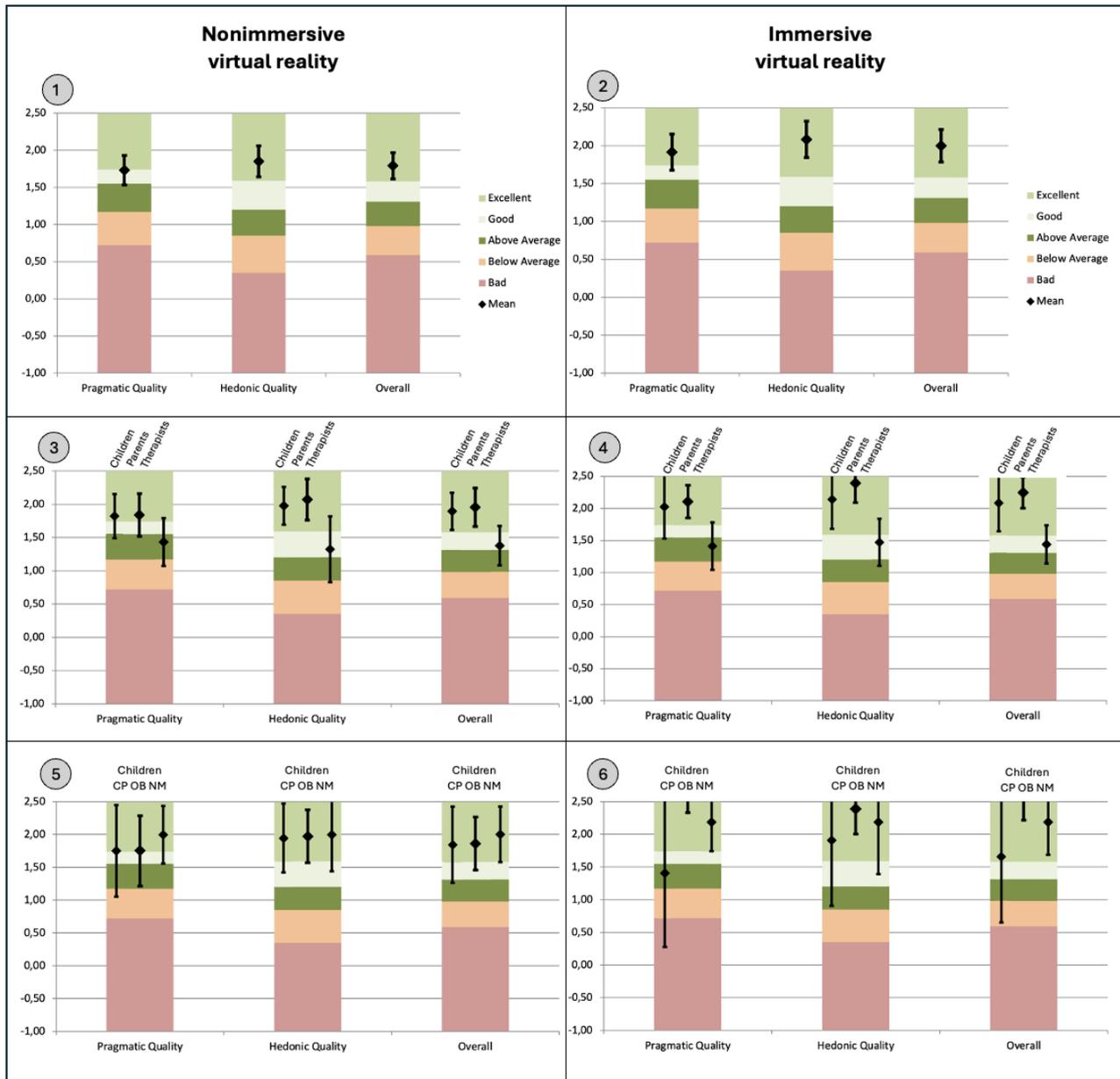
^bQ3: third quartile.

User Experience

The scores to the UEQ-S were all positive, mean values beyond 0.8 representing a positive evaluation according to the author [46]. For all participants, the scores for each scale, with nonimmersive and immersive virtual reality, were as follows: mean 1.8 (95% CI 1.6-2.0) and mean 2.0 (95% CI 1.8-2.2) for the overall UEQ-S, mean 1.7 (95% CI 1.5-1.9) and mean 1.9 (95% CI 1.7-2.2) for the pragmatic quality scale, and mean 1.9 (95% CI 1.6-2.1) and mean 2.1 (95% CI 1.8-2.3) for the hedonic quality scale, respectively. The overall UEQ-S, as well as the pragmatic and hedonic quality scales, were also positively evaluated in both immersive and nonimmersive virtual reality

separately in children, parents, and therapists, and in children's subgroups CP, NM, and OB. Mean (SD) UEQ-S responses are shown in Table S1 in [Multimedia Appendix 1](#). Compared with benchmark data, the overall scale means with immersive and nonimmersive virtual reality in all participants were in the range of 10% best results or in the "excellent" category ([Figure 3](#)). For all participants, there was a statistically significant difference in favor of immersive virtual reality for the overall UEQ-S ($P<.001$), as well as in the pragmatic ($P=.001$) and hedonic quality scales ($P<.001$). There was a significant difference in favor of immersive virtual reality for children ($P=.001$) and parents ($P<.001$) but not for therapists ($P=.41$).

Figure 3. Mean scores and 95% CI bars for the entire short version of the User Experience Questionnaire, pragmatic, and hedonic quality scales, relative to the benchmark. The device evaluated is the Amy treadmill, which incorporates both nonimmersive (left panels) and immersive (right panels) virtual reality. It was tested by children with chronic conditions. User experience data were collected from children, their parents, and therapists. Benchmark graphs come from the Excel tool created by the author [45]. (1) For all participants (N=74), when using nonimmersive virtual reality; (2) for all participants (N=74), when using immersive virtual reality; (3) for children (n=28), parents (n=28), and therapists (n=18), when using nonimmersive virtual reality; (4) for children (n=28), parents (n=28), and therapists (n=18), when using immersive virtual reality; (5) for children, in CP (n=11), OB (n=9), and NM (n=8) groups, when using nonimmersive virtual reality; and (6) for children, in CP (n=11), OB (n=9), and NM (n=8) groups, when using immersive virtual reality. The scale ranges from -3 (fully agree with the negative item) to +3 (fully agree with the positive item). The colored bars represent the ranges for the scale's mean values. CP: cerebral palsy; NM: neuromuscular; OB: obesity.



The overall UEQ-S was significantly better for children and parents than for therapists (P adjusted $<.001$). No difference was found between children and parents (P adjusted $>.99$). Compared with existing values from the benchmark dataset, the—overall, pragmatic, and hedonic quality—scale means with immersive and nonimmersive virtual reality in children and parents were in the “excellent” category. On the contrary, for therapists, the pragmatic quality scale means with immersive and nonimmersive virtual reality were in the “above average” category, and the overall and hedonic quality scale means with immersive and nonimmersive virtual reality were in the “good” category.

No statistical difference was found between the children’s subgroups ($P=.41$). A significant difference in favor of immersive virtual reality compared with nonimmersive virtual reality was found in the OB children’s subgroup ($P=.001$), as opposed to the CP group ($P=.80$) and NM group ($P=.10$).

Usability

Raw UMUX scores are shown in Table S2 in [Multimedia Appendix 1](#). Overall, participants rated the Amy treadmill’s usability as “good to excellent” on the SUS scale. Mean (95% CI) scores were 80 (78-82) in 2D and 79 (77-81) in 3D (Figure 4). There was no statistically significant difference between

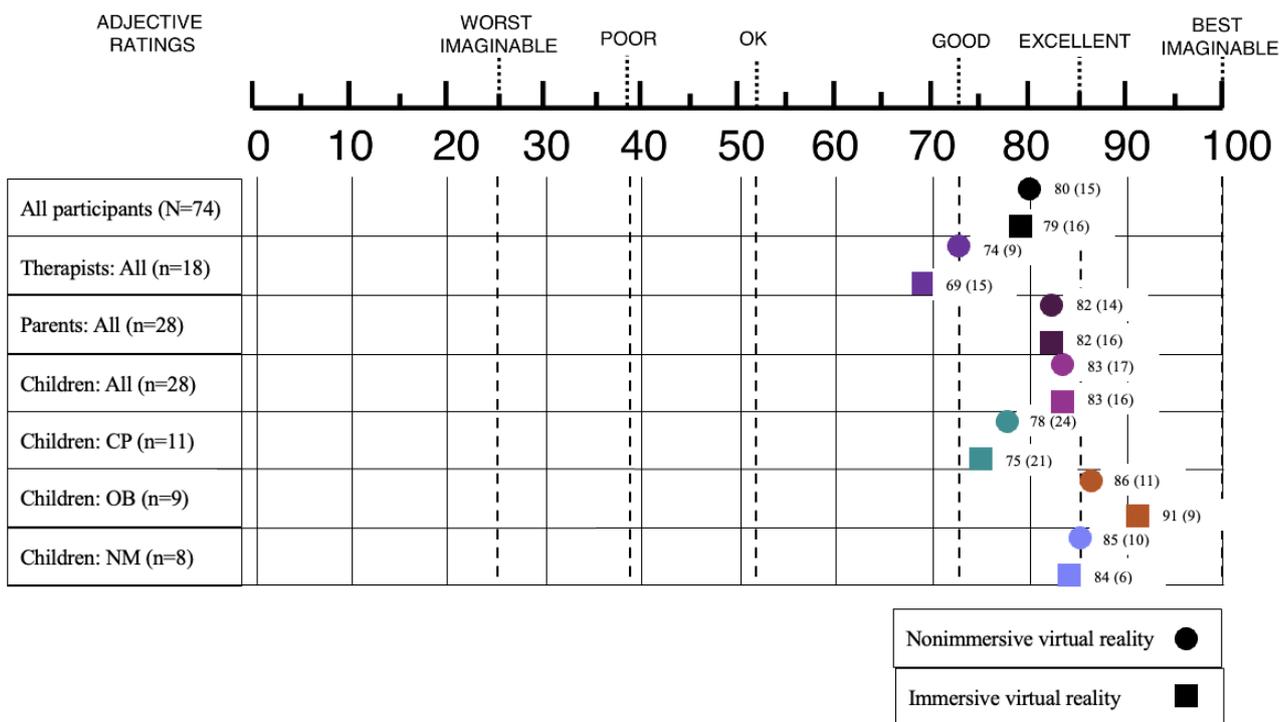
nonimmersive and immersive virtual reality for all participants ($P=.72$).

Children and parents reported similar positive experiences, mean (95% CI) scores being 83 (79-86) and 82 (79-84) in 2D, and 83 (80-86) and 82 (79-85) in 3D, respectively, that is “good to excellent” on the SUS scale. There was no statistically significant difference between nonimmersive and immersive virtual reality for either children ($P=.47$) or parents ($P=.60$), and there was no statistically significant difference between children and parents (P adjusted $>.99$). Therapists rated the system’s usability as “OK to good” on the SUS scale, with mean (95% CI) scores of 74 (72-76) in 2D and 69 (66-73) in 3D, with a significant difference in favor of nonimmersive virtual reality

($P<.02$). The usability was significantly better for children and parents than for therapists (P adjusted $<.001$).

Among children, those with obesity found the system more usable (mean 86, 95% CI 82-90 in 2D; mean 91, 95% CI 88-94 in 3D), followed by those with a neuromuscular disease (mean 85, 95% CI 82-89 in 2D; mean 84, 95% CI 82-86 in 3D) and cerebral palsy (mean 78, 95% CI 71-85 in 2D; mean 75, 95% CI 69-81 in 3D). A significant difference was found only between the OB and CP children’s subgroups (P adjusted $=.004$). A significant difference in favor of immersive virtual reality was found in the OB children’s subgroup ($P=.005$), as opposed to the CP group ($P=.50$) and NM group ($P=.30$).

Figure 4. Comparison of Usability Metric for User Experience scores converted to the System Usability Scale normative scale [47]. The score is between 0 and 100. Data are expressed as means (SDs). The device evaluated is the Amy treadmill, which incorporates both nonimmersive (rounds) and immersive (squares) virtual reality. It was tested by children with chronic conditions. Usability data were collected from children, their parents, and therapists. CP: cerebral palsy; NM: neuromuscular; OB: obesity.



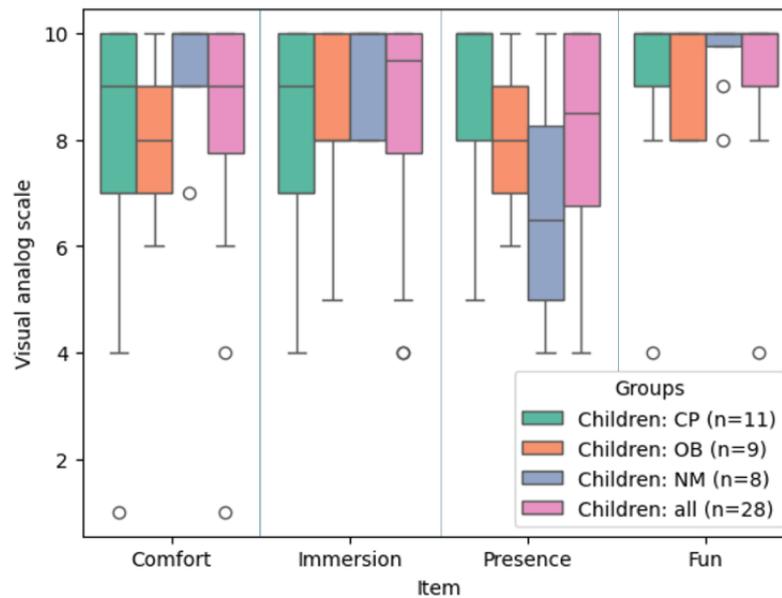
Cybersickness

Immersive virtual reality was well tolerated by children, apart from 1 participant in the cerebral palsy group. Overall, symptoms of cybersickness were minimal with a median total score on a 0-100 scale of 4 [3-12]. When present, these symptoms primarily involved oculomotor discomfort rather than disorientation, as indicated by the VRSQ, with a median of 8 (IQR 0-17) for oculomotor discomfort and 0 (IQR 0-7) for disorientation, respectively.

Comfort, Fun, Sense of Presence, and Immersion

Children experienced immersive virtual reality very positively in terms of comfort, immersion, presence, and fun (Figure 5). They felt comfortable during the game (median 9, IQR 8-10). The children reported no difficulty with immersion in the virtual environment (median 10, IQR 8-10). Children reported a strong sense of presence within the virtual environment (median 9, IQR 7-10) and expressed enjoyment while using the treadmill games (median 10, IQR 9-10). Results were similar in children with cerebral palsy, obesity, and neuromuscular disease ($P=.52$).

Figure 5. Box plots expressing the scores obtained from the customized questionnaire assessing children's comfort, immersion, sense of presence, and fun with immersive virtual reality on the Amy treadmill. The scale is a visual analogue scale from 1 to 10. CP: cerebral palsy; NM: neuromuscular; OB: obesity.

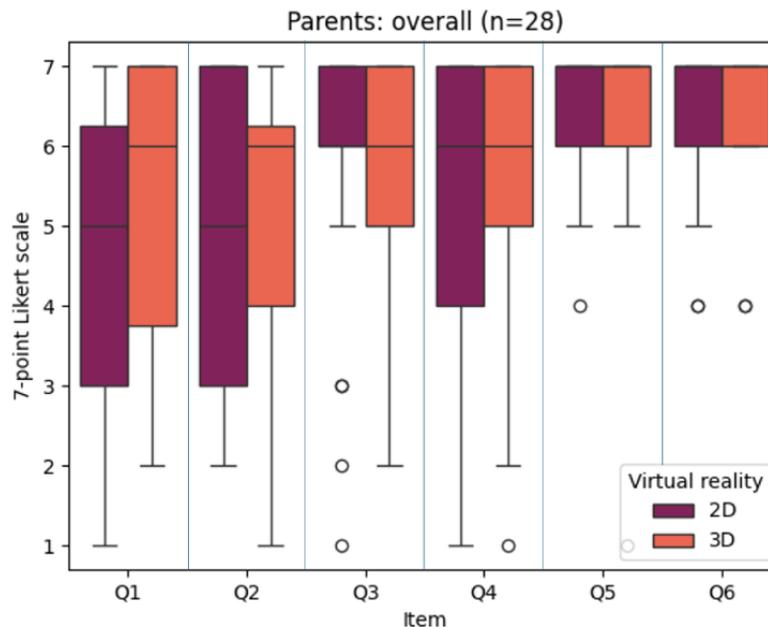


Acceptability

Parents' perception of the connected treadmill was positively assessed (Figure 6). According to the parents, out of 28 children, 8 (28%) exhibited impaired walking on the treadmill when using nonimmersive virtual reality versus 7 (25%) with immersive virtual reality. Parents reported that 8 children needed more support than usual when using nonimmersive virtual reality (8/28, 28%), compared with 5 children with immersive virtual reality (5/28, 18%). Most of the children did not require any external motivation to walk (24/28, 86% with nonimmersive

virtual reality and 25/28, 89% with immersive virtual reality). Their engagement seemed increased compared with conventional physiotherapy (24/28, 86% with nonimmersive and immersive virtual reality). Only 1 child seemed bored with immersive virtual reality. All the children remained interested with nonimmersive virtual reality. Apart from a neutral opinion, all parents felt that using the connected treadmill could be a good complement to conventional physiotherapy approaches for their child. Overall, for parents, results were similar between nonimmersive and immersive virtual reality ($P=.19$).

Figure 6. Box plots illustrating the scores obtained for each item in the customized questionnaire measuring parents' acceptability of their children's experience with the Amy treadmill under both nonimmersive and immersive virtual reality conditions. Q1: The child's movement on the Amy treadmill was not impaired compared with his usual walking. Q2: The child required no more support than usual. Q3: The child required no external motivation to walk. Q4: The child's involvement was greater than usual (at home or in the usual physiotherapy session). Q5: The child remained interested and did not seem bored. Q6: The use of the treadmill could be a good complement to conventional physiotherapy approaches for this child. The scale ranges from 1 (strongly disagree) to 7 (strongly agree). Q: quartile.



Discussion

Principal Findings

In this study, we evaluated the UX of a virtual reality–based treadmill as a rehabilitation tool in a sample of 74 individuals, including children with a chronic disease impacting physical health, their parents, and therapists. Participants reported an excellent UX from testing the connected treadmill about hedonic and pragmatic aspects. The findings showed that the virtual reality–based treadmill presented excellent usability. The immersive virtual reality system was well tolerated in terms of cybersickness; feelings of comfort, fun, sense of presence, and immersion were highly positive. For all participants, UX was better with the immersive virtual reality system than with the nonimmersive virtual reality system. The virtual reality–based treadmill was largely perceived positively by the parents. Children and parents reported greater UX than therapists.

Interpretation and Comparison With Prior Work

The overall UX evaluation of the virtual reality–based treadmill was excellent, as shown by the results of the UEQ-S. Compared with the benchmark reference data, the results on the connected treadmill were in the top 10%, indicating a product of excellent quality [45]. The pragmatic qualities of the connected treadmill were evaluated positively in terms of dependability, perspicuity, and efficiency. The virtual reality–based treadmill was easy to understand and use, secure, and predictable. This was supported by the results of the UEQ and the UMUX. The UMUX also adds the dimension of usefulness, as the system has been assessed as meeting the children's needs. The hedonic qualities of the connected treadmill were evaluated positively in terms of stimulation and novelty. Good rates in the UEQ-S were

reported on interest, fun, motivation, innovation. Feelings of comfort, immersion, sense of presence, and fun with immersive virtual reality were positively rated by all children. Moreover, the virtual reality headset was well tolerated. Only 1 child, aged 9 years, showed important signs of cybersickness, especially oculomotor discomfort. These results are in line with those reported by Ammann-Reiffer et al [23], who investigated the UX of an immersive virtual reality system to train everyday walking activities in children with a neurologic disease. This virtual reality–based treadmill, designed as a rehabilitation tool for gait and balance training through direct interaction between the treadmill and virtual reality games, generated a positive UX in terms of pragmatic and hedonic qualities, with positive emotions reported in all participants. Control of both immersive and nonimmersive environments is achieved directly through the treadmill, via weight shifts under static and dynamic conditions and adjustments in treadmill speed, providing a convincing UX. Therefore, it has the potential to be adopted by all end users and to motivate children to participate in physical rehabilitation, either more often or for an extended period. As previously showed in the study by Phelan et al [34], who tested immersive virtual reality in children with upper limb injuries, being entertained by virtual reality can increase children's participation to rehabilitation, as the fun induced by virtual reality keeps them distracted from the difficulties of exercising. This has to be confirmed by a subsequent feasibility study with long-term use of the device in the target environment, that is, at home.

From the parents' perspective, only a few children exhibited gait disturbances and required additional support (from the bars or the harness), both with nonimmersive and immersive virtual reality. Parents felt that their child was motivated, interested,

and engaged during the session, and they considered the use of the connected treadmill to be a valuable complement to conventional rehabilitation methods. The findings align with those of Phelan et al [34], where parents reported a positive UX and showed interest in using the virtual reality system at home. For all participants, there was a statistically significant difference in favor of immersive virtual reality in the UEQ but not in the UMUX questionnaire. The difference could therefore lie in the hedonic aspects that the UEQ investigates unlike the UMUX [40]. According to Winter et al [48], who tested the UX of treadmill training with semi-immersive and immersive virtual reality for patients with gait disorders, better results are explained by the higher sense of presence experienced in the immersive condition compared with the semi-immersive condition.

While the results were positive overall, they were significantly better in children and parents than in therapists, in both the UEQ and the UMUX. Possible explanations could be, on the one hand, functional—relating to the technical aspects of the treadmill—and, on the other hand, associated with the games, such as their design, interactivity, and coordination with the treadmill. Some participants complained about the mismatch between the speed of the games and the speed of the treadmill, which is limited to 3 km/hour. They reported some technical difficulties and simplistic esthetics of the different games. We hypothesize that during this first session, therapists anticipated using the connected treadmill in a more complex situation, that is, at home, under supervision, as part of a rehabilitation program lasting several weeks. As a result, and because of their technical knowledge, they may have had higher expectations of the virtual reality-based treadmill. Given their role as therapists, they may be looking for a device that is safe, reliable, as unobtrusive as possible so that it can be integrated into their workload, and as personalized as possible for patients. The engagement of therapists remains essential for the future implementation of a virtual reality rehabilitation device [28]. In the study by Phelan et al [34], therapists report that immersive virtual reality in children with upper limb injuries can be a good alternative rehabilitation tool and can promote positive interaction between patients and therapists. However, this is subject to several conditions: the system must be customizable to each patient, activities must be varied and progressive, and therapists must be properly trained [34]. Ammann-Reiffer et al [23] reported favorable results in terms of therapists' acceptability. They found that children were focused and motivated and, above all, were not restricted in their movements by the virtual reality system.

The results were in favor of some difference between the children's subgroups. Children with obesity showed better usability results than other children, especially with immersive virtual reality. Children with cerebral palsy and neuromuscular diseases were younger and more functionally limited than children with obesity. None of the obese children required a mobility aid. We hypothesize that younger age and greater functional impairment could make it more difficult to use the connected treadmill. Cerebral visual impairment, which is frequently observed among children with cerebral palsy [49], could account for the poorer results observed in this population,

although very limited data are available in the literature and further research is warranted.

Limitations

This study has several limitations. The participant group was heterogeneous, particularly among children, in terms of age and pathology. This enabled us to provide a range of opinions for a UX study but made comparisons difficult and is not intended to be generalized. Regarding UX, and especially usability, part of the virtual reality system was set up by the caregiver. As a result, children did not test the usability of the connected treadmill in its entirety. It could also be interesting to involve parents in the installation of the system in a future context of a home program where one of their roles would be to supervise the treadmill sessions. Ensuring usability is essential, especially for systems intended for long-term use. Two 20-minute sessions UX test were insufficient, and test sessions will need to be repeated throughout the design and improvement process to confirm the trend of these initial results and to project long-term use.

Acceptability from the therapists' point of view could not be assessed, as it was not possible to match the children with their respective therapists for logistical and organizational reasons. This could be interesting for future studies in order to supplement the results. Although the questionnaires have been translated into French and used in several French language studies and/or with children, they have not yet been formally psychometrically validated in an independently published study. The various items were not specifically adapted to the vocabulary and cognitive abilities of children. Some items of the customized questionnaires were possibly too complex or not clear enough, which may have led to inconsistent answers. This may have been the case in the parents' perception questionnaire, particularly for the first 2 questions, where a wide variety of responses was observed. Furthermore, the risk of the Hawthorne effect and social desirability bias on participants' responses to questionnaires cannot be ruled out. Methodologically, and in order to limit the order effect, it could have been valuable to randomize the 2 test sessions. It might have also been interesting to consider motivation for physical activity and the family's socioeconomic context, as these factors could potentially have an impact on subsequent use of the device in a several-week home rehabilitation program.

Future Directions

This initial UX session of the connected treadmill, conducted in a supervised environment, demonstrated promising potential within our sample. These findings represent a first step in the iterative codevelopment of the connected treadmill. The device and its UX will be refined based on user feedback to strengthen children's and parents' engagement and foster greater acceptance from therapists. Integrating the findings from both quantitative and qualitative studies may enhance accuracy through complementarity. Thus, future studies, quantitative and qualitative, will be necessary to assess UX after the company has developed the games and the treadmill—first in a supervised environment and subsequently in real-life situations in the children's homes. The objective, from a home program perspective, is to provide a comprehensive program featuring

numerous progressive games through an engaging interface, enabling children to participate actively while being supervised by therapists.

Conclusions

This is the first study to explore the UX of a treadmill designed as a playful game controller of virtual environments, in children with chronic diseases affecting physical health, in a user-centered design and a multidisciplinary, team-based approach. This initial test demonstrates promising potential for

using the connected treadmill as a rehabilitation tool. Improvements are required to better meet therapists' expectations. This underlines the importance of further iterations to achieve the best alignment between technological features and practical clinical context for those children. Consequently, UX will be reassessed after the games and the treadmill have been enhanced, first in a supervised environment and subsequently in real-life settings within the children's homes, through extended-duration trials.

Acknowledgments

The authors would like to express their gratitude to all the study participants for their support. The study received funding from France 2030 as part of the Tiers-Lieu MOBIS initiative (CHU ANGERS, France and Caisse des Dépôts). The authors declare the use of generative artificial intelligence (GAI) in the writing process. According to the GAIDeT taxonomy (2025), translation, proofreading, and editing were delegated to GAI tools under full human supervision. The GAI tool used was Perplexity AI (2025 version). Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

The study received funding from France 2030 as part of the Tiers-Lieu MOBIS initiative (CHU ANGERS, France, and Caisse des Dépôts).

Data Availability

The datasets generated during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: LC, MD, JD

Methodology: LC, MD, JD

Investigation: CH, LC, JD

Formal analysis: CH, LC

Data curation: CH, LC

Visualization: CH

Writing – original draft: CH

Writing – review & editing: CH, LC, AC, MD, JD

Funding acquisition: LC, MD, JD

Resources: LC, AC, MD, JD

Project administration: LC, MD, JD

Supervision: LC, MD, JD

Conflicts of Interest

The authors declare no conflicts of interest. The developers of the treadmill and virtual reality solution (EzyGain) were involved in the study design but not in the data collection, analysis, or writing of this paper. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Multimedia Appendix 1

Detail of the scores obtained for each item of the short version of the User Experience Questionnaire and the Usability Metric for User Experience.

[[DOCX File , 32 KB - games_v14i1e82953_app1.docx](#)]

References

1. Van Cleave J, Gortmaker SL, Perrin JM. Dynamics of obesity and chronic health conditions among children and youth. *JAMA* 2010;303(7):623-630. [doi: [10.1001/jama.2010.104](https://doi.org/10.1001/jama.2010.104)] [Medline: [20159870](https://pubmed.ncbi.nlm.nih.gov/20159870/)]
2. Graham HK, Rosenbaum P, Paneth N, Dan B, Lin JP, Damiano DL, et al. Cerebral palsy. *Nat Rev Dis Primers* 2016;2:15082 [[FREE Full text](#)] [doi: [10.1038/nrdp.2015.82](https://doi.org/10.1038/nrdp.2015.82)] [Medline: [27188686](https://pubmed.ncbi.nlm.nih.gov/27188686/)]

3. Daniels SR. The consequences of childhood overweight and obesity. *Future Child* 2006;16(1):47-67. [doi: [10.1353/foc.2006.0004](https://doi.org/10.1353/foc.2006.0004)] [Medline: [16532658](https://pubmed.ncbi.nlm.nih.gov/16532658/)]
4. Ryder S, Leadley RM, Armstrong N, Westwood M, de Kock S, Butt T, et al. The burden, epidemiology, costs and treatment for Duchenne muscular dystrophy: an evidence review. *Orphanet J Rare Dis* 2017;12(1):79 [FREE Full text] [doi: [10.1186/s13023-017-0631-3](https://doi.org/10.1186/s13023-017-0631-3)] [Medline: [28446219](https://pubmed.ncbi.nlm.nih.gov/28446219/)]
5. Fitzgerald D, Hickey C, Delahunt E, Walsh M, O'Brien T. Six-minute walk test in children with spastic cerebral palsy and children developing typically. *Pediatr Phys Ther* 2016;28(2):192-199. [doi: [10.1097/PEP.0000000000000224](https://doi.org/10.1097/PEP.0000000000000224)] [Medline: [26808959](https://pubmed.ncbi.nlm.nih.gov/26808959/)]
6. Kennedy RA, Carroll K, McGinley JL, Paterson KL. Walking and weakness in children: a narrative review of gait and functional ambulation in paediatric neuromuscular disease. *J Foot Ankle Res* 2020;13(1):10 [FREE Full text] [doi: [10.1186/s13047-020-0378-2](https://doi.org/10.1186/s13047-020-0378-2)] [Medline: [32122377](https://pubmed.ncbi.nlm.nih.gov/32122377/)]
7. Barros WMA, da Silva KG, Silva RKP, Souza APDS, da Silva ABJ, Silva MRM, et al. Effects of overweight/obesity on motor performance in children: a systematic review. *Front Endocrinol (Lausanne)* 2021;12:759165 [FREE Full text] [doi: [10.3389/fendo.2021.759165](https://doi.org/10.3389/fendo.2021.759165)] [Medline: [35126307](https://pubmed.ncbi.nlm.nih.gov/35126307/)]
8. World Health Organization. *International Classification of Functioning, Disability and Health: ICF*. Geneva: World Health Organization; 2001.
9. HAS. Rééducation et réadaptation de la fonction motrice de l'appareil locomoteur des personnes diagnostiquées de paralysie cérébrale [Rehabilitation and re-education of motor function of the locomotor system in people diagnosed with cerebral palsy]. *Good Practice Guidelines*. 2021. URL: <https://tinyurl.com/y3ja53j6> [accessed 2026-01-23]
10. Booth ATC, Buizer AI, Meyns P, Oude Lansink ILB, Steenbrink F, van der Krogt MM. The efficacy of functional gait training in children and young adults with cerebral palsy: a systematic review and meta-analysis. *Dev Med Child Neurol* 2018;60(9):866-883. [doi: [10.1111/dmcn.13708](https://doi.org/10.1111/dmcn.13708)] [Medline: [29512110](https://pubmed.ncbi.nlm.nih.gov/29512110/)]
11. Sudati IP, Sakzewski L, Fioroni Ribeiro da Silva C, Jackman M, Haddon M, Pool D, et al. Efficacy and threshold dose of intensive training targeting mobility for children with cerebral palsy: a systematic review and meta-analysis. *Dev Med Child Neurol* 2024;66(12):1542-1557 [FREE Full text] [doi: [10.1111/dmcn.16040](https://doi.org/10.1111/dmcn.16040)] [Medline: [39108099](https://pubmed.ncbi.nlm.nih.gov/39108099/)]
12. Novak I, McIntyre S, Morgan C, Campbell L, Dark L, Morton N, et al. A systematic review of interventions for children with cerebral palsy: state of the evidence. *Dev Med Child Neurol* 2013;55(10):885-910 [FREE Full text] [doi: [10.1111/dmcn.12246](https://doi.org/10.1111/dmcn.12246)] [Medline: [23962350](https://pubmed.ncbi.nlm.nih.gov/23962350/)]
13. Papadopoulou O, Desli E, Chrysochoou E, Kogias C, Liakos V, Sopiadou A, et al. Effect of an exercise intervention program on BMI and distance covered in the 6 min walking test (6MWT) in obese children. *Children (Basel)* 2024;11(12):1473 [FREE Full text] [doi: [10.3390/children11121473](https://doi.org/10.3390/children11121473)] [Medline: [39767902](https://pubmed.ncbi.nlm.nih.gov/39767902/)]
14. Sherief AEAA, Abd ElAziz HG, Ali MS. Efficacy of two intervention approaches on functional walking capacity and balance in children with Duchene muscular dystrophy. *J Musculoskelet Neuronal Interact* 2021;21(3):343-350 [FREE Full text] [Medline: [34465672](https://pubmed.ncbi.nlm.nih.gov/34465672/)]
15. Bektashi S, Monbaliu E, McIntyre S, Saloojee G, Hilberink SR, Tatishvili N, et al. Towards functional improvement of motor disorders associated with cerebral palsy. *Lancet Neurol* 2023;22(3):229-243. [doi: [10.1016/S1474-4422\(23\)00004-2](https://doi.org/10.1016/S1474-4422(23)00004-2)] [Medline: [36657477](https://pubmed.ncbi.nlm.nih.gov/36657477/)]
16. Hansen D, Hens W, Peeters S, Wittebrood C, Van Ussel S, Verleyen D, et al. Physical therapy as treatment for childhood obesity in primary health care: clinical recommendation from AXXON (Belgian Physical Therapy Association). *Phys Ther* 2016;96(6):850-864. [doi: [10.2522/ptj.20150206](https://doi.org/10.2522/ptj.20150206)] [Medline: [26637641](https://pubmed.ncbi.nlm.nih.gov/26637641/)]
17. Al-Mfarej D, Vojtech J, Roy S, Townsend E, Keysor JJ, Kuntz N, et al. A virtual reality exergame: clinician-guided breathing and relaxation for children with muscular dystrophy. 2023 IEEE Conf Virtual Real 3D User Interfaces Abstr Workshops 2023;2023:270-276 [FREE Full text] [doi: [10.1109/vrww58643.2023.00065](https://doi.org/10.1109/vrww58643.2023.00065)] [Medline: [38009078](https://pubmed.ncbi.nlm.nih.gov/38009078/)]
18. Beckers LWME, Geijen MME, Kleijnen J, A A Rameekers E, L A P Schnackers M, J E M Smeets R, et al. Feasibility and effectiveness of home-based therapy programmes for children with cerebral palsy: a systematic review. *BMJ Open* 2020;10(10):e035454 [FREE Full text] [doi: [10.1136/bmjopen-2019-035454](https://doi.org/10.1136/bmjopen-2019-035454)] [Medline: [33028544](https://pubmed.ncbi.nlm.nih.gov/33028544/)]
19. 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;13(11):724-733 [FREE Full text] [doi: [10.1111/ijpo.12438](https://doi.org/10.1111/ijpo.12438)] [Medline: [30027607](https://pubmed.ncbi.nlm.nih.gov/30027607/)]
20. Jackman M, Sakzewski L, Morgan C, Boyd RN, Brennan SE, Langdon K, et al. Interventions to improve physical function for children and young people with cerebral palsy: international clinical practice guideline. *Dev Med Child Neurol* 2022;64(5):536-549 [FREE Full text] [doi: [10.1111/dmcn.15055](https://doi.org/10.1111/dmcn.15055)] [Medline: [34549424](https://pubmed.ncbi.nlm.nih.gov/34549424/)]
21. Hao J, Huang B, Remis A, He Z. The application of virtual reality to home-based rehabilitation for children and adolescents with cerebral palsy: a systematic review and meta-analysis. *Physiother Theory Pract* 2024;40(7):1588-1608. [doi: [10.1080/09593985.2023.2184220](https://doi.org/10.1080/09593985.2023.2184220)] [Medline: [36847396](https://pubmed.ncbi.nlm.nih.gov/36847396/)]
22. Deutsch JE, Westcott McCoy S. Virtual reality and serious games in neurorehabilitation of children and adults: prevention, plasticity, and participation. *Pediatr Phys Ther* 2017;29 Suppl 3(Suppl 3 IV Step 2016 Conference Proceedings):S23-S36 [FREE Full text] [doi: [10.1097/PEP.0000000000000387](https://doi.org/10.1097/PEP.0000000000000387)] [Medline: [28654475](https://pubmed.ncbi.nlm.nih.gov/28654475/)]

23. Ammann-Reiffer C, Kläy A, Keller U. Virtual reality as a therapy tool for walking activities in pediatric neurorehabilitation: usability and user experience evaluation. *JMIR Serious Games* 2022;10(3):e38509 [FREE Full text] [doi: [10.2196/38509](https://doi.org/10.2196/38509)] [Medline: [35834316](https://pubmed.ncbi.nlm.nih.gov/35834316/)]
24. Lallemand C, Gronier G, Koenig V. L'expérience utilisateur : Un concept sans consensus ? Enquête sur le point de vue des professionnels. In: *Activités Humaines, Technologies et Bien-être*. Paris, France: ARPEGE Science Publishing; 2013:35-48.
25. Roto V. Demarcating User eXperience. : Springer Berlin Heidelberg; 2009 Presented at: Human-Computer Interaction - INTERACT 2009 12th IFIP TC 13 International Conference; 2009 August 24-28; Uppsala, Sweden p. 922-923. [doi: [10.1007/978-3-642-03658-3_112](https://doi.org/10.1007/978-3-642-03658-3_112)]
26. Levander XA, VanDerSchaaf H, Barragán VG, Choxi H, Hoffman A, Morgan E, et al. The role of human-centered design in healthcare innovation: a digital health equity case study. *J Gen Intern Med* 2024;39(4):690-695. [doi: [10.1007/s11606-023-08500-0](https://doi.org/10.1007/s11606-023-08500-0)] [Medline: [37973709](https://pubmed.ncbi.nlm.nih.gov/37973709/)]
27. Mota LAT, Silva MZ, Dos Santos M, Pfeifer LI. The processes and outcomes related to 'family-centred care' in neuromotor and functional rehabilitation contexts for children with cerebral palsy: a scoping review. *Child Care Health Dev* 2024;50(3):e13271. [doi: [10.1111/cch.13271](https://doi.org/10.1111/cch.13271)] [Medline: [38738842](https://pubmed.ncbi.nlm.nih.gov/38738842/)]
28. Proffitt R, Glegg S, Levac D, Lange B. End-user involvement in rehabilitation virtual reality implementation research. *J Enabling Technol* 2019;13(2):92-100 [FREE Full text] [doi: [10.1108/JET-10-2018-0050](https://doi.org/10.1108/JET-10-2018-0050)] [Medline: [31663039](https://pubmed.ncbi.nlm.nih.gov/31663039/)]
29. Balloufaud M, Boujut A, Marie R, Guinaldo A, Fourcade L, Hamonet-Torny J, et al. Augmented reality exergames for upcoming cognitive-motor rehabilitation: user-centered design approach and user experience of healthy children. *JMIR Rehabil Assist Technol* 2025;12:e69205 [FREE Full text] [doi: [10.2196/69205](https://doi.org/10.2196/69205)] [Medline: [39970421](https://pubmed.ncbi.nlm.nih.gov/39970421/)]
30. Single N, Mishelle Graham W, Kelson J, Sulikowski D. Effectiveness and user experience of immersive virtual reality in cognitive rehabilitation for attention-deficit/hyperactivity disorder: systematic review. *JMIR Ment Health* 2025;12:e71963 [FREE Full text] [doi: [10.2196/71963](https://doi.org/10.2196/71963)] [Medline: [41385780](https://pubmed.ncbi.nlm.nih.gov/41385780/)]
31. Rico-Olarte C, Pez DM, Keplinger S. User experience evaluations in rehabilitation video games for children: a systematic mapping of the literature. *Ger Med Data Sci Vis Bridg* 2017;243:13-17. [doi: [10.3233/978-1-61499-808-2-13](https://doi.org/10.3233/978-1-61499-808-2-13)]
32. Schmidt M, Fisher AP, Sensenbaugh J, Ling B, Rietta C, Babcock L, et al. User experience (re)design and evaluation of a self-guided, mobile health app for adolescents with mild traumatic brain injury. *J Form Des Learn* 2020;4(2):51-64 [FREE Full text] [doi: [10.1007/s41686-019-00038-x](https://doi.org/10.1007/s41686-019-00038-x)] [Medline: [33860150](https://pubmed.ncbi.nlm.nih.gov/33860150/)]
33. Shen J, Xiang H, Luna J, Grishchenko A, Patterson J, Strouse RV, et al. Virtual reality-based executive function rehabilitation system for children with traumatic brain injury: design and usability study. *JMIR Serious Games* 2020;8(3):e16947 [FREE Full text] [doi: [10.2196/16947](https://doi.org/10.2196/16947)] [Medline: [32447275](https://pubmed.ncbi.nlm.nih.gov/32447275/)]
34. Phelan I, Furness PJ, Dunn HD, Carrion-Plaza A, Matsangidou M, Dimitri P, et al. Immersive virtual reality in children with upper limb injuries: findings from a feasibility study. *J Pediatr Rehabil Med* 2021;14(3):401-414 [FREE Full text] [doi: [10.3233/PRM-190635](https://doi.org/10.3233/PRM-190635)] [Medline: [34151871](https://pubmed.ncbi.nlm.nih.gov/34151871/)]
35. Consales A, Biffi E, Nossa R, Pittaccio S, Lazzari F, Malosio M, et al. Inclusivity is child's play: pilot study on usability, acceptability and user experience of a sensory-motor PC game for children with cerebral palsy (GiocAbile). *Ital J Pediatr* 2024;50(1):263 [FREE Full text] [doi: [10.1186/s13052-024-01830-7](https://doi.org/10.1186/s13052-024-01830-7)] [Medline: [39707421](https://pubmed.ncbi.nlm.nih.gov/39707421/)]
36. Vlaka JH, Drop DLQ, Van Bommel J, Riva G, Wiederhold BK, Cipresso P, RATE-XR Expert Group. Reporting guidelines for the early-phase clinical evaluation of applications using extended reality: RATE-XR qualitative study guideline. *J Med Internet Res* 2024;26:e56790 [FREE Full text] [doi: [10.2196/56790](https://doi.org/10.2196/56790)] [Medline: [39612482](https://pubmed.ncbi.nlm.nih.gov/39612482/)]
37. Schrepp M, Hinderks A, Thomaschewski J. Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). *IJIMAI* 2017;4(6):103-108. [doi: [10.9781/ijimai.2017.09.001](https://doi.org/10.9781/ijimai.2017.09.001)]
38. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. : Springer Berlin Heidelberg; 2008 Presented at: Symposium of the Austrian HCI and Usability Engineering Group. Vol 5298. Lecture Notes in Computer Science; 2008 November 20-21; Graz, Austria p. 63-76. [doi: [10.1007/978-3-540-89350-9_6](https://doi.org/10.1007/978-3-540-89350-9_6)]
39. Finstad K. The usability metric for user experience. *Interact Comput* 2010;22(5):323-327. [doi: [10.1016/j.intcom.2010.04.004](https://doi.org/10.1016/j.intcom.2010.04.004)]
40. Schrepp M, Kollmorgen J, Thomaschewski J. A comparison of SUS, UMUX-LITE, and UEQ-S. *J User Exp* 2023;18(2):86-104. [doi: [10.5555/3604890.3604893](https://doi.org/10.5555/3604890.3604893)]
41. Kim HK, Park J, Choi Y, Choe M. Virtual Reality Sickness Questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. *Appl Ergon* 2018;69:66-73. [doi: [10.1016/j.apergo.2017.12.016](https://doi.org/10.1016/j.apergo.2017.12.016)] [Medline: [29477332](https://pubmed.ncbi.nlm.nih.gov/29477332/)]
42. Savoie-Zajc L. Comment peut-on construire un échantillonnage scientifiquement valide? [How can a scientifically valid sampling be constructed?]. *Recherches Qualitatives* 2007;5(1):99-111 [FREE Full text] [doi: [10.7202/1118446ar](https://doi.org/10.7202/1118446ar)]
43. Patton MQ. Enhancing the quality and credibility of qualitative analysis. *Health Serv Res* 1999;34(5 Pt 2):1189-1208 [FREE Full text] [Medline: [10591279](https://pubmed.ncbi.nlm.nih.gov/10591279/)]
44. Díaz-Oreiro I, López G, Quesada L, Guerrero L. Standardized questionnaires for user experience evaluation: a systematic literature review. *Proceedings* 2019;31(1):14. [doi: [10.3390/proceedings2019031014](https://doi.org/10.3390/proceedings2019031014)]
45. Schrepp M, Hinderks A, Thomaschewski J. Construction of a benchmark for the User Experience Questionnaire (UEQ). *IJIMAI* 2017;4(4):40-44. [doi: [10.9781/ijimai.2017.445](https://doi.org/10.9781/ijimai.2017.445)]
46. Schrepp M. User experience questionnaire handbook. Version 11. 2023. URL: <https://www.ueq-online.org/material/handbook.pdf> [accessed 2026-01-23]

47. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an Adjective Rating Scale. *J Usability Stud* 2009;4(3):114-123. [doi: [10.5555/2835587.2835589](https://doi.org/10.5555/2835587.2835589)]
48. Winter C, Kern F, Gall D, Latoschik ME, Pauli P, Käthner I. Immersive virtual reality during gait rehabilitation increases walking speed and motivation: a usability evaluation with healthy participants and patients with multiple sclerosis and stroke. *J Neuroeng Rehabil* 2021;18(1):68 [FREE Full text] [doi: [10.1186/s12984-021-00848-w](https://doi.org/10.1186/s12984-021-00848-w)] [Medline: [33888148](https://pubmed.ncbi.nlm.nih.gov/33888148/)]
49. Dufresne D, Dagenais L, Shevell MI, REPACQ Consortium. Spectrum of visual disorders in a population-based cerebral palsy cohort. *Pediatr Neurol* 2014;50(4):324-328. [doi: [10.1016/j.pediatrneurol.2013.11.022](https://doi.org/10.1016/j.pediatrneurol.2013.11.022)] [Medline: [24468636](https://pubmed.ncbi.nlm.nih.gov/24468636/)]

Abbreviations

CE: European Conformity
CP: cerebral palsy
ISO: International Organization for Standardization
NM: neuromuscular
OB: obesity
SUS: System Usability Scale
UEQ: User Experience Questionnaire
UMUX: Usability Metric for User Experience
UX: user experience
VRSQ: Virtual Reality Sickness Questionnaire

Edited by S Brini; submitted 02.Sep.2025; peer-reviewed by C Pons, CJR Siah; comments to author 16.Oct.2025; accepted 22.Jan.2026; published 29.Apr.2026.

Please cite as:

Hennequin C, Carcreff L, Christiaens A, Dinomais M, Demas J

User Experience of a Virtual Reality–Based Treadmill for Children With a Chronic Disease Affecting Physical Health: Cross-Sectional Feasibility Study

JMIR Serious Games 2026;14:e82953

URL: <https://games.jmir.org/2026/1/e82953>

doi: [10.2196/82953](https://doi.org/10.2196/82953)

PMID:

©Capucine Hennequin, Lena Carcreff, Adélie Christiaens, Mickaël Dinomais, Josselin Demas. Originally published in JMIR Serious Games (<https://games.jmir.org>), 29.Apr.2026. 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.

Analysis of Spatiotemporal Features in a Virtual Navigation Game Across Different Age Groups: Quantitative Research

Xiaofeng Qiao^{1,2,3,4}, BEng; Shan Tian^{1,2,3,4*}, PhD; Min Tang^{1,2,3,4}, MEng; Shipei He^{1,2,3,4}, MSc; Jinghui Wang^{1,2,3,4}, MEng; Linyuan Fan^{1,2,3,4}, BEng; Yuanjie Zhu^{1,2,3,4}, MEng; Zhiyang Zhang^{1,2,3,4}, MD; Songjun Du^{1,2,3,4}, BEng; Chaojie Dong^{1,2,3,4}, BEng; Yepu Chen^{1,2,3,4}, BD; Xiaoyu Liu^{1,2,3,4*}, PhD

¹School of Biological Science and Medical Engineering, Beihang University, 5 Building in Beihang University, 37 Xueyuan Road, Beijing, China

²Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education, Beihang University, Beijing, China

³Key Laboratory of Innovation and Transformation of Advanced Medical Devices, Ministry of Industry and Information Technology, Beijing, China

⁴National Medical Innovation Platform for Industry-Education Integration in Advanced Medical Devices (Interdiscipline of Medicine and Engineering), Beijing, China

*these authors contributed equally

Corresponding Author:

Xiaoyu Liu, PhD

School of Biological Science and Medical Engineering, Beihang University, 5 Building in Beihang University, 37 Xueyuan Road, Beijing, China

Abstract

Background: With rapid urbanization, the proliferation of densely arranged buildings and increasingly homogeneous architectural designs has made disorientation and navigation difficulties more common, especially for older adults. Meanwhile, advances in virtual reality technology now allow researchers to create highly immersive navigation games, offering opportunities for assessing cognitive abilities and examining how environmental factors shape navigation behavior.

Objective: This study aimed to design a virtual reality-based navigation game capable of assessing cognitive abilities through navigation behavior and quantitatively examining how environmental configurations influence navigation patterns in different age groups.

Methods: We designed a virtual goal-directed navigation game and recruited 2 groups, younger adults (n=18) and older adults (n=21), to complete identical wayfinding tasks. Before the formal experiment, participants completed cognitive assessments and received training. To characterize navigational behavior, k-means clustering was applied to classify navigation states and extract behaviorally meaningful navigation measurements, which were then examined for correlations with cognitive test scores. To quantify the effects of environmental structure, space syntax analysis was conducted to calculate line-based and grid-based experienced metrics for each participant, and their associations with navigation performance were examined. Additionally, between-group differences in navigation performance and experienced metrics were evaluated across age groups.

Results: Our results revealed that navigation behavior performance, particularly navigation efficiency, was significantly influenced by cognitive abilities and was strongly associated with several cognitive tests: the Montreal Cognitive Assessment ($r=0.495$, $P=.04$), Trail Making Test Part A ($r=-0.761$, $P=.001$), and the Mental Rotation Test ($r=0.848$, $P<.001$). In terms of environmental influences, experienced axial integration (EAI) and experienced visual integration (EVI) demonstrated significant age-related differences: EAI ($z=-2.43$, $P=.01$) and EVI ($t=2.48$, $P=.02$). Moreover, navigation efficiency exhibited distinct age-specific correlations with experienced metrics: among older adults, navigation efficiency was negatively associated with EVI ($r=-0.48$, $P=.04$), and young adults showed negative correlation between navigation efficiency and EAI ($r=-0.64$, $P=.005$).

Conclusions: Our findings demonstrate that k-means clustering provides an effective approach for classifying navigation states and extracting quantitative behavioral indicators for assessing cognitive abilities. In addition, the environment-based experienced metrics derived from space syntax analysis revealed distinct age-related navigation patterns, highlighting how spatial configuration shapes wayfinding behavior across age groups. These results establish an important foundation for future applications in clinical cognitive assessment and rehabilitation, as well as the design of age-friendly urban environments.

(JMIR Serious Games 2026;14:e83128) doi:[10.2196/83128](https://doi.org/10.2196/83128)

KEYWORDS

spatial navigation; environmental structure; serious games; space syntax; cognitive assessment; urban planning and design; aging

Introduction

With rapid urbanization, the proliferation of densely packed buildings with homogeneous architectural designs has increased the likelihood of disorientation and navigation difficulties. Irregular street networks and the frequent modification of road layouts further disrupt the formation of cognitive maps and hinder spatial orientation [1]. Spatial navigation, the ability to determine an appropriate route and purposefully move toward a specific destination, has increasingly become a considerable challenge, particularly for older adults [2,3].

Significant declines in navigation ability have been recognized as one of the earliest indicators of mild cognitive impairment and Alzheimer disease [4,5]. Individuals experiencing navigation difficulties tend to confine themselves to familiar environments, resulting in critical limitations on their ability to engage in daily activities [5]. Neuroscientific evidence underscores the vulnerability of the entorhinal cortex, a crucial region involved in spatial navigation, to neurodegenerative processes [6]. Grid cells within the entorhinal cortex, which are essential for spatial representation and path integration, exhibit notable functional impairments as individuals age. These age-related deficits in grid cell activity may disrupt the formation of stable cognitive maps, leading to difficulties in orientation, route planning, and wayfinding [7]. Early and rapid identification of navigation impairments is of critical importance.

Evidence from previous studies has indicated that behavioral performance in navigation tasks is influenced by both extrinsic factors (eg, spatial properties of the environment, visual accessibility of the landscape) [8,9] and intrinsic factors (eg, age, cognitive abilities) [10,11]. These factors jointly determine how individuals gather and interpret spatial information, select navigation strategies, and execute wayfinding decisions. Extensive research has demonstrated that environmental cues [12], such as the configuration of landmarks [8], locations [13], and paths [14,15], play a critical role in guiding navigation behavior as well as the neural processes underlying spatial orientation [16,17]. Importantly, age-related differences in the use of environmental cues are well documented [18]. Older adults tend to rely more on associating specific directional responses with particular landmark stimuli [19]; when geometric cues conflict with landmark cues, they are more likely to maintain the landmark-response association [20]. This reliance reflects an egocentric navigation strategy, in which spatial relations are encoded from a body-centered perspective, and actions are guided by subject-to-object associations [18,21,22]. In contrast, younger adults more effectively integrate both geometric and landmark cues, showing greater flexibility in switching between allocentric (map-based) and egocentric (route-based) frames of reference [18].

Although many studies have explored how modifying environmental elements affects navigation behavior, a consistent and intuitive way to characterize spatial properties across different settings is lacking. Space syntax, a well-established framework for quantifying spatial configuration [23], offers a powerful solution. It has been widely applied in urban design [24], path planning [25], and wayfinding research [26]. Within

this framework, axial map analysis (AMA) [13] characterizes spatial structure through line-based representations, whereas visibility graph analysis (VGA) [9,27,28] captures grid-based visibility relationships. Key parameters, including integration, depth, and connectivity, reflect how accessible, shallow, or interconnected different spatial segments are. Prior research has demonstrated the value of these metrics for predicting exploration and wayfinding behavior. Among them, integration is particularly important, as it indicates how accessible a location can be reached from all others and is considered a key determinant of navigation difficulty [29]. Moreover, experienced integration has been shown to indicate the extent to which individuals concentrate their movement within globally integrated regions [13]. Spending more time in highly integrated areas enables individuals to form more accurate cognitive maps [14], resulting in higher experienced integration. Building on this theoretical foundation, we further introduce a series of experienced metrics derived from both axial and visual analyses to quantify the environmental structures that influence participants' actual navigation behavior.

Recent advancements in virtual reality (VR) technology have enabled the creation of highly immersive virtual environments (VEs) that closely replicate the real world with high fidelity [30,31] and have been widely applied in the field of spatial navigation. Researchers can precisely control environmental variables such as landmark distribution, path complexity, spatial layout, and visual cues in immersive VEs [32], making it possible to investigate navigation behavior under the influence of different factors. Moreover, VR technology allows for the collection of a wide range of behavioral data, such as completion time, movement trajectories, error rates, and specific behavioral reactions at given moments [33]. By constructing immersive VEs and a task framework, researchers have developed a variety of experimental paradigms designed for specific navigation tasks, including the Virtual Floor Maze Test [34], Morris Water Maze task [35], and object-location memory task [7,36]. Efficient VR-based navigation tasks are regarded as valuable tools for assessing and enhancing spatial navigation [34].

In recent years, navigation tasks have increasingly been gamified [37,38], with successful applications reported in both the health care domain and geographic science research. For example, the mobile game *Sea Hero Quest* has been widely used for large-scale assessment of human navigation ability [39], providing detailed analyses of navigation performance across age groups and cultural backgrounds [12,40-42]. In parallel, gamification approaches combined with location-based services have demonstrated effectiveness in improving users' understanding of urban maps, orientation, and wayfinding skills in real-world city environments [43,44]. The introduction of serious games (SGs) has enhanced the engagement of navigation tasks while preserving scientific validity. By incorporating clearly defined goals, standardized task structures, and meaningful feedback within immersive environments, SG-based VR navigation paradigms provide effective platforms for assessment, training, and rehabilitation [45].

In this study, we intentionally integrated structured navigation tasks into a goal-oriented game framework. We developed a goal-oriented VR-based navigation game, aimed at effectively

assessing individual cognitive abilities and exploring how environmental configurations influence navigation behavior. Specifically, the objectives were to (1) evaluate the feasibility and reliability of the VR navigation game as a cognitive assessment tool; (2) compare navigation behaviors between younger and older adults; and (3) examine how key environmental factors shape navigation performance across age groups. These findings are expected to provide methodological support for cognitive assessment and training, as well as empirical evidence to inform spatial planning in aging cities.

We hypothesized that navigation behavior would be closely associated with cognitive abilities, such that individuals with higher cognitive capacity would demonstrate more efficient wayfinding performance. We further expected that younger and older adults would exhibit distinct navigation patterns, reflecting age-related differences in cognitive processes and strategy use. Finally, we anticipated that quantitatively measured environmental configuration would exert effects on navigation efficiency, and that these effects would differ significantly between age groups.

Methods

Ethical Considerations

This research involved human participants. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Beihang University (BM20230063). Strict privacy protection measures were applied, and all data were anonymized with no personally identifiable information included. Prior to participation, all individuals were fully informed about the study procedures, potential risks, and their right to withdraw at any time without penalty. All participants provided written informed consent and were compensated at a rate of 100 CNY (approximately US \$14) per hour.

Considering that older adults constitute a potentially vulnerable population, additional safeguards were implemented for the older adult group. These included the presence of at least one qualified health professional to provide immediate medical support, repeated verbal reports to monitor fatigue, dizziness, and overall comfort, and the option for participants to terminate the experiment immediately upon experiencing any discomfort.

Participants

The study was conducted with 21 healthy older adults (14 males and 7 females) aged 65 to 80 years (mean age 71.79, SD 4.57 years) and 18 healthy younger adults (10 males and 8 females) aged 21 to 33 years (mean age 26.50, SD 2.99 years). All participants had no history of neurological or psychiatric disorders. Older participants were recruited from the community and were at least 65 years old and able to complete daily activities independently. One older participant with a history of cerebral hemorrhage was excluded prior to the study, and another withdrew due to intolerance to VR-induced dizziness. Younger participants were recruited through public advertisements posted on the university campus. Individuals who expressed interest were screened using the same inclusion criteria applied to all participants, including normal or corrected-to-normal vision, no history of neurological or

psychiatric disorders, and no prior diagnosis of cognitive impairment. Additionally, individual information, encompassing age, gender, education, and prior VR experience, was systematically gathered from each participant.

System Configuration

The virtual navigation game used in this study was developed with Unity 2019.4 (Unity Technologies) and deployed via the HTC VIVE Pro head-mounted display. The visual scene was rendered at 90 Hz and presented at a single-eye resolution of 1440×1600 pixels (2880×1600 binocular) with a 110° field of view, providing an immersive experience. For the experimental setup, the head-mounted display was connected to a high-performance desktop PC via a wired connection using the standard VIVE link box to ensure low latency and stable rendering. The system was equipped with an Intel Core i7-9700K CPU, an NVIDIA GeForce RTX 2070 GPU, 32 GB of RAM, and the Windows 10 operating system, enabling smooth rendering and reliable behavioral data acquisition.

The VE was designed as a simplified yet ecologically valid urban block (200 vm×150 vm, virtual meters), featuring residential buildings, street networks, greenery, and small parks. The intentional integration of these elements provided sufficient spatial complexity to elicit meaningful wayfinding behavior while minimizing potential discomfort or anxiety associated with navigation in a VE [46].

Given that older participants may experience limitations in sustained physical walking, we adopted a hybrid navigation control scheme combining the handheld controller and natural facing direction [47]. The VIVE controller features an ergonomic design with a circular touchpad. To minimize operational complexity, participants were instructed to press the touchpad to move forward in the direction they were facing, and releasing it stopped movement immediately. Horizontal body rotation was mapped one-to-one to virtual orientation, allowing participants to turn naturally without additional button input. Throughout the experiment, all participants were instructed to avoid movement in the physical space and to navigate using the controller and body rotation only. Most participants completed the game while standing, whereas three older participants, due to advanced age, performed the experiment while seated in a swivel chair that allowed 360° physical rotation. This simplified control scheme reduced cognitive and motor demands while ensuring consistent interaction across participants.

Experimental Procedure

The experiment consisted of three phases: (1) neuropsychological assessment, (2) pretraining, and (3) formal experiment. The neuropsychological assessment phase aimed to evaluate the cognitive abilities of older participants using standardized measures of general cognitive skills and empirically validated tests. During the pretraining phase, participants were instructed to navigate freely within the VE for 15 minutes. Navigation proficiency was considered achieved when participants demonstrated the ability to independently and fluently operate the controller and navigate to designated target locations, qualifying them to proceed to the formal experimental

phase. In the formal experimental phase, participants were required to sequentially navigate to multiple target locations, with spatiotemporal data recorded at fixed intervals throughout the navigation process. Older participants completed all 3 phases, while younger participants, who were healthy university students with presumed normal cognitive function, only participated in phases 2 and 3.

Before the formal navigation tasks, general cognitive functioning was quantified using the Mini-Mental State Examination (MMSE) [48,49] and the Montreal Cognitive Assessment (MoCA) [50]. Specifically, the MMSE administered in this study was the standardized Chinese adaptation based on the international MMSE framework, which has been validated through prior pilot testing and is widely used in studies involving Chinese populations [48,49,51]. The MoCA Full Chinese (Beijing) Version 7.1 was the official Beijing version published on the MoCA website [52]. MoCA scores were adjusted according to participants' education level, with one additional point added for individuals with 12 years of formal education or fewer. Nonverbal memory was further assessed using the delayed recall condition of the Rey-Osterrieth Complex Figure Test, while visuospatial functioning was evaluated through the Copy condition of the Rey-Osterrieth Complex Figure (ROCF-C) test [53]. To examine executive function and working memory, Part A of the Trail Making Test (TMTA) was

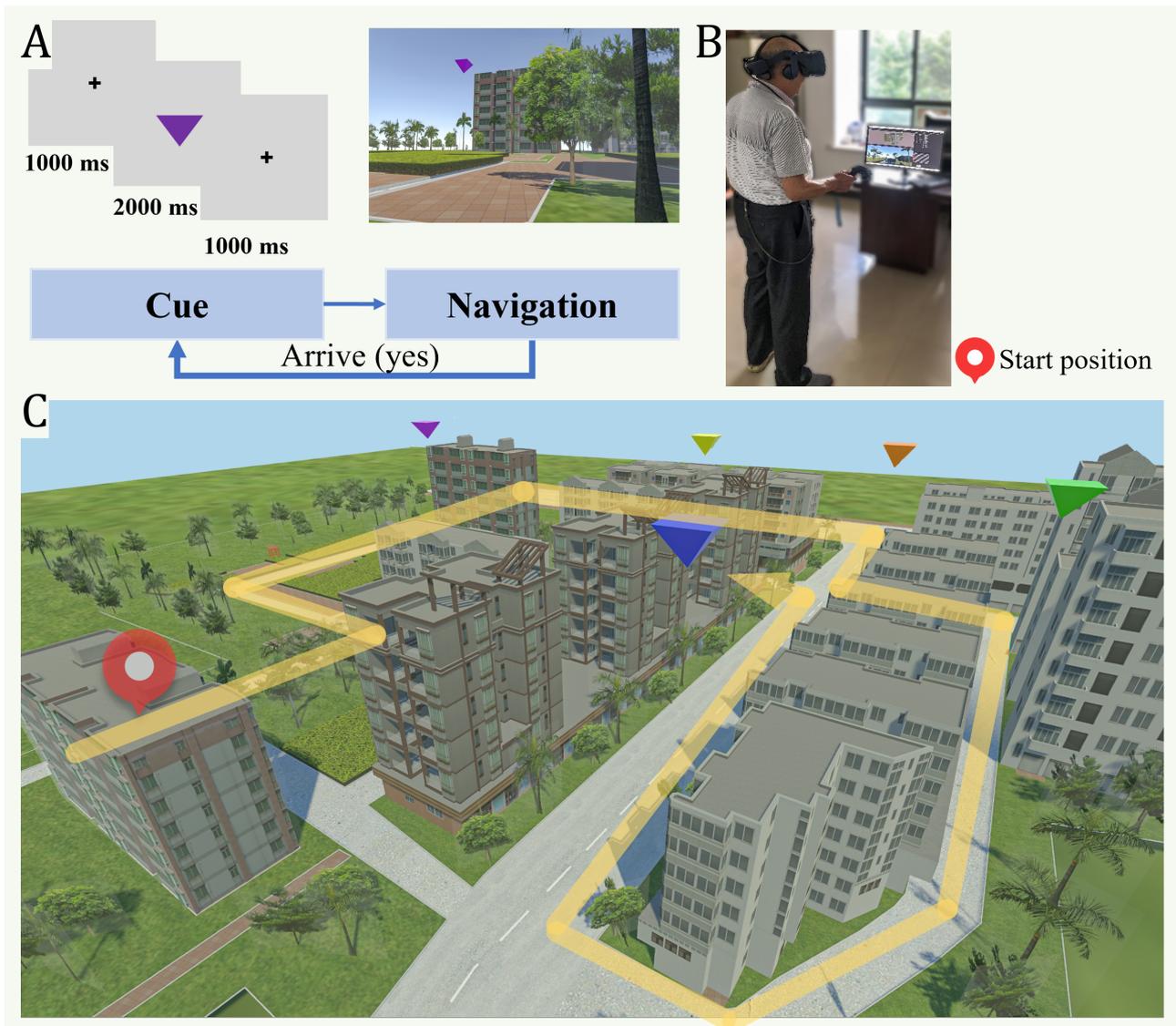
administered [54]. Additionally, spatial rotation ability was assessed using a standardized Mental Rotation Test (MRT) [55]. These assessments provided a comprehensive cognitive profile for each participant prior to engagement in the virtual navigation experiment.

During the pre-experimental phase, identically shaped but differently colored target objects were placed at various locations within the VE (Figure 1). Participants were instructed to memorize the spatial locations corresponding to each colored object. The specific verbal instructions can be found in Multimedia Appendix 1. During the formal experimental phase, 5 target locations previously encountered during the training phase were used (Figure 2C). Target locations were presented in a fixed sequential order, allowing for consistent route comparisons across participants. Each trial began with a fixation cross (1000 ms), followed by the target cue (2000 ms), another fixation cross (1000 ms), and then participants were required to navigate to the designated target location (Figure 2A). Upon reaching the target, the current trial was concluded, and the next trial was automatically initiated. This sequence continued until all tasks were completed. Participants were instructed to navigate as quickly and accurately as possible (Figure 2B), and a congratulatory animation was displayed after each trial to prevent boredom.

Figure 1. Pretraining phase: top-down view of the scene (middle) and views of the target from four different locations (left and right).



Figure 2. Formal experiment: (A) The sequence of a single trial. Following the presentation of a cue, participants were required to navigate to the target location. (B) Photograph of an older participant during the experiment. (C) Layout of the virtual city environment, including the shortest route planned to all target locations.



Environmental Structure Analysis

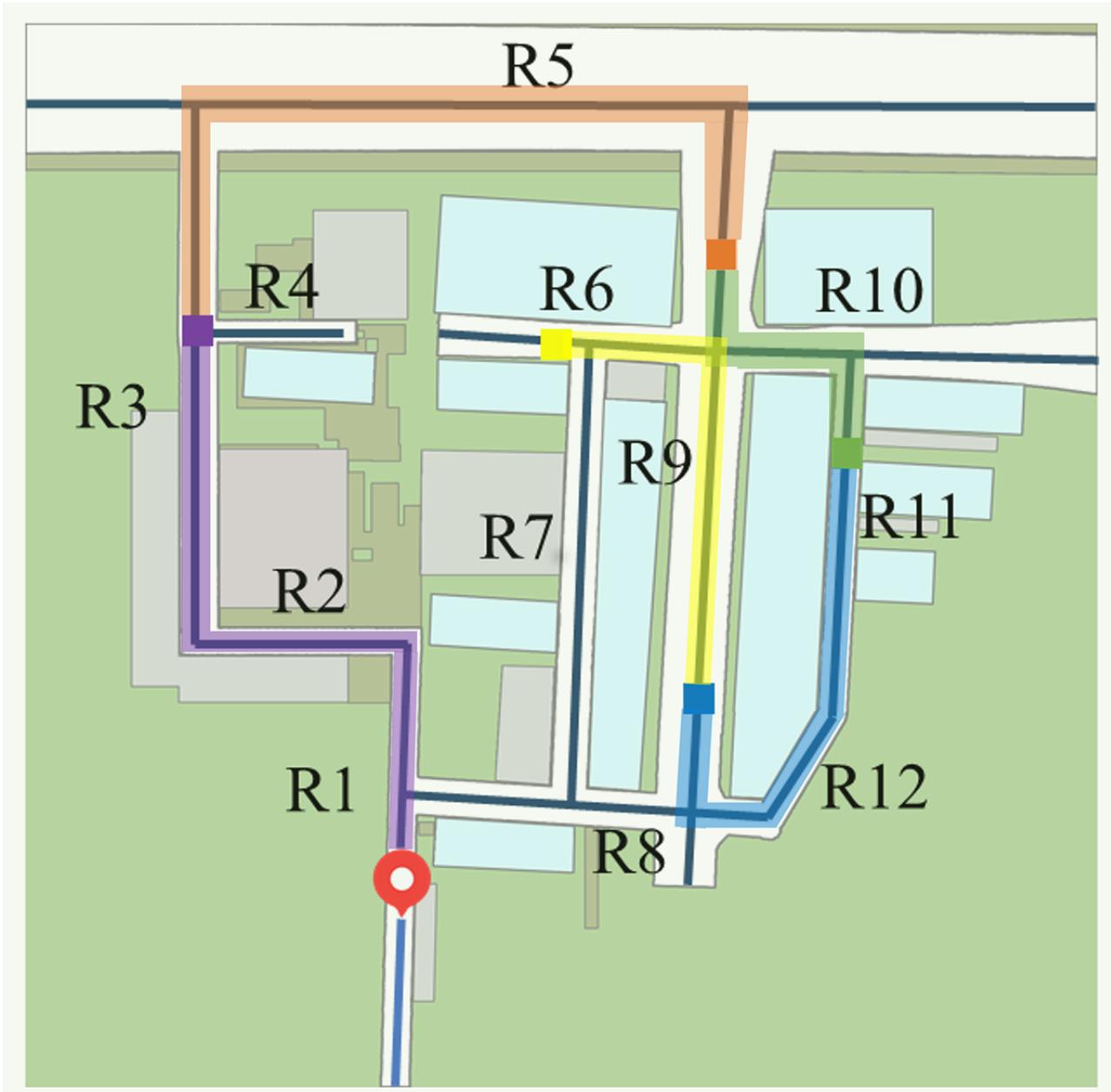
Space Syntax

Space Syntax provides a robust and systematic framework for examining the spatial arrangement of environments and the interrelationships between spaces. This methodology has been extensively applied to explore how spatial configurations affect human behavior [26,56,57], particularly how individuals encode, represent, and retrieve spatial information during navigation. In this study, both AMA and the VGA of the VE were carried out using the DepthmapX software (Space Syntax Laboratory).

Axial Map Analysis

The axial map provides a graphical representation of spatial structure, where streets within an urban environment are abstracted as axial lines, defined as the minimal set of longest straight lines of sight from a given location [23,58]. These lines collectively cover the entire space, capturing all possible movement paths. Based on the 2D layout of the VE street network, we constructed an axial map with a total of 12 axial lines delineated (Figure 3). For analytical purposes, key metrics such as axial integration, axial connectivity, and axial mean depth were calculated to characterize the accessibility and flow potential of individual environmental segments.

Figure 3. The axial map constructed from the 2D layout of the virtual environment street network.



Visibility Graph Analysis

VGA provides an analytical approach for quantifying the spatial characteristics of an environment by assessing the extent to which any given point within the space is visible from any other point [59]. In this study, the structural layout of the VE was first modeled using AutoCAD. Subsequently, a 1 m×1 m grid was generated in DepthmapX to perform the analysis. This process enables the computation of key spatial metrics at each node, including visual integration, visual connectivity, and visual mean depth.

Measurement

Metrics Setting

A set of behavioral and environment-based metrics was calculated in the virtual navigation game to evaluate the effectiveness of the game in assessing cognitive abilities and

to examine age-related differences in navigation performance and sensitivity to environmental features. Spatiotemporal data were recorded for each participant at a fixed sampling frequency. Based on these data, the navigation behavioral metrics and environment-based navigation metrics were defined as follows.

Wayfinding, Transition, and Moving States

During navigation, we continuously recorded each participant's spatial position (x_i, y_i) at timestamp t_i . For each recorded point i (excluding the starting and ending points), the instantaneous speed was computed as the average of the velocity estimates derived from two adjacent time intervals.

$$v_i = \frac{\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \cdot t_i - t_{i-1} + \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \cdot t_{i+1} - t_i}{2}$$

The instantaneous speed series for each participant was then classified into 3 distinct categories using the k-means clustering algorithm, yielding low, medium, and high speed clusters. These clusters corresponded to meaningful navigation states:

Wayfinding (low speed: searching, hesitation, or reorientation), Transition (medium speed: directional adjustment or preparatory acceleration), and Moving (high speed: continuous, goal-directed locomotion). This classification allowed us to segment the entire trajectory into functionally interpretable behavioral states.

Navigation Efficiency

Navigation efficiency is defined as the ratio of effective navigation time (moving time and transition time) to the total duration of the navigation game (navigation time).

$$\text{Navigation efficiency} = \frac{\text{Moving time} + \text{Transition time}}{\text{Navigation time}}$$

This metric serves as an indicator of the capability to navigate effectively within the environment. Higher values reflect a more streamlined and goal-oriented navigation process, characterized by reduced time spent on behaviors such as hesitation, reorientation, or uncertainty. By quantifying navigation efficiency, we can evaluate the extent to which individuals construct accurate cognitive maps and adopt effective spatial strategies during navigation tasks.

Experienced Axial Metrics

Axial analysis is derived from a representation of the environment using the fewest and longest lines of sight, which capture key structural characteristics of an urban street network. In this study, we used the axial values computed from the axial map shown in Figure 3. Three axial measures were considered: axial integration, which quantifies the accessibility of an axial line from the rest of the system; axial connectivity, which denotes the number of axial lines directly linked to a given line; and axial mean depth, which represents the arithmetic mean of the topological depths from each axial line to all others. To examine how participants' navigation behavior was shaped by the underlying street configuration, we computed the average

value of each metric along the participant's trajectory, resulting in 3 experienced axial metrics: experienced axial integration (EAI), experienced axial connectivity (EAC), and experienced mean axial depth (EMAD).

$$\text{EAI} = \frac{1}{n} \sum_{i=1}^n \text{IA}(x_i, y_i)$$

$$\text{EAC} = \frac{1}{n} \sum_{i=1}^n \text{CA}(x_i, y_i)$$

$$\text{EMAD} = \frac{1}{n} \sum_{i=1}^n \text{DA}(x_i, y_i)$$

The point (x_i, y_i) denotes the coordinates of the participant at sample i , and IA, CA, and DA represent the integration, connectivity, and mean depth values of the axial line associated with that position. The variable n indicates the total number of sampled points along the participant's trajectory.

Experienced Visual Metrics

VGA represents an environment as a grid-based visibility network, enabling the quantification of visual relationships between locations. The visual integration, visual connectivity, and visual mean depth values used in this study were derived from the VGA map shown in Figure 4. To examine how participants' navigation behavior was influenced by visually accessible spatial properties, we computed the average value of each VGA metric along the participant's trajectory, yielding experienced visual integration (EVI), experienced visual connectivity, and experienced visual mean depth.

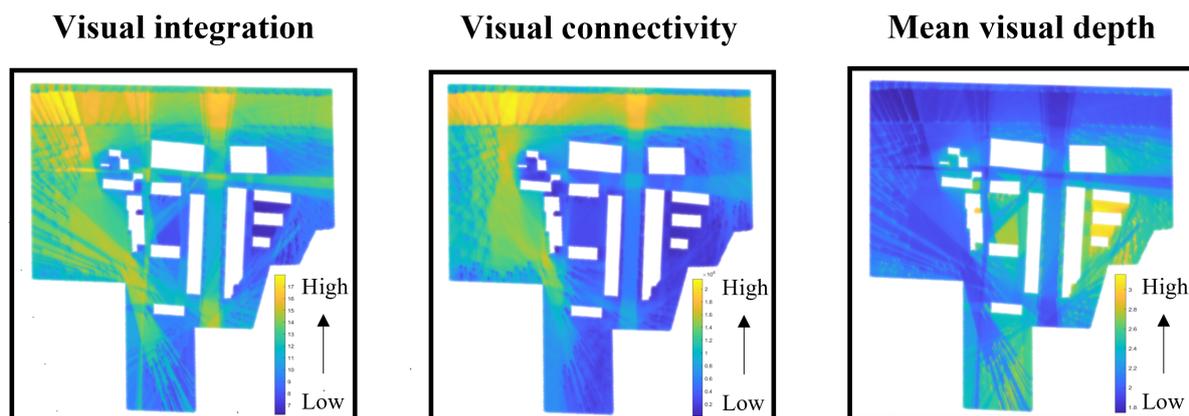
$$\text{EVI} = \frac{1}{n} \sum_{i=1}^n \text{Iv}(x_i, y_i)$$

$$\text{EVC} = \frac{1}{n} \sum_{i=1}^n \text{Cv}(x_i, y_i)$$

$$\text{EVMD} = \frac{1}{n} \sum_{i=1}^n \text{Dv}(x_i, y_i)$$

Iv, Cv, and Dv represent visual integration, visual connectivity, and visual mean depth values associated with participants' position (x_i, y_i) in the VGA grid.

Figure 4. The visualization of the virtual urban spatial structure using visibility graph analysis.



Statistical Analysis

Data from the experiment were processed using custom scripts developed in MATLAB, and all statistical analyses were performed using SPSS (IBM Corp). The normality of each variable was assessed through visual inspection of histograms and the Shapiro–Wilk test. Spearman's rank correlation analyses were conducted to examine the associations between navigation-related behavioral measures and cognitive test scores. To evaluate age-related differences, independent-samples t tests

were applied to variables that met normality assumptions, whereas the Mann-Whitney U test was used for variables that were not normally distributed. In addition, correlations between navigation efficiency and the various experienced spatial metrics were examined using Pearson correlations for normally distributed variables and Spearman rank correlations for nonnormally distributed variables.

Results

Participants

All 19 older participants and 18 young participants successfully completed the experiment without experiencing severe motion sickness. Table 1 summarizes the cognitive characteristics of the older group, including the means and SDs for various cognitive assessment scales. Notably, none of the participants

had prior experience with VR technology, which reduced potential biases associated with equipment familiarity.

After clustering the data, points labeled as wayfinding state exhibited clear spatial concentration and age-related differences (Figure 5). This spatial density map visualizes wayfinding patterns, offering an intuitive overview of participants' spatial behavior based on spatiotemporal features. The red-outlined areas indicate the top 1% of regions with the highest density values, highlighting zones of intensified wayfinding activity within the VE.

Table 1. Characteristics of participants of different groups.

	Older adults	Young adults
Age (years), mean (SD)	71.79 (4.57)	26.50 (2.99)
Sex (male/female), n	14/5	10/8
Education >12 years (yes/no), n	11/8	18/0
VR ^a experience (yes/no), n	0/19	0/18
MoCA ^b , mean (SD)	21.84 (5.47)	— ^c
MMSE ^d , mean (SD)	27.16 (3.60)	—
TMTA ^e , mean (SD)	62.06 (21.27)	—
ROCF-D ^f /ROCF-C ^g , mean (SD)	31.63 (4.59)/18.60 (5.62)	—
MRT ^h , mean (SD)	68.58 (13.62)	—

^aVR: virtual reality.

^bMoCA: Montreal Cognitive Assessment.

^cNot available.

^dMMSE: Mini-Mental State Examination.

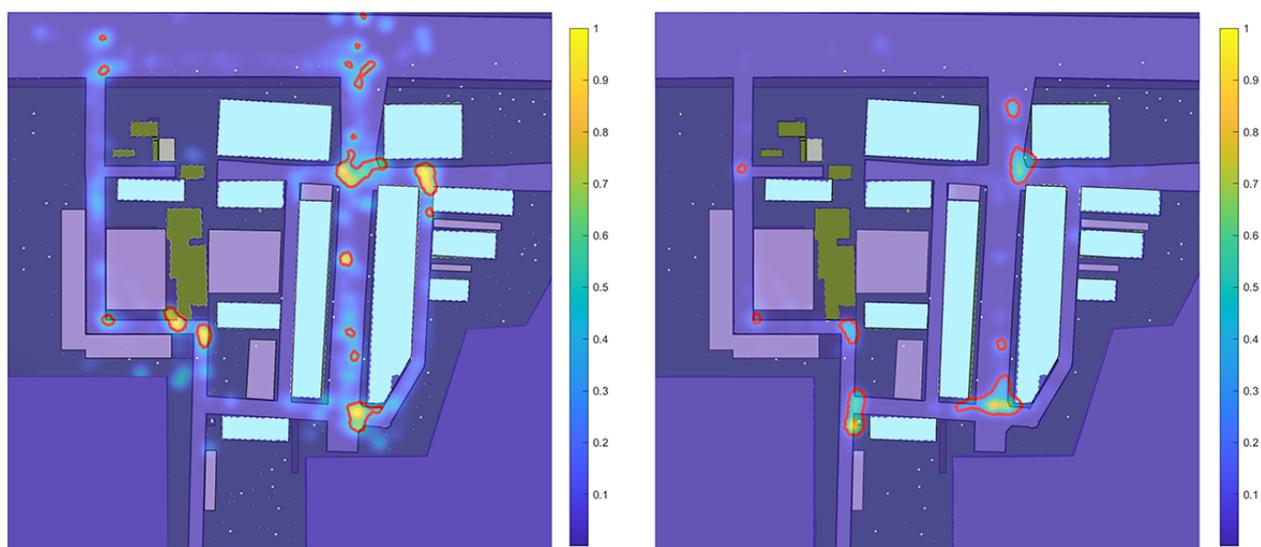
^eTMTA: Part A of the Trail Making Test.

^fROCF-D: delayed recall condition of the Rey-Osterrieth Complex Figure.

^gROCF-C: Copy condition of the Rey-Osterrieth Complex Figure.

^hMRT: Mental Rotation Test.

Figure 5. Spatial density distribution of wayfinding points (left: older adults; right: young adults).



The Correlation Between Navigation Behavior and Cognitive Abilities in Older Participants

Spearman rank correlation analysis was conducted to examine the relationships between navigation behavior metrics and cognitive assessment scores (Table 2). No significant correlations were observed between either distance or error and any of the cognitive scales. However, temporal-based navigation parameters performed well. Navigation time was significantly associated with both TMTA ($r=0.503$, $P=.047$) and the MRT ($r=-0.526$, $P=.03$). Wayfinding time showed strong correlations

with several cognitive measures, including TMTA ($r=0.716$, $P=.002$), MoCA ($r=-0.549$, $P=.02$), ROCF-C ($r=-0.588$, $P=.01$), MMSE ($r=-0.458$, $P=.046$) and MRT ($r=-0.751$, $P<.001$). Additionally, navigation efficiency demonstrated significant associations with MoCA ($r=0.495$, $P=.04$) and ROCF-C ($r=0.658$, $P=.003$) scores, and particularly strong correlations with TMTA ($r=-0.761$, $P=.001$) and MRT ($r=0.848$, $P<.001$). These findings confirmed that cognitive abilities play a crucial role in influencing navigation efficiency, including navigation time and wayfinding time.

Table . The correlation between spatial navigation and cognitive scales.

	MoCA ^a	MMSE ^b	ROCF-C ^c	ROCF-D ^d	TMTA ^e	MRT ^f	Distance	Error	Navigation time	Wayfinding time
MMSE										
<i>r</i>	0.885 ^g	1								
<i>P</i> value	<.001	— ^h								
ROCF-C										
<i>r</i>	0.470 ⁱ	0.364	1							
<i>P</i> value	.049	.14	—							
ROCF-D										
<i>r</i>	0.344	0.084	0.603 ^j	1						
<i>P</i> value	.16	.74	.008	—						
TMTA										
<i>r</i>	-0.663 ^j	-0.496	-0.380	-0.081	1					
<i>P</i> value	.005	.05	.15	.76	—					
MRT										
<i>r</i>	0.409	0.301	0.701 ^j	0.349	-0.591 ⁱ	1				
<i>P</i> value	.09	.22	.001	.16	.02	—				
Distance										
<i>r</i>	-0.287	-0.201	-0.013	-0.122	-0.105	0.103	1			
<i>P</i> value	.25	.42	.96	.63	.70	.68	—			
Error										
<i>r</i>	-0.278	-0.098	0.049	-0.203	-0.07	0.313	0.873 ^g	1		
<i>P</i> value	.26	.70	.85	.42	.80	.21	<.001	—		
Navigation time										
<i>r</i>	-0.399	-0.286	-0.356	-0.247	0.503 ⁱ	-0.526 ⁱ	0.655 ^j	0.473 ⁱ	1	
<i>P</i> value	.10	.25	.15	.32	.047	.02	.003	.047	—	
Wayfinding time										
<i>r</i>	-0.549 ⁱ	-0.458 ⁱ	-0.588 ⁱ	-0.296	0.716 ^j	-0.751 ^g	0.426	0.216	0.920 ^g	1
<i>P</i> value	.02	.06	.01	.23	.002	<.001	.08	.39	<.001	—
Navigation efficiency										
<i>r</i>	0.495 ⁱ	0.445	0.658 ^j	0.206	-0.761 ^j	0.848 ^g	-0.067	0.129	-0.709 ^j	-0.913 ^g
<i>P</i> value	.04	.06	.003	.41	.001	<.001	.79	.61	.001	<.001

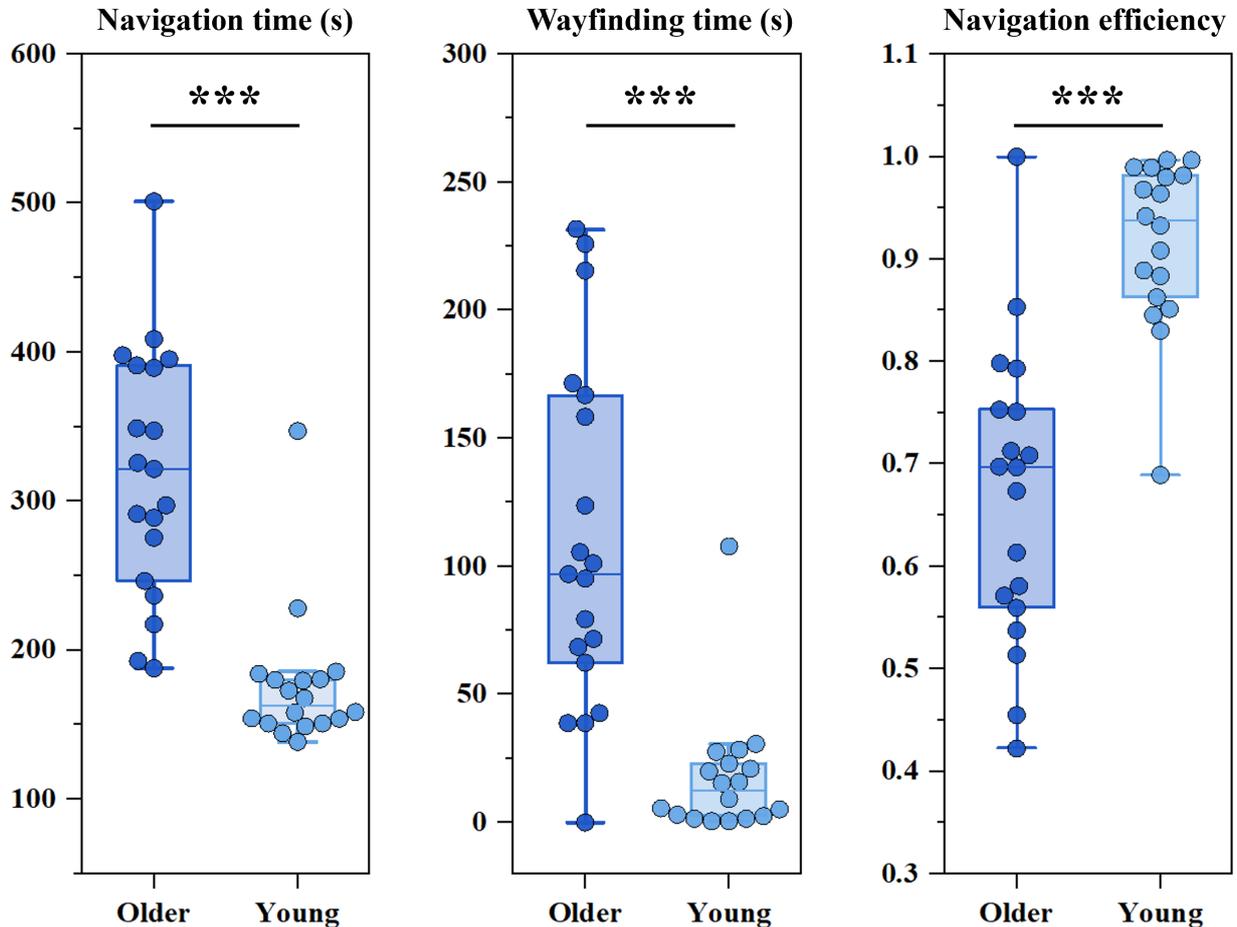
^aMoCA: Montreal Cognitive Assessment.^bMMSE: Mini-Mental State Examination.^cROCF-C: Copy condition of the Rey-Osterrieth Complex Figure.^dROCF-D: delayed recall condition of the Rey-Osterrieth Complex Figure.^eTMTA: Part A of the Trail Making Test.^fMRT: Mental Rotation Test.^g*P*<.001^hNot applicable.ⁱ*P*<.05^j*P*<.01.

Analysis of Age-Related Decline in Wayfinding Ability

The Mann-Whitney U test was used to assess differences in navigation performance between age groups (Figure 6). Compared to young participants, older adults, who exhibited

significantly lower cognitive ability, demonstrated significantly longer navigation time ($z=-4.771$, $P<.001$) and wayfinding times ($z=-4.315$, $P<.001$), as well as reduced navigation efficiency ($z=-4.285$, $P<.001$).

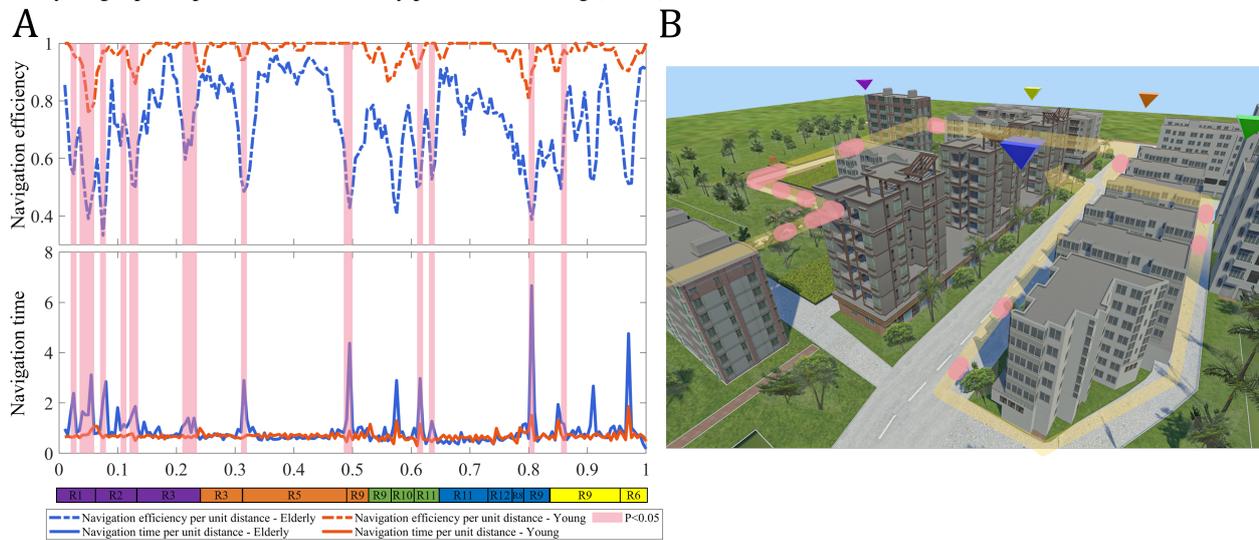
Figure 6. Box plots illustrating differences in navigation performance between older and younger participants.



The above metrics reflect participants' overall wayfinding performance. However, navigation behavior is not solely influenced by internal factors such as cognitive ability. It also varies depending on the location of participants within the VE. Figure 7A illustrates differences in the navigation process across age groups during the formal experiment. To investigate this, participants' spatial positions are mapped onto the shortest navigation routes (excluding trials involving route errors), and the time spent per unit distance is calculated for each individual. Based on these data, we plot curves of navigation time and

navigation efficiency for each age group. Furthermore, a permutation test is conducted on navigation efficiency within each unit distance between older and young adults. The shaded regions on Figure 7B correspond to areas where significant differences are observed ($P<.05$) in Figure 7A, revealing spatial zones where age-related differences in navigation behavior are most pronounced. These results highlight age-related differences in spatial behavior and reinforce the utility of these behavioral metrics as sensitive indicators of cognitive function.

Figure 7. Comparison of the navigation process between older and younger participants during the formal experiment. (A) The x-axis represents the normalized value along the shortest distance across 5 consecutive trials, and the y-axis shows both navigation time and navigation efficiency. The colored bar at the bottom corresponds to the target color for each trial, and the labels R1, R2, ... indicate the axial lines that constitute the shortest navigation route for that trial. (B) Pink-shaded areas indicate route segments where significant differences in navigation efficiency were found between older and younger participants, as determined by permutation testing ($P < .05$).



Effects of Environmental Structure on Navigation Behavior in Different Age Groups

The influence of environmental characteristics on navigation performance is analyzed from the perspective of spatial configuration using both AMA and VGA. Between-group comparisons revealed significant age-related differences in several experienced spatial measures. Older adults showed higher EAI ($z = -2.43$, $P = .01$) and EVI ($t = 2.48$, $P = .02$), whereas young adults exhibited higher EMAD ($z = -2.13$, $P = .03$) and EMVD ($t = -2.49$, $P = .02$). No significant group differences were found for experienced axial connectivity and experienced visual connectivity.

Correlation analyses further revealed distinct age-specific associations between navigation efficiency and experienced metrics. Among older adults, navigation efficiency showed a significant negative correlation with EVI ($r = -0.48$, $P = .04$) and a positive correlation with EMVD ($r = 0.49$, $P = .03$). In contrast, young adults demonstrated strong associations with axial properties, with navigation efficiency negatively correlated with EAI ($r = -0.64$, $P = .005$) and positively correlated with EMAD ($r = .72$, $P < 0.001$). Young adults also exhibited a significant positive relationship between navigation efficiency and EMVD ($r = 0.57$, $P = .01$).

Discussion

Summary and Explanation of Findings

This study investigated how cognitive abilities and environmental structure jointly shape navigational performance in a virtual navigation game across different age groups. Consistent with prior research, our findings demonstrate that performance in navigation games, particularly navigation efficiency, is strongly influenced by individual cognitive capacity, as well as the spatial configuration of the environment. These results underscore the relevance of navigation behavior

as an indicator of cognitive status and highlight the potential of VR-based navigation tasks as effective tools for cognitive assessment. Moreover, the insights gained from this work also provide valuable guidance for future urban design and navigation-game development.

The behavioral metrics used in this study, including traversal distance, number of errors, and navigation time, are widely used in human navigation research [34,60]. Spatiotemporal representation has long been essential for analyzing navigation behavior, and the present study extends conventional approaches by classifying participants' navigation states using an individualized method. Navigation tasks are generally considered to consist of 2 main components [61]: locomotion and wayfinding. Locomotion refers to the physical execution of movement in the space, whereas wayfinding is involved in determining and planning a route to a desired destination [62]. Traditional approaches typically distinguish these components using fixed absolute speed thresholds [47,63]. To account for individual differences in navigation ability, we applied k-means clustering to each participant's instantaneous speed distribution to derive personalized relative speed thresholds. k-Means is a classical clustering algorithm that iteratively assigns data points to the nearest cluster center [64]. To assess the reliability of this classification, clustering solutions with different numbers of clusters (k) were evaluated across all younger and older participants using the sum of squared errors (SSE) [65] and the silhouette coefficient [66] (Figure S3 in Multimedia Appendix 2). The SSE quantifies within-cluster variance by measuring the squared distance between each data point and its assigned cluster center, whereas the silhouette coefficient provides a complementary assessment of clustering quality by comparing consistency within clusters with separation between other clusters [67]. When $k = 3$, the SSE curve showed a clear inflection point, indicating a markedly diminished rate of decrease in within-cluster variance. Meanwhile, the silhouette coefficient remained relatively high and exhibited more stable variance, suggesting an optimal balance between cluster cohesion and

separation. In addition to these statistical indicators, the three clusters aligned well with meaningful navigation states observed in participants' movement patterns (Figure S2 in [Multimedia Appendix 3](#)): (1) low-speed hesitation or searching, (2) medium-speed transitional adjustments, and (3) high-speed goal-directed movement. Taken together, both the quantitative metrics and the behavioral interpretability support the individualized clustering approach, which enabled the effective categorization of each sampled time point into low-, medium-, or high-speed states, thereby identifying regions within the environment where participants associated with wayfinding, transition, or continuous movement.

Effective spatial navigation relies on the integration of multiple abilities, including visual perception, spatial orientation, learning, and memory [68]. Wayfinding encompasses all of the ways in which people orient themselves in physical space and navigate from place to place. As anticipated, wayfinding ability, as a cognitive element of the navigation process [69], is effectively captured by parameters such as wayfinding time and navigation efficiency, both of which are shown to be sensitive indicators of cognitive capacity in the current study. As summarized in [Table 2](#), cognitive scales including the MMSE, MoCA, ROCF-C, TMTA, and MRT exhibited robust correlations with these navigation metrics. Among these, the MoCA and MMSE are widely used for comprehensive cognitive assessment; the TMTA primarily reflects executive functioning; ROCF-C performance is sensitive to visuoconstructive and geometric processing abilities; and the MRT assesses mental rotation capacity. These cognitive functions collectively support navigation by contributing to processes such as initial route planning, continuous movement control, reorientation, and decision-making. SG-based VR navigation tasks elicit continuous, ecologically relevant behavior under dynamic task constraints, enabling the simultaneous assessment of multiple cognitive functions within an integrated behavioral context. Quantifying wayfinding performance provides an indirect yet meaningful reflection of an individual's cognitive functioning [34]. Longer wayfinding times and lower navigation efficiency were associated with poorer cognitive performance in our older participants. Taken together, these findings highlight the potential of well-designed navigation games to serve as effective tools for cognitive assessment [70].

Human navigation in outdoor environments involves a series of distinctive behaviors, such as retracing steps, hesitation, and reorientation [71]. These behaviors become particularly evident in complex spatial settings where individuals must process environmental cues and make directional decisions. Wayfinding time serves as a key indicator for revealing navigation uncertainty, effectively pinpointing locations where external environmental factors impact individuals' sense of directional uncertainty ([Figure 4](#)). We observe that prolonged wayfinding times were predominantly concentrated in the turning areas of the path and the starting or ending locations of each trial. These highlighted areas in the figure represent regions where participants are more likely to encounter decision-making bottlenecks, experiencing uncertainty in selecting the correct path [72]. Furthermore, integrating the analysis in [Figure 6](#), we

find significant differences in navigation performance between older and younger participants in certain regions.

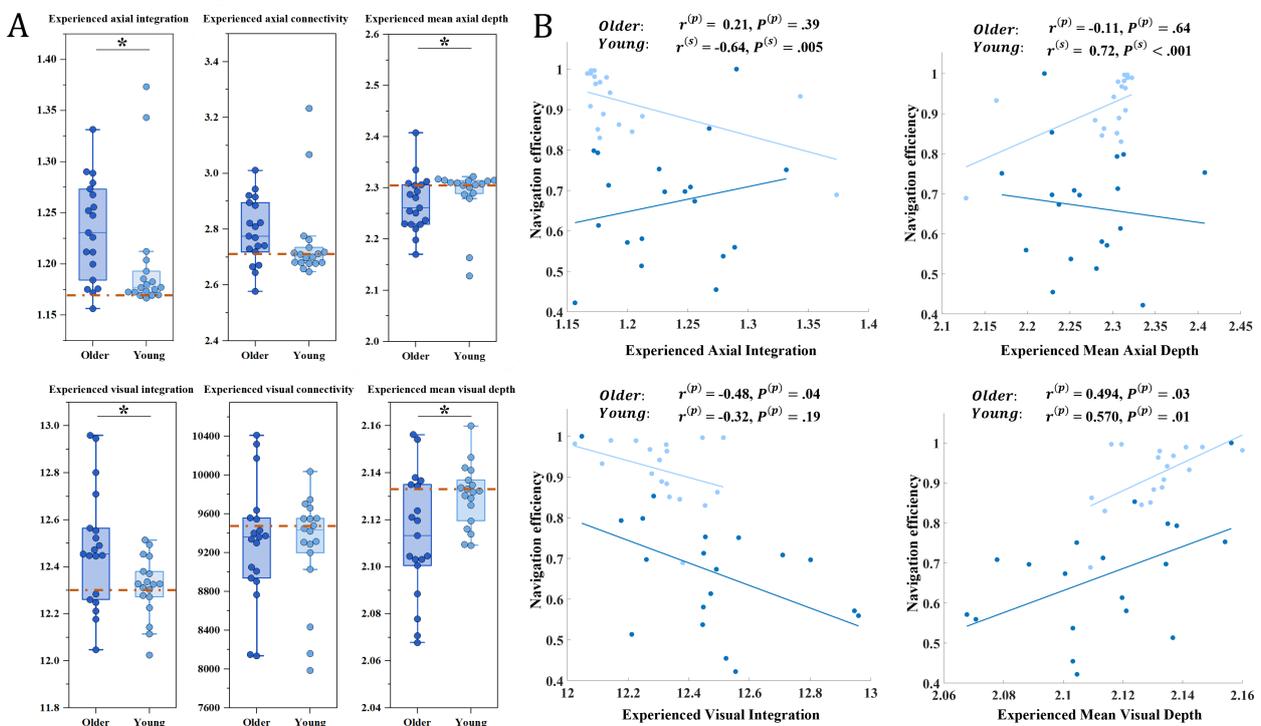
Synthesizing findings from previous research, we propose that these results may be explained by age-related differences in navigation strategy preferences. Human navigation relies on 2 primary strategies: allocentric (survey-based) and egocentric (route-based). Younger adults tend to rely more on the allocentric strategy [18] and show greater flexibility in switching between allocentric and egocentric frames of reference [73]. However, aging is associated with a decline in allocentric navigation abilities, accompanied by a preferential shift toward egocentric strategies [18,21,22], which may contribute to less efficient route planning and increased reliance on familiar or visually accessible paths. These strategy differences align closely with the behavioral patterns observed in our study. Young participants adopted a more proactive and goal-directed approach: they typically scanned the environment at the beginning of each trial, extracted relevant spatial cues, and formed a planned route before moving. In contrast, older participants showed a stimulus-response navigation pattern [74,75], making decisions only upon arriving at intersections rather than planning ahead. This led to hesitation at nearly every decision point ([Figure 6B](#)), reflecting greater reliance on immediate cues and reduced anticipation of upcoming turns. Consequently, participants spent more time searching for landmarks or other environmental cues, leading to decreased navigation performance [76,77]. Such behaviors are consistent with age-related declines in spatial working memory and executive function [78], which limit their ability to efficiently integrate spatial cues.

To explore how environmental structure influences navigation across age groups, we used line-based (AMA) and grid-based (VGA) analytic approaches. Both methods quantify spatial configuration and offer complementary perspectives on how spatial structure shapes navigation behavior. Within this framework, integration serves as a core space syntax metric capturing the degree to which a path is connected to its surrounding environment [13,29]. In this study, experienced integration reflects the actual spatial integration values of the segments that participants traversed, thereby characterizing the dynamic interplay between environmental structure and navigational behavior. Older adults exhibited significantly higher EAI and EVI than young adults, suggesting a tendency to navigate toward more integrated areas rather than taking shorter or more direct routes. Prior studies have shown that spending more time in highly integrated areas facilitates the formation of accurate cognitive maps [13]. With aging, older adults require more time to form cognitive maps and encode spatial information [79], which may partly explain their tendency to remain within highly integrated axial regions during navigation. From another perspective, although integration is the normalization form of mean depth [25], the two measures capture different aspects of spatial structure. Mean depth quantifies the average topological steps from a location to all others, with higher values indicating deeper and harder-to-reach positions. Older adults showed lower EMAD and EMVD values, suggesting a tendency to remain within relatively shallower, more accessible areas.

In our navigation game, all participants were instructed to reach the target location as quickly as possible. By comparing each group's experienced metrics with those of the shortest path (Figure 8A), we found that the experienced metrics of younger participants closely matched the optimal path. This pattern is consistent with behavioral observations showing that young participants, driven by the goal of rapid completion, tended to select more direct and efficient routes. Correlation analyses further support these environment-mediated navigation patterns. Among young adults, EAI was negatively correlated with navigation efficiency, indicating that more movement through globally integrated axial lines was associated with increased wayfinding behaviors (eg, hesitation, pausing, reorientation),

thereby reducing overall efficiency. For older adults, EVI showed a significant negative correlation with navigation efficiency, suggesting that more extensive traversal within visually integrated regions corresponded to heightened hesitation or searching behaviors. The results for EMAD and EMVD exhibited similar trends: achieving higher navigation efficiency and shorter routes requires participants to strategically decide whether to traverse spatially advantageous or disadvantageous areas. Younger adults tended to accept navigating through deeper or less accessible regions to optimize route efficiency [56], whereas older adults were more likely to remain within visually accessible areas, even when these were not part of the shortest path.

Figure 8. Experienced metrics across age groups and their associations with navigation efficiency. (A) Group differences in 6 experienced spatial metrics, experienced axial integration (EAI), experienced axial connectivity, experienced mean axial depth (EMAD), experienced visual integration (EVI), experienced visual connectivity, and experienced visual mean depth (EVMD) are shown for older adults (dark blue) and young adults (light blue). The brown dashed line represents the experienced metric obtained from the shortest path. (B) Correlations between navigation efficiency and 4 key experienced metrics (EAI, EMAD, EVI, and EVMD) are presented separately for older adults (dark blue) and young adults (light blue). Pearson correlations are denoted by (p) and Spearman correlations by (s).



This subjective tendency in navigational behavior indicates that participants do not passively respond only to environmental structure; rather, their navigational choices also reflect intentional, strategy-driven preferences [13]. Such findings point to a bidirectional relationship between spatial configuration and navigation behavior, where environmental structure shapes movement patterns, and individuals' strategic preferences, in turn, influence the parts of the environment they traverse. Understanding whether environmental properties or individual behavior play the dominant role in guiding navigation will require further empirical investigation.

Additionally, our findings extend this literature by demonstrating age-specific patterns in the relationship between experienced metrics and navigation efficiency. Navigation efficiency in younger adults was strongly associated with line-based (AMA) metrics such as EAI and EMAD, whereas older adults showed

stronger associations with grid-based (VGA) metrics, such as EVI and EMVD. These differences can be meaningfully interpreted through the lens of allocentric and egocentric navigation strategies. Younger adults tend to rely more on allocentric representations, enabling them to rapidly extract structural information from the environment, integrate axial connectivity, and flexibly switch between global structure and local visual information. This cognitive flexibility allows them to traverse deeper or less accessible areas. In contrast, aging is associated with an increased reliance on egocentric navigation. Older adults become more dependent on the visual field and immediate cues, which aligns with their stronger correlation between navigation efficiency and visual-based experienced metrics (EVI and EMVD). Their tendency to remain within visually accessible, highly integrated regions suggests that they

prioritize environments that offer clear visual guidance and reduced cognitive demands [14].

Together, space syntax highlights the fundamental role of spatial configuration in shaping navigation behavior, and our findings further show that its influence varies with age, ultimately affecting navigation performance. These insights have important implications not only for developing more effective cognitive assessment tools but also for designing age-friendly urban environments that support safe, efficient, and accessible navigation. Strengthening the coordination and integration of street space is essential for urban street-design guidelines and road-network optimization [80]. Through axial and grid-based analyses, space syntax provides a powerful framework for identifying high integration. Such information can guide the placement of landmarks, directional signage, and other environmental cues that facilitate cognitive map formation and enhance navigability for diverse user groups. Moreover, because spatial integration is closely tied to navigation difficulty [29], purposeful modulation of integration levels can help calibrate the cognitive demands of navigation tasks. From an SG perspective, the majority of participants verbally reported a high level of immersion and task engagement, suggesting that the SG-based framework effectively supported sustained interaction with the spatial environment. By integrating principles of spatial configuration into an immersive VR navigation game, this approach may provide promising opportunities for the development of gamified tools for cognitive assessment, training, and rehabilitation, particularly for individuals with spatial or cognitive impairments.

Limitations and Future Work

This research highlights the critical role of cognitive abilities and spatial configuration in shaping navigation behavior. However, several limitations should be acknowledged. First, the participant sample may not fully represent the broader population. The young participants were primarily between 20 and 30 years of age, whereas the older group ranged from 65 to 85 years, leaving a substantial gap between ages 30 and 65 years, which was not included. This omission limits the generalizability of the findings, as individuals in midlife may exhibit distinct navigation patterns and cognitive characteristics that were not captured in the current analysis [72]. Second, the individualized k-means clustering approach adopted in this study is particularly suited to navigation games characterized by relatively stable movement speeds. It may be less appropriate for tasks involving pronounced, irregular, or continuous speed transitions, in which clear and interpretable cluster boundaries are difficult to establish. Moreover, although a 3-cluster solution showed good behavioral interpretability within the current

navigation game, the generalizability of this clustering scheme to other navigation games or more complex real-world environments remains to be validated in future studies. Third, discrepancies between virtual and real-world navigation are unavoidable [76]. The absence of environmental features such as directional signs, combined with the lack of natural physical movement, may influence how spatial information is encoded, updated, and retrieved [81]. Finally, the study did not explicitly account for the influence of demographic factors such as gender, educational background, or geographic and cultural experience. Ignoring these variables could obscure important individual differences and interactions relevant to wayfinding performance.

Future research should aim to address these limitations by recruiting a more demographically diverse participant sample, including middle-aged individuals, and by designing VEs with greater ecological validity. Incorporating navigation methods that more closely approximate natural physical movement may yield richer behavioral data and uncover additional nuances in navigational performance. Additionally, systematically examining the role of experienced metrics could provide deeper insights into urban design, spatial planning, and navigation-aid development in both virtual and real-world contexts. From a rehabilitation perspective, the game demonstrates promising potential for virtual navigation tasks as a controlled and adaptable platform for personalized cognitive assessment and training. Further research could investigate the incorporation of adaptive task difficulty and longitudinal monitoring into VR-based rehabilitation, thereby supporting the maintenance or recovery of spatial and cognitive functions.

Conclusions

This study provides quantitative evidence for how environmental configuration and cognitive abilities jointly shape navigational behavior in VEs. By applying an individualized k-means clustering approach, we extracted key behavioral indicators, such as wayfinding time and navigation efficiency, that proved sensitive to cognitive decline. This method offers a more personalized and potentially more effective framework for assessing cognitive function through navigation tasks. Through spatial analysis, we further identified age-specific regions of wayfinding difficulty and quantified the extent to which environmental structure influences navigation behavior. These findings reveal distinct navigation patterns between younger and older adults and underscore the value of integrating behavioral analytics with spatial-syntax metrics. Overall, the results lay an important foundation for future applications in clinical cognitive assessment and rehabilitation, as well as in the development of age-friendly urban environments.

Funding

This work was supported by the National Nature Science Foundation of China (12372326, 12472302, T2288101) and the Beijing Municipal Natural Science Foundation (L244020).

Authors' Contributions

XL and ST contributed equally as co-corresponding authors. ST can be contacted by email at shantian@buaa.edu.cn or by phone at +86-15010725395.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Instructions presented on the desktop during each phase of the navigation game.

[[PNG File, 268 KB - games_v14i1e83128_app1.png](#)]

Multimedia Appendix 2

(A) and (B) show the sum of squared errors and silhouette coefficient, respectively, for younger and older adults across different values of k .

[[PNG File, 253 KB - games_v14i1e83128_app2.png](#)]

Multimedia Appendix 3

(A) and (B) show all instantaneous speed samples from older and young participants, respectively. Blue, green, and yellow points represent low-, medium-, and high-speed clusters, respectively, and black points indicate the cluster centroids. (C) and (D) present the probability distributions of navigation speed for older and younger participants, respectively, with dark blue representing the group mean and light blue indicating individual distributions.

[[PNG File, 2365 KB - games_v14i1e83128_app3.png](#)]

References

1. Liu B, Dong WH, Zhan ZC, Wang SK, Meng LQ. Differences in the gaze behaviours of pedestrians navigating between regular and irregular road patterns. *ISPRS Int J Geo-Inf* 2020 Jan;9(1):45. [doi: [10.3390/ijgi9010045](#)]
2. Lester AW, Moffat SD, Wiener JM, Barnes CA, Wolbers T. The aging navigational system. *Neuron* 2017 Aug 30;95(5):1019-1035. [doi: [10.1016/j.neuron.2017.06.037](#)] [Medline: [28858613](#)]
3. Moffat SD. Aging and spatial navigation: what do we know and where do we go? *Neuropsychol Rev* 2009 Dec;19(4):478-489. [doi: [10.1007/s11065-009-9120-3](#)] [Medline: [19936933](#)]
4. Guariglia CC, Nitrini R. Topographical disorientation in Alzheimer's disease. *Arq Neuropsiquiatr* 2009 Dec;67(4):967-972. [doi: [10.1590/s0004-282x2009000600001](#)] [Medline: [20069202](#)]
5. Iaria G. Developmental topographical disorientation: lost every day. *Lancet Neurol* 2013 Aug;12(8):745. [doi: [10.1016/S1474-4422\(13\)70133-9](#)]
6. Igarashi KM. Entorhinal cortex dysfunction in Alzheimer's disease. *Trends Neurosci* 2023 Feb;46(2):124-136. [doi: [10.1016/j.tins.2022.11.006](#)] [Medline: [36513524](#)]
7. Stangl M, Achtzehn J, Huber K, Dietrich C, Tempelmann C, Wolbers T. Compromised grid-cell-like representations in old age as a key mechanism to explain age-related navigational deficits. *Curr Biol* 2018 Apr 2;28(7):1108-1115. [doi: [10.1016/j.cub.2018.02.038](#)] [Medline: [29551413](#)]
8. Kinkhabwala AA, Gu Y, Aronov D, Tank DW. Visual cue-related activity of cells in the medial entorhinal cortex during navigation in virtual reality. *Elife* 2020 Mar 9;9:e43140. [doi: [10.7554/eLife.43140](#)] [Medline: [32149601](#)]
9. Zhang TT, Lian ZF, Xu YN. Combining GPS and space syntax analysis to improve understanding of visitor temporal-spatial behaviour: a case study of the Lion Grove in China. *Landsc Res* 2020 May 18;45(4):534-546. [doi: [10.1080/01426397.2020.1730775](#)]
10. Hegarty M, He CXY, Boone AP, Yu SY, Jacobs EG, Chrastil ER. Understanding differences in wayfinding strategies. *Top Cogn Sci* 2023 Jan;15(1):102-119. [doi: [10.1111/tops.12592](#)] [Medline: [34973064](#)]
11. Kafri M, Weiss PL, Zeilig G, Bondi M, Baum-Cohen I, Kizony R. Performance in complex life situations: effects of age, cognition, and walking speed in virtual versus real life environments. *J Neuroeng Rehabil* 2021 Feb 8;18(1):30. [doi: [10.1186/s12984-021-00830-6](#)] [Medline: [33557894](#)]
12. Coutrot A, Manley E, Goodroe S, et al. Entropy of city street networks linked to future spatial navigation ability. *Nature New Biol* 2022 Apr 7;604(7904):104-110. [doi: [10.1038/s41586-022-04486-7](#)]
13. Brunec IK, Nantais MM, Sutton JE, Epstein RA, Newcombe NS. Exploration patterns shape cognitive map learning. *Cognition* 2023 Apr;233:105360. [doi: [10.1016/j.cognition.2022.105360](#)] [Medline: [36549130](#)]
14. Pagkratidou M, Galati A, Avraamides M. Do environmental characteristics predict spatial memory about unfamiliar environments? *Spat Cogn Comput* 2020 Jan 2;20(1):1-32. [doi: [10.1080/13875868.2019.1676248](#)]
15. Merhav M, Fisher-Gewirtzman D. How pathways' configuration impacts wayfinding in young and older adults. *J Environ Psychol* 2023 Sep;90:102065. [doi: [10.1016/j.jenvp.2023.102065](#)]

16. Krupic J, Bauza M, Burton S, Barry C, O'Keefe J. Grid cell symmetry is shaped by environmental geometry. *Nature New Biol* 2015 Feb 12;518(7538):232-235. [doi: [10.1038/nature14153](https://doi.org/10.1038/nature14153)] [Medline: [25673417](https://pubmed.ncbi.nlm.nih.gov/25673417/)]
17. Epstein RA, Patai EZ, Julian JB, Spiers HJ. The cognitive map in humans: spatial navigation and beyond. *Nat Neurosci* 2017 Oct 26;20(11):1504-1513. [doi: [10.1038/nn.4656](https://doi.org/10.1038/nn.4656)] [Medline: [29073650](https://pubmed.ncbi.nlm.nih.gov/29073650/)]
18. Rodgers MK, Sindone JA III, Moffat SD. Effects of age on navigation strategy. *Neurobiol Aging* 2012 Jan;33(1):202. [doi: [10.1016/j.neurobiolaging.2010.07.021](https://doi.org/10.1016/j.neurobiolaging.2010.07.021)] [Medline: [20832911](https://pubmed.ncbi.nlm.nih.gov/20832911/)]
19. Tversky B. Structures of mental spaces: how people think about space. *Environ Behav* 2003;35(1):66-80. [doi: [10.1177/0013916502238865](https://doi.org/10.1177/0013916502238865)]
20. Bécu M, Sheynikhovich D, Tatur G, et al. Age-related preference for geometric spatial cues during real-world navigation. *Nat Hum Behav* 2020 Jan;4(1):88-99. [doi: [10.1038/s41562-019-0718-z](https://doi.org/10.1038/s41562-019-0718-z)] [Medline: [31548677](https://pubmed.ncbi.nlm.nih.gov/31548677/)]
21. Harris MA, Wiener JM, Wolbers T. Aging specifically impairs switching to an allocentric navigational strategy. *Front Aging Neurosci* 2012;4(29):29. [doi: [10.3389/fnagi.2012.00029](https://doi.org/10.3389/fnagi.2012.00029)] [Medline: [23125833](https://pubmed.ncbi.nlm.nih.gov/23125833/)]
22. van der Ham IJM, Claessen MHG. How age relates to spatial navigation performance: functional and methodological considerations. *Ageing Res Rev* 2020 Mar;58:101020. [doi: [10.1016/j.arr.2020.101020](https://doi.org/10.1016/j.arr.2020.101020)] [Medline: [31954190](https://pubmed.ncbi.nlm.nih.gov/31954190/)]
23. Hillier B, Hanson J. *The Social Logic of Space*: Cambridge University Press; 1984. [doi: [10.1017/CBO9780511597237](https://doi.org/10.1017/CBO9780511597237)]
24. Jiang B, Claramunt C, Klarqvist B. Integration of space syntax into GIS for modelling urban spaces. *Int J Appl Earth Obs Geoinf* 2000;2(3-4):161-171. [doi: [10.1016/S0303-2434\(00\)85010-2](https://doi.org/10.1016/S0303-2434(00)85010-2)]
25. Hillier B. *Space Is the Machine: A Configurational Theory of Architecture*: Cambridge University Press; 1998. URL: <https://discovery.ucl.ac.uk/id/eprint/3881/1/SITM.pdf> [accessed 2026-03-09]
26. Peponis J, Zimring C, Choi YK. Finding the building in wayfinding. *Environ Behav* 1990 Sep;22(5):555-590. [doi: [10.1177/0013916590225001](https://doi.org/10.1177/0013916590225001)]
27. De Cock L, Van de Weghe N, Ooms K, et al. Linking the cognitive load induced by route instruction types and building configuration during indoor route guidance, a usability study in VR. *Int J Geogr Inf Sci* 2022 Oct 3;36(10):1978-2008. [doi: [10.1080/13658816.2022.2032080](https://doi.org/10.1080/13658816.2022.2032080)]
28. Li R, Klippel A. Wayfinding behaviors in complex buildings: the impact of environmental legibility and familiarity. *Environ Behav* 2016 Apr;48(3):482-510. [doi: [10.1177/0013916514550243](https://doi.org/10.1177/0013916514550243)]
29. Yesiltepe D, Fernández Velasco P, Coutrot A, et al. Entropy and a sub-group of geometric measures of paths predict the navigability of an environment. *Cognition* 2023 Jul;236:105443. [doi: [10.1016/j.cognition.2023.105443](https://doi.org/10.1016/j.cognition.2023.105443)] [Medline: [37003236](https://pubmed.ncbi.nlm.nih.gov/37003236/)]
30. Neo JRJ, Won AS, Shepley MM. Designing immersive virtual environments for human behavior research. *Front Virtual Real* 2021 Mar;2:603750. [doi: [10.3389/frvir.2021.603750](https://doi.org/10.3389/frvir.2021.603750)]
31. Slater M, Wilbur S. A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence: Teleop Virtual Environ* 1997 Dec;6(6):603-616. [doi: [10.1162/pres.1997.6.6.603](https://doi.org/10.1162/pres.1997.6.6.603)]
32. Checa D, Bustillo A. A review of immersive virtual reality serious games to enhance learning and training. *Multimed Tools Appl* 2020 Mar;79(9-10):5501-5527. [doi: [10.1007/s11042-019-08348-9](https://doi.org/10.1007/s11042-019-08348-9)]
33. Martinez ES, Wu AS, McMahan RP. Research trends in virtual reality locomotion techniques. Presented at: 2022 IEEE on Conference Virtual Reality and 3D User Interfaces (VR); Mar 12-16, 2022. [doi: [10.1109/VR51125.2022.00046](https://doi.org/10.1109/VR51125.2022.00046)]
34. Martelli D, Prado A, Xia BX, Verghese J, Agrawal SK. Development of a virtual floor maze test-effects of distal visual cues and correlations with executive function in healthy adults. *IEEE Trans Neural Syst Rehabil Eng* 2019 Oct;27(10):2229-2236. [doi: [10.1109/TNSRE.2019.2938103](https://doi.org/10.1109/TNSRE.2019.2938103)] [Medline: [31478863](https://pubmed.ncbi.nlm.nih.gov/31478863/)]
35. Roth D, Purps CF, Neumann WJ. A virtual Morris water maze to study neurodegenerative disorders. Presented at: 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct); Nov 9-13, 2020. [doi: [10.1109/ISMAR-Adjunct51615.2020.00048](https://doi.org/10.1109/ISMAR-Adjunct51615.2020.00048)]
36. Jacobs J, Weidemann CT, Miller JF, et al. Direct recordings of grid-like neuronal activity in human spatial navigation. *Nat Neurosci* 2013 Sep;16(9):1188-1190. [doi: [10.1038/nn.3466](https://doi.org/10.1038/nn.3466)] [Medline: [23912946](https://pubmed.ncbi.nlm.nih.gov/23912946/)]
37. Irshad S, Perkiş A, Azam W. Wayfinding in virtual reality serious game: an exploratory study in the context of user perceived experiences. *Appl Sci (Basel)* 2021;11(17):7822. [doi: [10.3390/app11177822](https://doi.org/10.3390/app11177822)]
38. De Leeuw D, De Maeyer P, De Cock L. A gamification-based approach on indoor wayfinding research. *ISPRS Int J Geo-Inf* 2020;9(7):423. [doi: [10.3390/ijgi9070423](https://doi.org/10.3390/ijgi9070423)]
39. Coughlan G, Coutrot A, Khondoker M, Minihane AM, Spiers H, Hornberger M. Toward personalized cognitive diagnostics of at-genetic-risk Alzheimer's disease. *Proc Natl Acad Sci U S A* 2019 May 7;116(19):9285-9292. [doi: [10.1073/pnas.1901600116](https://doi.org/10.1073/pnas.1901600116)] [Medline: [31015296](https://pubmed.ncbi.nlm.nih.gov/31015296/)]
40. Coutrot A, Silva R, Manley E, et al. Global determinants of navigation ability. *Curr Biol* 2018 Sep 10;28(17):2861-2866.e4. [doi: [10.1016/j.cub.2018.06.009](https://doi.org/10.1016/j.cub.2018.06.009)] [Medline: [30100340](https://pubmed.ncbi.nlm.nih.gov/30100340/)]
41. West GL, Patai ZE, Coutrot A, Hornberger M, Bohbot VD, Spiers HJ. Landmark-dependent navigation strategy declines across the human life-span: evidence from over 37,000 participants. *J Cogn Neurosci* 2023 Mar 1;35(3):452-467. [doi: [10.1162/jocn_a_01956](https://doi.org/10.1162/jocn_a_01956)] [Medline: [36603038](https://pubmed.ncbi.nlm.nih.gov/36603038/)]
42. Tragantzopoulou P, Giannouli V. Spatial orientation assessment in the elderly: a comprehensive review of current tests. *Brain Sci* 2024 Sep 5;14(9):898. [doi: [10.3390/brainsci14090898](https://doi.org/10.3390/brainsci14090898)] [Medline: [39335393](https://pubmed.ncbi.nlm.nih.gov/39335393/)]

43. Bartoschek T, Schwering A, Li R, Münzer S, Carlos V. OriGami: a mobile geogame for spatial literacy. In: Ahlqvist O, Schlieder C, editors. *Geogames and Geoplay: Game-Based Approaches to the Analysis of Geo-Information*: Springer; 2018:37-62. [doi: [10.1007/978-3-319-22774-0_3](https://doi.org/10.1007/978-3-319-22774-0_3)]
44. Schlieder C, Kiefer P, Matyas S. Geogames: designing location-based games from classic board games. *IEEE Intell Syst* 2006;21(5):40-46. [doi: [10.1109/MIS.2006.93](https://doi.org/10.1109/MIS.2006.93)]
45. Sánchez-Escudero JP, Galvis-Herrera AM, Sánchez-Trujillo D, et al. Virtual reality and serious videogame-based instruments for assessing spatial navigation in Alzheimer's disease: a systematic review of psychometric properties. *Neuropsychol Rev* 2025 Mar;35(1):77-101. [doi: [10.1007/s11065-024-09633-7](https://doi.org/10.1007/s11065-024-09633-7)] [Medline: [38403731](https://pubmed.ncbi.nlm.nih.gov/38403731/)]
46. Wilczyńska D, Walczak-Kozłowska T, Alarcón D, et al. The role of immersive experience in anxiety reduction: evidence from virtual reality sessions. *Brain Sci* 2025 Jan;15(1):14. [doi: [10.3390/brainsci15010014](https://doi.org/10.3390/brainsci15010014)]
47. Zhu S, Lakshminarasimhan KJ, Arfaei N, Angelaki DE. Eye movements reveal spatiotemporal dynamics of visually-informed planning in navigation. *Elife* 2022 May 3;11:e73097. [doi: [10.7554/eLife.73097](https://doi.org/10.7554/eLife.73097)] [Medline: [35503099](https://pubmed.ncbi.nlm.nih.gov/35503099/)]
48. Yi Z, Vaupel JW. Functional capacity and self-evaluation of health and life of oldest old in China. *J Soc Issues* 2002 Jan;58(4):733-748. [doi: [10.1111/1540-4560.00287](https://doi.org/10.1111/1540-4560.00287)]
49. Duan J, Lv YB, Gao X, et al. Association of cognitive impairment and elderly mortality: differences between two cohorts ascertained 6-years apart in China. *BMC Geriatr* 2020 Jan 28;20(1):29. [doi: [10.1186/s12877-020-1424-4](https://doi.org/10.1186/s12877-020-1424-4)] [Medline: [31992221](https://pubmed.ncbi.nlm.nih.gov/31992221/)]
50. Nasreddine ZS, Phillips NA, Bédirian V, 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](https://doi.org/10.1111/j.1532-5415.2005.53221.x)]
51. Li Z, Gong X, Wang S, et al. Cognitive impairment assessed by Mini-Mental State Examination predicts all-cause and CVD mortality in Chinese older adults: a 10-year follow-up study. *Front Public Health* 2022;10:908120. [doi: [10.3389/fpubh.2022.908120](https://doi.org/10.3389/fpubh.2022.908120)] [Medline: [36518570](https://pubmed.ncbi.nlm.nih.gov/36518570/)]
52. Paper versions. MoCA Cognition. URL: <https://mocacognition.com/paper/> [accessed 2026-03-13]
53. Shin MS, Park SY, Park SR, Seol SH, Kwon JS. Clinical and empirical applications of the Rey-Osterrieth Complex Figure Test. *Nat Protoc* 2006;1(2):892-899. [doi: [10.1038/nprot.2006.115](https://doi.org/10.1038/nprot.2006.115)] [Medline: [17406322](https://pubmed.ncbi.nlm.nih.gov/17406322/)]
54. Giovagnoli AR, Del Pesce M, Mascheroni S, Simoncelli M, Laiacina M, Capitani E. Trail making test: normative values from 287 normal adult controls. *Ital J Neurol Sci* 1996 Aug;17(4):305-309. [doi: [10.1007/BF01997792](https://doi.org/10.1007/BF01997792)] [Medline: [8915764](https://pubmed.ncbi.nlm.nih.gov/8915764/)]
55. Shepard RN, Metzler J. Mental rotation of three-dimensional objects. *Science* 1971 Feb 19;171(3972):701-703. [doi: [10.1126/science.171.3972.701](https://doi.org/10.1126/science.171.3972.701)] [Medline: [5540314](https://pubmed.ncbi.nlm.nih.gov/5540314/)]
56. Hölscher C, Brösamle M, Vrachliotis G. Challenges in multilevel wayfinding: a case study with the space syntax technique. *Environ Plann B Plann Des* 2012 Feb;39(1):63-82. [doi: [10.1068/b34050t](https://doi.org/10.1068/b34050t)]
57. Tomé A, Kuipers M, Pinheiro T, Nunes M, Heitor T. Space-use analysis through computer vision. *Autom Constr* 2015 Sep;57:80-97. [doi: [10.1016/j.autcon.2015.04.013](https://doi.org/10.1016/j.autcon.2015.04.013)]
58. Hirtle SC, Jonides J. Evidence of hierarchies in cognitive maps. *Mem Cognit* 1985 May;13(3):208-217. [doi: [10.3758/bf03197683](https://doi.org/10.3758/bf03197683)] [Medline: [4046821](https://pubmed.ncbi.nlm.nih.gov/4046821/)]
59. Bendjedidi S, Bada Y, Meziani R. Urban plaza design process using space syntax analysis: El-Houria plaza, Biskra, Algeria. *Int Rev Spat Plan Sustain Dev A Plan Strateg Des* 2019;7(2):125-142. [doi: [10.14246/irspsda.7.2_125](https://doi.org/10.14246/irspsda.7.2_125)]
60. Iftikhar H, Shah P, Luximon Y. Human wayfinding behaviour and metrics in complex environments: a systematic literature review. *Archit Sci Rev* 2021;64:452-643. [doi: [10.1080/00038628.2020.1777386](https://doi.org/10.1080/00038628.2020.1777386)]
61. Montello DR. Navigation. In: Shah P, Miyake A, editors. *The Cambridge Handbook of Visuospatial Thinking*: Cambridge University Press; 2005:257-294. [doi: [10.1017/CBO9780511610448.008](https://doi.org/10.1017/CBO9780511610448.008)]
62. Xia J, Arrowsmith C, Jackson M, Cartwright W. The wayfinding process relationships between decision-making and landmark utility. *Tour Manag* 2008 Jun;29(3):445-457. [doi: [10.1016/j.tourman.2007.05.010](https://doi.org/10.1016/j.tourman.2007.05.010)]
63. Irmischer IJ, Clarke KC. Measuring and modeling the speed of human navigation. *Cartogr Geogr Inf Sci* 2018 Mar 4;45(2):177-186. [doi: [10.1080/15230406.2017.1292150](https://doi.org/10.1080/15230406.2017.1292150)]
64. MacQueen J. Some methods for classification and analysis of multivariate observations. : University of California Press; 1967 URL: <https://scispace.com/pdf/some-methods-for-classification-and-analysis-of-multivariate-4pswti19oz.pdf> [accessed 2026-03-09]
65. Archdeacon TJ. *Correlation and Regression Analysis: A Historian's Guide*: University of Wisconsin Press; 1994. URL: <https://history.wisc.edu/publications/correlation-and-regression-analysis-a-historians-guide/> [accessed 2026-03-09]
66. Thinsungnoen T, Kaoungku N, Durongdumronchai P, Kerdprasop K, Kerdprasop N. The clustering validity with silhouette and sum of squared errors. In: Kerdprasop N, editor. 2015 Presented at: International Conference on Industrial Application Engineering 2015; Mar 28-31, 2015. [doi: [10.12792/iciae2015.012](https://doi.org/10.12792/iciae2015.012)]
67. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987 Nov;20:53-65. [doi: [10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)]
68. Claessen MHG, van der Ham IJM. Classification of navigation impairment: a systematic review of neuropsychological case studies. *Neurosci Biobehav Rev* 2017 Feb;73:81-97. [doi: [10.1016/j.neubiorev.2016.12.015](https://doi.org/10.1016/j.neubiorev.2016.12.015)] [Medline: [27993606](https://pubmed.ncbi.nlm.nih.gov/27993606/)]
69. Wiener JM, Büchner SJ, Hölscher C. Taxonomy of human wayfinding tasks: a knowledge-based approach. *Spat Cogn Comput* 2009 May 20;9(2):152-165. [doi: [10.1080/13875860902906496](https://doi.org/10.1080/13875860902906496)]

70. Cogné M, Taillade M, N’Kaoua B, et al. The contribution of virtual reality to the diagnosis of spatial navigation disorders and to the study of the role of navigational aids: a systematic literature review. *Ann Phys Rehabil Med* 2017 Jun;60(3):164-176. [doi: [10.1016/j.rehab.2015.12.004](https://doi.org/10.1016/j.rehab.2015.12.004)] [Medline: [27017533](https://pubmed.ncbi.nlm.nih.gov/27017533/)]
71. Wei Y, Liu J, Jin L, et al. Individual behavior and attention distribution during wayfinding for emergency shelter: an eye-tracking study. *Sustainability* 2023 Aug;15(15):11880. [doi: [10.3390/su151511880](https://doi.org/10.3390/su151511880)]
72. Kalantari S, Mostafavi A, Xu TB, Lee AS, Yang Q. Comparing spatial navigation in a virtual environment vs. an identical real environment across the adult lifespan. *Comput Human Behav* 2024 Aug;157:108210. [doi: [10.1016/j.chb.2024.108210](https://doi.org/10.1016/j.chb.2024.108210)]
73. Byrne P, Becker S, Burgess N. Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychol Rev* 2007 Apr;114(2):340-375. [doi: [10.1037/0033-295X.114.2.340](https://doi.org/10.1037/0033-295X.114.2.340)] [Medline: [17500630](https://pubmed.ncbi.nlm.nih.gov/17500630/)]
74. Luo CM, Proctor RW. How different direct association routes influence the indirect route in the same Simon-like task. *Psychol Res* 2019 Nov;83(8):1733-1748. [doi: [10.1007/s00426-018-1024-5](https://doi.org/10.1007/s00426-018-1024-5)]
75. Jin WJ, Qin H, Zhang K, Chen XW. Spatial navigation. In: Wang H, editor. *Neural Circuits of Innate Behaviors*: Springer; 2020:63-90. [doi: [10.1007/978-981-15-7086-5_7](https://doi.org/10.1007/978-981-15-7086-5_7)]
76. Savino GL, Emanuel N, Kowalzik S, et al. Comparing pedestrian navigation methods in virtual reality and real life. Presented at: ICMI '19: 2019 International Conference on Multimodal Interaction; Oct 14-18, 2019. [doi: [10.1145/3340555.335374](https://doi.org/10.1145/3340555.335374)]
77. Hirsh JB, Mar RA, Peterson JB. Psychological entropy: a framework for understanding uncertainty-related anxiety. *Psychol Rev* 2012 Apr;119(2):304-320. [doi: [10.1037/a0026767](https://doi.org/10.1037/a0026767)] [Medline: [22250757](https://pubmed.ncbi.nlm.nih.gov/22250757/)]
78. Lacreuse A, Raz N, Schmidtke D, Hopkins WD, Herndon JG. Age-related decline in executive function as a hallmark of cognitive ageing in primates: an overview of cognitive and neurobiological studies. *Philos Trans R Soc Lond B Biol Sci* 2020 Nov 9;375(1811):20190618. [doi: [10.1098/rstb.2019.0618](https://doi.org/10.1098/rstb.2019.0618)] [Medline: [32951543](https://pubmed.ncbi.nlm.nih.gov/32951543/)]
79. Xu TB, Govani VN, Kalantari S. Age differences in pedestrian navigational skills and performance: a systemic review and meta-analysis. *Ageing Res Rev* 2025 Feb;104:102591. [doi: [10.1016/j.arr.2024.102591](https://doi.org/10.1016/j.arr.2024.102591)] [Medline: [39710072](https://pubmed.ncbi.nlm.nih.gov/39710072/)]
80. Ma Q, Wei X, Ren G. Street design guidelines and urban road system improvement: from traffic volume to space quality. *Transp Res Procedia* 2025;82:3266-3282. [doi: [10.1016/j.trpro.2024.12.095](https://doi.org/10.1016/j.trpro.2024.12.095)]
81. van der Ham IJM, Faber AME, Venselaar M, van Kreveld MJ, Löffler M. Ecological validity of virtual environments to assess human navigation ability. *Front Psychol* 2015;6(637):637. [doi: [10.3389/fpsyg.2015.00637](https://doi.org/10.3389/fpsyg.2015.00637)] [Medline: [26074831](https://pubmed.ncbi.nlm.nih.gov/26074831/)]

Abbreviations

AMA: axial map analysis
EAI: experienced axial integration
EMAD: experienced mean axial depth
EVI: experienced visual integration
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
MRT: Mental Rotation Test
ROCF-C: Copy condition of the Rey-Osterrieth Complex Figure
SG: serious game
SSE: sum of squared errors
TMTA: Part A of the Trail Making Test
VE: virtual environment
VGA: visibility graph analysis
VR: virtual reality

Edited by M Schmid; submitted 30.Aug.2025; peer-reviewed by Q Yang, TB Xu; revised version received 03.Feb.2026; accepted 15.Feb.2026; published 02.Apr.2026.

Please cite as:

Qiao X, Tian S, Tang M, He S, Wang J, Fan L, Zhu Y, Zhang Z, Du S, Dong C, Chen Y, Liu X
Analysis of Spatiotemporal Features in a Virtual Navigation Game Across Different Age Groups: Quantitative Research
JMIR Serious Games 2026;14:e83128
URL: <https://games.jmir.org/2026/1/e83128>
doi: [10.2196/83128](https://doi.org/10.2196/83128)

© Xiaofeng Qiao, Shan Tian, Min Tang, Shipai He, Jinghui Wang, Linyuan Fan, Yuanjie Zhu, Zhiyang Zhang, Songjun Du, Chaojie Dong, Yepu Chen, Xiaoyu Liu. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 2.Apr.2026. 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.

Gamified Versus Nongamified Metaverse Learning for Breast Health Knowledge in Women: Randomized Controlled Trial

Rui Li, MA; Sabzali Musa Kahn, PhD; Bingyu Duan, MA; Seng Yue Wong, PhD

Dean Office, Academy of Malay Studies, Universiti Malaya, Kuala Lumpur, Malaysia

Corresponding Author:

Seng Yue Wong, PhD

Dean Office, Academy of Malay Studies, Universiti Malaya, Kuala Lumpur, Malaysia

Abstract

Background: The metaverse provides an immersive, interactive medium for health education, but most studies evaluate immersion and gamification together. Randomized evidence disentangling their separate effects on immediate learning and short-term retention in breast health education is lacking.

Objective: This study aimed to isolate the effects of gamification, over and above an identical immersive metaverse environment, on immediate gains and 4-week retention of women's breast health knowledge.

Methods: This 2-arm, parallel, individually randomized controlled trial was conducted in Hangzhou, China. Eligible participants were women aged ≥ 18 years who were interested in breast health and able to use a personal computer with internet access. A total of 80 women were recruited via the Xiaohongshu social media platform; 8 withdrew before randomization or did not complete the baseline assessment, and the remaining 72 women were randomized to a gamified metaverse (GM) group or a nongamified metaverse (NGM) group using a computer-generated 1:1 sequence. Both groups used the Mammaverse platform with identical educational content and exposure time. Breast health knowledge was assessed at baseline (T1), immediately postintervention (T2), and 4-week follow-up (T3) using the same questionnaire. The primary outcome was a change in knowledge score. Linear mixed-effects models were used, with age, education, family history of breast cancer, prior training, and baseline knowledge as covariates. Participants and investigators were not blinded.

Results: All randomized participants completed follow-up and were included in the analysis (GM group: $n=36$; NGM group: $n=36$), with no loss to follow-up. Knowledge scores improved in both groups, but gains from T1 to T2 were larger in the GM group than in the NGM group (Hedges $g=0.65$, 95% CI 0.18 - 1.12; $P=.007$). From T2 to T3, there was no between-group difference in change scores ($P=.91$). However, at 4 weeks, the GM group retained higher absolute knowledge than the NGM group (estimated marginal means 15.7 vs 13.0). No intervention-related adverse events were reported.

Conclusions: This study marks the first application of gamification in breast self-examination education for ordinary Chinese women within a 3D desktop metaverse. By comparing gamified and nongamified versions under identical metaverse platform conditions, it expands the application boundaries of the GM group in breast health education. Gamification significantly enhanced immediate acquisition of breast health knowledge but did not provide additional advantages for short-term retention. However, the gamified group maintained higher absolute knowledge levels at the 4-week follow-up. Overall, in the 3D desktop metaverse, immersive experiences provide foundational effects, while gamification delivers immediate gains. To further optimize long-term retention, memory consolidation strategies should be integrated into the gamified framework.

Trial Registration: ClinicalTrials.gov NCT06930898; <https://clinicaltrials.gov/study/NCT06930898>

(*JMIR Serious Games* 2026;14:e75318) doi:[10.2196/75318](https://doi.org/10.2196/75318)

KEYWORDS

gamification; metaverse; health education; digital intervention; breast health knowledge

Introduction

Digital health interventions have emerged as a vital approach for health education and behavior change [1]. In recent years, immersive technologies such as the metaverse have garnered significant attention for their ability to deliver heightened presence and interactivity, thereby effectively addressing the limitations of traditional educational models in practical

application and contextual learning [2]. In medical education, metaverse technology has demonstrated reusable learning benefits and accessibility advantages in anatomy and clinical skills training [3-5]. It also offers new implementation pathways for public health education, focusing on privacy protection, remote accessibility, and resource equity [6,7].

Meanwhile, gamification, as an effective motivational strategy, enhances engagement and behavioral adherence through designs

incorporating points, goals, feedback, leaderboards, and story-driven tasks. In health and medical education, gamification has repeatedly been shown to improve short- and medium-term memory retention as well as learning engagement [8,9]. However, evidence regarding its long-term effects remains inconsistent. Some studies indicate that purely competitive gamification designs may induce learning anxiety [10,11] or fail to effectively enhance long-term retention due to concerns about privacy, team collaboration, and technical barriers [12]. These phenomena collectively suggest that gamification effects may primarily amplify immediate learning drive, whereas long-term knowledge consolidation likely relies more on other cognitive mechanisms.

Although immersive technologies and gamified designs are often integrated in practice, research on the medical metaverse has primarily focused on specialized domains such as clinical training and surgical simulation [13-15], with limited systematic validation for public-facing breast health education targeting women. Enhancing breast health information literacy and knowledge is a critical prerequisite influencing women's screening intentions and health decisions [16]. However, existing digital interventions in this field are predominantly limited to two-dimensional web pages, minigames, or conceptual explorations [17-19].

More importantly, most studies combine immersion and gamification into a single integrated intervention or compare them only with traditional teaching methods, making it difficult to isolate the independent contributions of each mechanism [20-22]. Therefore, when observing positive effects on knowledge enhancement, a critical question arises: to what extent should these gains be attributed to the immersive experience provided by the metaverse, and to what extent to the motivational mechanisms of gamification? Distinguishing the unique contributions of different design elements is crucial for optimizing intervention efficiency, precisely controlling development costs, and guiding future personalized recommendations.

Based on this research, although a published mixed methods study has preliminarily confirmed the tool's feasibility and potential motivational mechanisms, the attribution of its positive effects remains unclear [23]. Therefore, this study was designed as a randomized controlled trial (RCT). By introducing an NGM control group, it aims to systematically isolate the unique contributions of immersion and gamification effects to answer the core question: does a gamified intervention in a metaverse environment significantly outperform a metaverse-only environment in enhancing and sustaining breast health knowledge?

This study proposes the following 2 hypotheses:

- H1: In immediate post-intervention measurements, the enhancement of breast health knowledge delivered by the GM group will be significantly greater than that of the NGM group.
- H2: In delayed measurements, the retention effect of breast health knowledge in the GM group will be significantly superior to that of the NGM group.

The unique contributions of this study are as follows: (1) conducting an element-level causal test of immersion versus gamification within the same metaverse platform; (2) distinguishing and quantifying immediate and retention effects using linear mixed models (LMMs) and estimated marginal means (EMMs); and (3) providing replicable digital intervention evidence for public health scenarios such as women's breast health education, along with design strategies for knowledge retention.

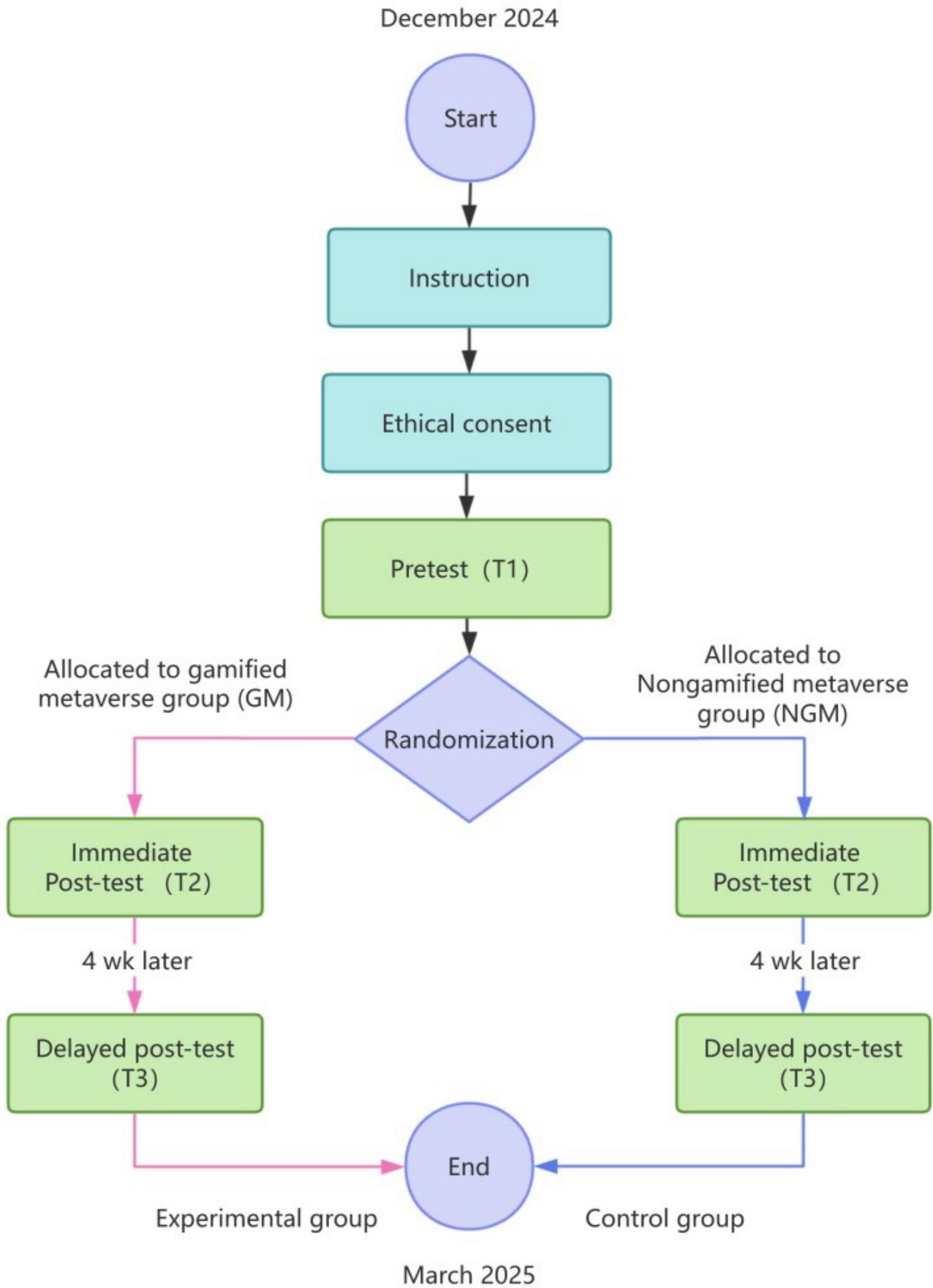
Methods

Study Design

This was a 2-arm, parallel-group RCT with 1:1 allocation. Participants were randomly assigned to 2 groups: the experimental group received a gamified health intervention within the metaverse (GM group), whereas the control group used a nongamified version of the same metaverse platform (NGM group). To assess both immediate and sustained effects of the intervention, 3 measurement points were established: preintervention (T1), immediate postintervention (T2), and delayed postintervention (T3), conducted 4 weeks after the intervention.

This study adhered to the CONSORT (Consolidated Standards of Reporting Trials) 2025 statement ([Checklist 1](#)) and the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth, version 1.6.1) guidelines ([Checklist 2](#)) to enhance the reproducibility and transparency of research on web-based and digital health interventions [24,25]. In accordance with these guidelines, the study was retrospectively registered at ClinicalTrials.gov (NCT06930898). This study did not involve patient or public participation in the design, implementation, or reporting of research findings. Although a small number of participants self-reported having breast-related health issues, they were included as members of the general adult population rather than as patient representatives or research collaborators. Throughout the trial, all participants received the same intervention procedures and measurement schedules, and no significant modifications were made to the study procedures or outcome indicators based on interim results. The overall study flow is illustrated in [Figure 1](#).

Figure 1. Study procedure and timeline.



Setting

Participant recruitment and data collection were conducted online in Hangzhou, Zhejiang Province, China. The intervention was delivered through a custom-developed 3D platform named “Mammoverse,” which integrates a metaverse-based immersive learning environment with gamified design principles.

The platform employs a semi-immersive 3D scene presentation mode, in which users navigate using a keyboard and mouse to complete learning activities, including exploration, tasks, quizzes, and feedback. This level of immersion strikes an optimal balance between information delivery and ease of interaction, enhancing participant focus while effectively reducing cognitive load, thereby supporting long-term knowledge retention [26]. Given the advantages of personal computer (PC) devices in stability and controllable interaction, the intervention was implemented remotely via a web-based platform, allowing participants to complete all learning and assessment procedures in their home or study environments.

Participants

Participants were recruited using a combination of snowball and convenience sampling to ensure sample feasibility and diversity. Researchers posted recruitment notices on Xiaohongshu, where eligible and willing individuals could contact the research team via private message, complete an application form, and provide basic demographic information. To broaden coverage, the recruitment process allowed participants to invite friends or family members to join.

A total of 72 eligible participants completed baseline measurements and were randomized. Researchers completed the study using the online random number generation platform Research Randomizer [27]. After researchers set parameters as 2 groups, 36 participants per group, ID range 1 to 72. The platform automatically generated 2 nonoverlapping random ID sequences, participants were sequentially assigned to either the GM intervention group or the NGM control group based on the system output. Because the randomization sequence was automatically generated by the online system, researchers cannot predict subsequent allocation outcomes prior to assignment. Grouping operations were then performed based on the pregenerated sequence. As the 2 intervention modes were presented as distinct buttons on the platform interface, participants could identify their assigned intervention group upon accessing the platform; therefore, blinding could not be implemented for either participants or intervention implementers.

Given the increasingly younger age of individuals at risk for breast health issues and the rising probability of such risks with age, the inclusion criteria for this study were as follows: (1) women aged 18 years or older; (2) interest in breast health topics; (3) ownership and proficiency in operating a PC, mouse, keyboard, and stable internet connection; and (4) voluntary participation in the study and signing of an electronic informed

consent form. Exclusion criteria applied to participants who were unable to meet any of the above requirements.

To ensure sufficient statistical power, this study conducted a pretest sample size estimation using G*Power version 3.1.9.6 software prior to implementation [28]. Based on a medium effect size ($f=0.25$), a significance level $\alpha=.05$, and test power of 0.95, the analysis model was specified as a repeated measures design with a between-group factor (GM group vs NGM group) and a within-subject time factor (T1, T2, T3), allowing for the examination of the interaction between group and time. Calculations indicated a minimum sample size of 44 participants (22 per group) to meet the statistical requirements. Ultimately, 72 participants (36 per group) were recruited, significantly exceeding the minimum requirement. This larger sample size enhances the stability and generalizability of the findings. Furthermore, the sample size of this study aligns with that of previous RCTs in the same field, enhancing the comparability and external validity of the findings [29].

Intervention

Both intervention conditions were based on the “Mammoverse” educational platform, an immersive 3D virtual clinic that enables users to interact with virtual characters and environments (Multimedia Appendix 1). This platform facilitates interactive learning through virtual characters within an immersive 3D virtual clinic setting. The overall design process of the platform has been systematically detailed in prior mixed methods research [23].

To examine the specific effects of knowledge enhancement, this study implemented 2 parallel intervention conditions. In the GM group, participants experienced the full gamified version described in the hybrid study [23]. This version employed a narrative-driven task flow to guide learning, where participants progressively mastered knowledge points through story-based missions. It also incorporated comprehensive incentive and feedback mechanisms, such as points, badges, and virtual rewards. In contrast, the NGM group used the same 3D educational content and scene structure but removed all gamification features and task guidance elements. In this version, participants could only freely explore the virtual clinic environment to acquire breast health knowledge, without narrative tasks or reward feedback mechanisms. Both intervention conditions were standardized in terms of platform configuration, interaction procedures, and delivery format, ensuring that the intervention protocols were fully replicable.

Both interventions ran on PCs using identical interaction methods, such as W/A/S/D keys for movement, the E key for interaction, and mouse operations. No harms or unintended effects were anticipated for either intervention. Participants were instructed to report any discomfort or adverse experiences during the study period. However, no such events were reported. Key functional differences between the 2 intervention conditions are summarized in Table 1.

Table . Functional comparison of the gamified and nongamified metaverse breast health education platforms.

Function	Gamified metaverse group	Nongamified metaverse group
Common core		
3D Immersive clinic	Yes	Yes
Core health knowledge content	Yes	Yes
Key differentiators		
Narrative context (story-driven experience)	Yes	No
Task-based process guidance	Yes	No
Points and rewards system	Yes	No
Badges	Yes	No
Learning modalities	Task-driven	Free exploration

Instrument

The breast health knowledge questionnaire used in this study was adapted from the instrument developed by McCance et al [30]. It comprised 22 items organized into 4 sections: Section A collected demographic information (4 items); Section B assessed knowledge of breast cancer risk factors and screening (6 items); Section C assessed knowledge of breast self-examination (BSE) procedures and key techniques (10 items); and Section D assessed recognition of abnormal breast changes and symptoms (2 items). Excluding the demographic items, the remaining 18 items were single-choice or true or false knowledge questions. Each correct response was scored as 1, and each incorrect or “don’t know” response as 0, yielding a total knowledge score ranging from 0 to 18, with higher scores indicating greater breast health knowledge. Because all participants were native Mandarin Chinese speakers, the questionnaire was developed and administered in Chinese.

Data Collection

The primary outcome measures of this study were the total scores for breast health knowledge at each time point and their changes relative to baseline. No secondary outcome measures were predefined. Therefore, data collection primarily centered on the 3 measurements of the breast health knowledge questionnaire. Participants first completed an online questionnaire primarily focused on breast health knowledge to assess their baseline knowledge level prior to the intervention. Immediately following the intervention, participants completed the same knowledge questionnaire as the pretest to evaluate the immediate impact on knowledge enhancement. Four weeks after the intervention concluded, all participants completed the same questionnaire again to assess the short-term retention of knowledge. All questionnaires were distributed online in PDF format, with participants submitting electronic versions through designated channels. The entire data collection process was remotely monitored by the research team to ensure completion. However, researchers did not provide any guidance or prompts during the questionnaire completion to guarantee data objectivity and consistency.

Data Analysis

All data were analyzed using IBM SPSS Statistics (version 28.0; IBM Corp), with complete and missing-value-free sample data;

thus, all randomly assigned participants were included in the primary analysis. First, descriptive statistics, including frequencies and percentages, were applied to participants’ demographic variables. Each item on the Breast Health Knowledge Questionnaire was scored as correct (1 point) or incorrect/unknown (0 points), with total scores at each time point representing knowledge levels.

To evaluate intervention effectiveness, LMMs were used to analyze score changes across groups and time points. The model included “knowledge score” as the dependent variable, with “group” (GM group vs NGM group) and “time” (T1, T2, T3) as fixed effects, and “participant ID” as a random effect to control for baseline differences between individuals. The primary focus was on examining the interaction effect between “group and time.” A significant interaction term ($P < .05$) would indicate statistically distinct trajectories of knowledge score change between the 2 groups. Post hoc comparisons of group differences were then conducted to further analyze score changes between T1 to T2 and T2 to T3, testing hypotheses H1 and H2. The above between-group difference comparisons represent exploratory post hoc analyses; this study did not prespecify or implement subgroup analyses or sensitivity analyses.

Before the model analysis, LMM assumptions were verified. The Q–Q plot indicated near-normal residual distribution, and the scatterplot of residuals versus predicted values showed no discernible patterns or heteroscedasticity, confirming good model fit (Multimedia Appendix 2). In addition to the prespecified intergroup comparisons described above, no subgroup analyses or sensitivity analyses were planned or performed in this study. All tests were 2-tailed, with a significance level set at $P < .05$.

Ethical Considerations

This study was registered at ClinicalTrials.gov (NCT06930898). The study protocol and statistical analysis plan were reviewed and approved by the University of Malaya Research Ethics Committee (Ethics ID: UM.TNC2/UMREC_3967). The full protocol and statistical analysis plan are not publicly accessible but can be obtained from the corresponding author upon reasonable request. All study procedures adhered to the principles of the Declaration of Helsinki and the ethical guidelines for human subjects research.

Part of the analyses in this manuscript involved the use of deidentified data collected in a previous study conducted under the same ethics approval. In the original study, written informed consent was obtained from all participants, and the consent form explicitly stated that their data might be used for future related research. Therefore, the secondary analysis of the existing dataset fell within the scope of the original ethics approval, and no new informed consent or additional ethics review was required.

The study method adjustments implemented in this RCT, including the newly added randomization procedure and 4-week follow-up assessment, were reviewed and approved by the ethics committee as protocol amendments. For the RCT component of the study, all participants received full information about the study and electronically signed informed consent forms prior to participation.

All research data were deidentified prior to analysis and contained no personally identifiable information. Data were used solely for academic research purposes. Participants who completed all assessments received a compensation of 20 RMB (approximately US \$2.8). All illustrative images, interface

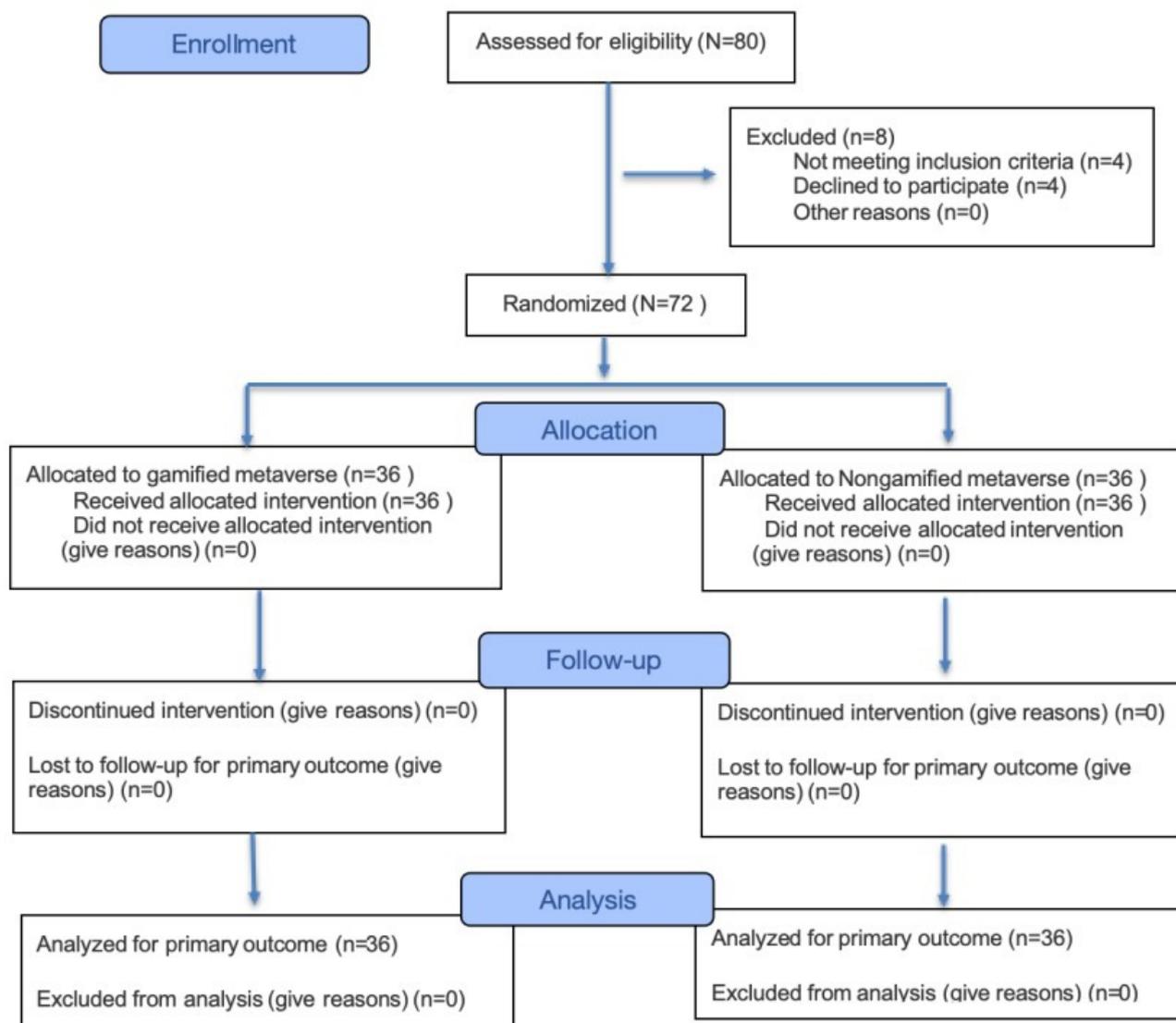
screenshots, and supplementary materials included in this manuscript contain no facial features or other identifiable personal information.

Results

Participant Flow and Recruitment

The recruitment and follow-up process for this study were completed between December 2024 and March 2025. A total of 80 women underwent eligibility assessment, of whom 8 were excluded prior to random assignment: 4 did not meet inclusion criteria and 4 withdrew for personal reasons before completing baseline measurements. The remaining 72 eligible participants completed the T1 pretest and were randomly assigned to either the GM group (n=36) or the NGM group (n=36). All randomly assigned participants received the intervention according to their assignment and completed follow-up assessments. No participants were lost to follow-up or additionally excluded after randomization. Consequently, all 72 participants from both groups were included in the primary outcome analysis. Participant flow is shown in Figure 2.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) 2025 flow diagram of participant recruitment, allocation, follow-up, and analysis.



Intervention Delivery and Adherence

The intervention was implemented according to a predefined protocol. After all participants completed baseline testing, researchers sent remote access links based on random assignment. Participants then independently experienced the Mammaverse platform via PCs for approximately 20 minutes.

Prior to the formal intervention, researchers provided brief, standardized verbal instructions on basic operations, such as login procedures, interface navigation, and fundamental interactions, to ensure participants could successfully access the platform. Beyond these foundational operational guidelines, researchers did not engage in real-time instruction or intervene in participants' learning content throughout the entire intervention period.

All randomly assigned participants received their respective interventions as per the study protocol. The intervention group completed gamified immersive learning tasks automatically presented by the platform, featuring incentive mechanisms such as points, badges, and leaderboards. The control group completed nongamified learning tasks identical in content to the intervention group but lacking gamification elements. The task flow and duration were identical for both groups. All participants completed all learning tasks required by the system prior to T2 (immediate test) and T3 (4 wk delayed test), with no dropouts or incomplete tasks, indicating good intervention

adherence. During the study, no intervention-related adverse events or other harms were reported in either group. Neither group received additional educational materials or parallel interventions related to breast health, nor did they receive any concurrent care that could potentially affect learning outcomes, thereby ensuring the independence of the intervention effects.

Baseline Characteristics

This study included 72 female participants who completed follow-up, with 36 in each of the GM group and the NGM group (Table 2). Both groups exhibited similar age distributions, with the majority aged 20 to 30 years (GM group: $n=29$, 80.56%; NGM group: $n=28$, 77.78%). Educational attainment was predominantly bachelor's degree or higher: 61.11% ($n=22$) in GM group and 33.33% ($n=12$) in NGM group held bachelor's degrees, while 52.78% ($n=19$) and 30.56% ($n=11$), respectively, held master's degrees or higher. Most participants had no family history of breast cancer ($n=31$, 86.11% in both groups), and a low proportion had received breast health-related training (GM group: $n=6$, 16.67%; NGM group: $n=4$, 11.11%). Regarding baseline outcomes, T1 breast health knowledge scores were comparable between groups (GM group: mean 9, SD 2.6; NGM group: mean 8.3, SD 3.8), with a standardized difference of 0.20, indicating overall good balance after randomization. As preplanned, subsequent effect assessments will adjust for T1 as a covariate in the model.

Table 2. Baseline characteristics of women in the gamified and nongamified metaverse groups (N=72).

Variable	Gamified metaverse group (n=36)	Nongamified metaverse group (n=36)
Age (y), n (%)		
<20	2 (5.56)	5 (13.89)
20 - 30	29 (80.56)	28 (77.78)
31 - 40	3 (8.33)	2 (5.56)
>40	2 (5.56)	1 (2.78)
Education, n (%)		
Preuniversity	1 (2.78)	1 (2.78)
College	1 (2.78)	5 (13.89)
Undergraduate	22 (61.11)	19 (52.78)
Postgraduate and above	12 (33.33)	11 (30.56)
Family history of breast cancer, n (%)		
Yes	3 (8.33)	1 (2.78)
No	31 (86.11)	31 (86.11)
Uncertain	2 (5.56)	4 (11.11)
Education or training related to breast health, n (%)		
Yes	6 (16.67)	4 (11.11)
No	28 (77.78)	31 (86.11)
Uncertain	2 (5.56)	1 (2.78)
Baseline outcome, mean (SD)		
T1 breast-health knowledge ^a	9.0 (2.6)	8.3 (3.8)

^aFor T1, the standardized mean difference was 0.20; no significance test is performed. The primary analysis adjusted for T1 as a covariate.

Immediate Knowledge Improvement: LMM Analysis

As shown in Table 3, the LMM revealed a significant main effect of group ($F_{1,64.75}=29.110$; $P<.001$), indicating that overall,

GM group participants scored significantly higher on knowledge measures than NGM group participants. This establishes the overall superiority of the gamified intervention.

Table . Linear mixed-effects model of breast health knowledge scores by group and time.

Effect	<i>F</i> test (<i>df</i>)	<i>P</i> value	Interpretation
Group	29.11 (1, 64.75)	<.001	Overall higher knowledge scores in the gamified group
Time	212.72 (2, 113.69)	<.001	Knowledge scores increased significantly over time
Group×Time	5.53 (2, 113.69)	.005	The improvement pattern differed between groups (significant interaction)

The main effect of time was significant ($F_{2, 113.69}=212.70$; $P<.001$), confirming that knowledge scores increased significantly over time during the intervention period (T1-T3), reflecting the effectiveness of the learning process.

The interaction between group and time was also significant ($F_{2, 113.69}=5.53$, $P=.005$). This key finding indicates that the growth patterns of knowledge scores were not parallel across

groups, with the gamified group exhibiting a significantly different improvement curve compared to the nongamified group. Therefore, we conducted between-group comparisons for T2-T1 and T3-T2 (see Table 4). Finally, after adjusting for age, education level, family history of breast cancer, prior training experience, and baseline (T1) scores, the primary findings remained consistent, with all $P>.28$ (Multimedia Appendix 3).

Table . Immediate knowledge improvement (T2-T1) and 4-week retention (T3-T2) in the gamified and nongamified groups.

Comparison	Group	Test ^a	Effect size ^b	<i>P</i> value	Interpretation	
	GM ^c (n=36)	NGM ^d (n=36)				
T2-T1 (immediate improvement)	8.47 (2.5) ^e	6.25 (4.05) ^e	$t_{58.35}=2.8$	Hedges $g=0.65$ (95% CI 0.18 - 1.12)	.007	Greater improvement in gamified group
T3-T2 (retention)	36.22 ^f	36.78 ^f	$z=-0.11$	$r=-0.02$.91	No difference in retention

^aTwo-sided $\alpha=.05$.

^bEffect size: Hedges g (for t test), r (for Mann-Whitney test).

^cGM: gamified metaverse.

^dNGM: nongamified metaverse.

^eMean (SD).

^fMean rank.

Table 5 presents the EMMs, showing significant improvements in both groups postintervention: GM group increased from 8.97 to 17.44, while NGM group rose from 8.33 to 14.58. At follow-up (T3), both groups experienced slight declines, but the GM group remained at 15.72, still higher than the NGM group at 12.97. Furthermore, the GM group maintained its

advantage at both post-test and follow-up, particularly in timely improvement. The GM group (17.44) averaged approximately 3 points higher than the NGM group (14.58), indicating a greater enhancement. This suggests gamification delivers additional immediate gains within the same metaverse environment, supporting H1.

Table . Estimated marginal means (EMMs) of breast health knowledge scores by group and time^a.

Group	T1 EMM (SE; 95% CI)	T2 EMM (SE; 95% CI)	T3 EMM (SE; 95% CI)
GM ^b	8.97 (0.41; 8.16 - 9.78)	17.44 (0.41; 16.64 - 18.25)	15.72 (0.41; 14.91 - 16.53)
NGM ^c	8.33 (0.41; 7.53 - 9.14)	14.58 (0.41; 13.78 - 15.39)	12.97 (0.41; 12.16 - 13.78)

^aEMMs from the linear mixed model (restricted maximum likelihood estimation; Satterthwaite df); SEs are shown in parentheses; 2-sided $\alpha=.05$.

^bGM: gamified metaverse.

^cNGM: nongamified metaverse.

Knowledge Retention: Post Hoc Tests

Table 4 presents between-group comparisons for the enhancement and maintenance phases. For immediate improvement (T2–T1), Welch independent-samples *t* test was applied due to unequal variance (Levene $P=.006$). Results indicated that the GM group's improvement (mean 8.47, SD 2.50) significantly exceeded the NGM group's (mean 6.25, SD 4.05): $t_{58,35}=2.80$, $P=.007$, with an effect size (Hedges $g=0.65$, 95% CI 0.18 - 1.12), indicating a medium effect. This demonstrates that incorporating gamification elements within the same metaverse learning environment significantly enhances immediate knowledge acquisition, directly supporting H1.

The retention phase (T3–T2) employed the Mann-Whitney *U* test to compare knowledge retention or decline rates between groups: $z=-0.11$, $P=.91$, rank-biserial $r=-0.02$. Results indicated a negligible effect approaching zero, with no significant difference, suggesting comparable overall knowledge retention across groups during follow-up. Thus, H2 was not supported.

Thus, both groups gained learning benefits in the metaverse environment, but gamification delivered an additional, medium-sized immediate gain. Regarding follow-up retention, no significant intergroup differences emerged, indicating gamification's limited impact on long-term retention.

Discussion

Findings

This RCT isolated and quantified the independent effects of immersive environments versus gamification mechanisms on breast health knowledge acquisition within the same Mammaverse platform. Under identical content and duration, the GM group significantly outperformed the NGM group only in "immediate enhancement," supporting H1. However, it did not demonstrate superiority in "retention rate," though scores at follow-up remained higher than the NGM group, failing to support H2. Despite comparable retention rates, the gamified group maintained higher knowledge levels at follow-up than the control (EMMs: 15.72 vs 12.97). These findings remained consistent after controlling for key covariates, including age, education level, family history of breast cancer, prior training experience, and baseline scores, further enhancing the robustness and reliability of the results.

From a mechanism perspective, both groups achieved significant improvements from T1 to T2. This finding strongly demonstrates that the metaverse scenario itself, as a highly immersive and interactive cognitive simulation environment, possesses fundamental pedagogical efficacy [31]. Building upon this foundation, gamification enhances short-term motivation and practice intensity through elements such as goal and task setting, scenario-based narratives, question-answer adjustments, and real-time feedback, thereby generating incremental effects during the immediate phase.

Based on H1, immediate postintervention measurements, the GM group significantly enhanced breast health knowledge more than the NGM group. The gamified "incremental effect" likely enhances learners' attention allocation and task persistence

through goal-setting and immediate feedback, reducing 'activation costs' to generate greater knowledge gains in a short time. Long-term retention relies more on consolidation strategies like spaced repetition and retrieval practice [32]. This aligns with Self-Determination Theory's fulfillment mechanisms of competence, autonomy, and relatedness, manifesting as greater immediate knowledge gains under identical content and duration. This "immediate gain" compounded with the "baseline gain" from the immersive environment resulted in more pronounced improvements at T2 [33–35].

The findings of this study regarding gamification's "instant gratification" align closely with the widely recognized positive effects of gamification on short-term knowledge acquisition and motivation enhancement in existing literature. The findings support Haruna's [9] observation that gamified platforms provide interactive and incentive mechanisms to effectively enhance students' learning engagement and knowledge acquisition. This immediate boost confirms gamification's role as a motivation-driven tool, significantly increasing user participation motivation and short-term focus through elements like points and rewards [36]. Particularly in health education, studies by Anderson, Romero-Alemán, Marcos, and others demonstrate that gamified interventions significantly improve participants' cognitive understanding and short-term retention of health knowledge. This aligns with the greater immediate gains observed in breast health knowledge among the GM group in this study [11,37,38]. This finding reduces the "startup costs" of learning by satisfying competence and autonomy in Self-Determination Theory, thereby maximizing knowledge transfer in a short timeframe. This aligns with Wang et al's [39] perspective that gamification enhances sustained engagement by influencing user perceptions. Although Tong and Hee [40] previously employed a two-dimensional online game for breast cancer education, this study demonstrated that gamified mechanisms (motivation-driven) layered upon an immersive environment (cognitive vehicle) yield significantly greater immediate knowledge gains than purely immersive learning when comparing GM group with NGM group. This effectively distinguishes the critical contributions of immersion versus gamification.

Based on H2, delayed measurement revealed that the breast health knowledge retention effect in the GM group did not significantly outperform that in the NGM group. The "baseline gains" observed in the immersive metaverse may stem from heightened presence and actionable contextual practice. On one hand, the 3D interactive environment and actionable contextual practice significantly enhanced learners' sense of presence, facilitating the formation of embodied cognition [41]. On the other hand, immersive experiences and immediate visual feedback help generate richer contextual encoding cues in the brain, effectively promoting active knowledge processing [42]. In other words, the immersive metaverse provides a high-fidelity cognitive vehicle for content learning, ensuring learners achieve significant knowledge gains with equivalent content and time investment [43].

The findings on long-term knowledge retention in this study provide new empirical evidence for understanding the cognitive vehicle role of immersive technologies (metaverse) and the

long-term limitations of gamification. The GM group's knowledge retention did not significantly outperform the NGM group, supporting existing literature indicating limitations of pure gamification mechanisms in long-term knowledge retention [8,10]. This suggests that while gamification's immediate incentives can enhance short-term focus, its effects on intrinsic motivation and long-term memory consolidation are inferior to consolidation strategies like spaced repetition or retrieval practice [44]. Sestino et al [12] also noted that the long-term efficacy of gamification remains to be fully validated. The primary contribution of this study lies in identifying the foundational role of the metaverse environment in knowledge retention, rather than gamification mechanisms. The similar long-term retention performance observed in both the GM and NGM groups strongly supports the notion that the metaverse itself serves as an effective cognitive carrier. This finding resonates with conclusions from Werner, Iwanaga, and others regarding virtual reality (VR)'s enhanced learning outcomes in medical training through contextual simulation and 3D interaction, emphasizing embodied learning's role in consolidating complex knowledge [4,5]. The heightened sense of presence and interactive contextualized practice provided by immersive environments helps form more robust memory anchors in the brain, thereby effectively promoting active knowledge processing and long-term retrieval [45].

This study confirms that the immersive metaverse serves as an efficient vehicle for foundational gains, providing robust support for learning. Gamification, building upon this foundation, generates immediate gains primarily by amplifying present learning motivation to achieve a cumulative effect. It is particularly suitable for health education scenarios requiring rapid acquisition of core knowledge, such as campus outreach, introductory courses, mobile microlearning, and resource-constrained environments. Designs should prioritize combining immersive environments with gamification to enhance immediate effectiveness. When targeting retention, reinforce this framework with consolidation modules such as spaced repetition, retrieval practice, delayed feedback, periodic recall challenges, and personalized reminder notifications. This transforms instant gains into sustained advantages.

Limitations

It is important to emphasize that this study did not directly measure potential mediating variables such as motivation, self-efficacy, or cognitive load. Consequently, it was not possible to directly test the mediating role of "motivation or immersion." Therefore, future research could incorporate such variables to examine mediating effects, thereby providing more causal empirical evidence regarding the contribution pathways of gamification and metaverse environments to learning outcomes under identical conditions. Second, the follow-up period in this study was limited to 4 weeks, restricting the assessment of longer-term, sustained knowledge retention effects. The existing data cannot determine the Mammaverse platform's resistance to knowledge forgetting over months or even half a year. Additionally, this study was registered retrospectively after the trial commenced. Although the study content and procedures fully align with the ethical approval documentation, this procedural factor has somewhat limited the

transparency of the research report and should be noted when interpreting the results. Concurrently, minor protocol amendments were implemented during the study period with ethics committee approval. None of these amendments involved core elements such as recruitment criteria, intervention content, or sample size. Consequently, they are not anticipated to substantially impact the study's internal validity. In summary, this study robustly confirms that GM group environments significantly enhance short-term knowledge acquisition and sustain high follow-up retention levels. However, to substantially improve long-term knowledge retention rates, future instructional designs must strategically integrate cognitive-based science memory consolidation modules onto existing motivational frameworks.

Innovation and Contribution

In terms of innovation, past explorations in breast health have primarily focused on patient education, lacking research aimed at public awareness dissemination, and employing relatively limited approaches. Existing studies have either examined gamification applications in breast health, such as online breast cancer education games, serious games for self-advocacy and emotional coping, computerized brain training games, step-counting gamified interventions, and breast cancer diagnosis simulation games, to enhance patients' breast cancer knowledge, emotional regulation, and lifestyle [40,46-50]. Alternatively, research focuses on metaverse applications in breast health, including metaverse-based genetic counseling panels for breast cancer, augmented reality-assisted breast surgery education and remote guidance, young women's demand for peer support in the metaverse, digital twin breast cancer models, and privacy-preserving metaverse medical data platforms for breast cancer [17-19,51,52]. However, no studies have yet combined gamification with metaverse platforms for BSE education. This research represents the first empirical study on GM group applications in breast health education, embodying its core innovation at the applied level.

Compared to existing research, GM group studies in recent years have predominantly utilized head-mounted display VR [53-55], or mobile augmented reality as primary platforms [22,56,57], while explorations specifically using 3D desktop metaverses as platforms remain relatively scarce. Existing 3D desktop studies either employ observational designs or single-group pre-post comparisons without control groups [58-60] or contrast 3D desktop environments with traditional lectures or case-based learning [29,61]. Furthermore, explorations of these GM groups in health education have largely been absent in mainland China and predominantly focused on nursing [62,63], predominantly utilizing VR-based approaches. Unlike previous studies, this research establishes both gamified and nongamified versions within the same 3D desktop metaverse platform, conducting an RCT with ordinary women in mainland China. This expands the application scope of gamified metaverses and introduces novel dissemination methods for breast health education and BSE knowledge promotion. Furthermore, prior research has indicated that superficial gamification elements yield no additional benefits in highly immersive VR environments [64]. However, this study identified significant differences within a semi-immersive

desktop environment, validating the effectiveness of GM group platforms for public education on breast health knowledge.

In the real world, against the backdrop of an increasingly younger onset of breast cancer, this study demonstrates that a high-quality immersive 3D desktop metaverse serves as an efficient and scalable platform for breast health education. When resources are limited, and the primary goal is knowledge dissemination, it can serve as a foundational platform for hospital health screening centers, community initiatives, and online programs. By incorporating gamification elements such as tasks, points, and rewards in moderation, engagement levels and short-term learning outcomes can be significantly enhanced. At the same time, the findings remind us that relying solely on one-time gamified experiences is insufficient to ensure long-term knowledge retention and consistent BSE behavior. In real-world practice, task design must be integrated with actual self-examination behaviors. Overall, this study demonstrates that the 3D desktop metaverse possesses a degree of universality and equity across different age groups, educational levels, and risk exposure populations, providing a viable reference for future digital health education initiatives.

Conclusions

As a gamified learning platform based on a metaverse environment, Mammaverse demonstrates clear benefits in enhancing women's breast health knowledge. Under conditions of equivalent content and learning duration, the GM group achieves greater immediate knowledge gains compared to the NGM group. While retention rates were similar between groups during follow-up, the gamified group maintained higher absolute knowledge levels at the four-week mark, indicating its comprehensive advantage in both short-term learning outcomes and sustained retention. To focus on long-term retention, integrating memory consolidation modules into the existing framework is recommended to convert immediate advantages into sustained retention. Mammaverse offers a viable pathway for health education. By introducing a structured perspective that positions "immersion as a foundational gain and gamification as an immediate gain" into public health education, this study provides empirical support for promoting efficient, scalable breast health interventions in resource-constrained and mobile environments.

Acknowledgments

During the preparation of this manuscript, the authors used ChatGPT for grammar and tense checking, language polishing, conversion of references to Vancouver style, and minor refinement of table titles. After using this tool, the authors critically reviewed and edited all content and accept full responsibility for the manuscript.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

Deidentified individual participant data, along with the data dictionary, statistical code, and other relevant materials that support the findings of this study, are not publicly available but can be obtained from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the Mammaverse intervention application interface illustrating key user interactions.

[[PNG File, 1134 KB - games_v14i1e75318_app1.png](#)]

Multimedia Appendix 2

Linear mixed model diagnostics for the primary outcome: standardized residual Q-Q plot (A) and residuals versus fitted values (B).

[[PNG File, 263 KB - games_v14i1e75318_app2.png](#)]

Multimedia Appendix 3

Sensitivity analyses for the primary outcome: type III tests of fixed effects from the adjusted model (dependent variable: T2_3score).

[[PNG File, 191 KB - games_v14i1e75318_app3.png](#)]

Checklist 1

CONSORT checklist.

[[PDF File, 126 KB - games_v14i1e75318_app4.pdf](#)]

Checklist 2

CONSORT-eHEALTH checklist (V 1.6.1).

[\[PDF File, 1127 KB - games_v14i1e75318_app5.pdf\]](#)**References**

1. Yoshikawa H, Wuermli AJ, Britto PR, et al. Effects of the global coronavirus disease-2019 pandemic on early childhood development: short- and long-term risks and mitigating program and policy actions. *J Pediatr* 2020 Aug;223:188-193. [doi: [10.1016/j.jpeds.2020.05.020](https://doi.org/10.1016/j.jpeds.2020.05.020)] [Medline: [32439312](https://pubmed.ncbi.nlm.nih.gov/32439312/)]
2. Kye B, Han N, Kim E, Park Y, Jo S. Educational applications of metaverse: possibilities and limitations. *J Educ Eval Health Prof* 2021;18:32. [doi: [10.3352/jeehp.2021.18.32](https://doi.org/10.3352/jeehp.2021.18.32)] [Medline: [34897242](https://pubmed.ncbi.nlm.nih.gov/34897242/)]
3. Hamid A, Rusdhy FSA, Yusuf PA. Metaverse in medical education. *Med J Indones* 2023;32(2):67-74. [doi: [10.13181/mji.com.236932](https://doi.org/10.13181/mji.com.236932)]
4. Iwanaga J, Muo EC, Tabira Y, et al. Who really needs a metaverse in anatomy education? A review with preliminary survey results. *Clin Anat* 2023 Jan;36(1):77-82. [doi: [10.1002/ca.23949](https://doi.org/10.1002/ca.23949)] [Medline: [36087277](https://pubmed.ncbi.nlm.nih.gov/36087277/)]
5. Werner H, Cossi P, Schor E, et al. EP32.10: The use of metaverse as a tool in teaching uterine malformations in a medical setting: a Brazilian experience. *Ultrasound in Obstet & Gyne* 2023 Oct;62(S1):305-305 [FREE Full text] [doi: [10.1002/uog.27225](https://doi.org/10.1002/uog.27225)]
6. Thomason J. Metahealth-how will the metaverse change health care. *J Metaverse* 2021;1:13-16 [FREE Full text]
7. Kim EJ, Kim JY. The metaverse for healthcare: trends, applications, and future directions of digital therapeutics for urology. *Int Neurourol J* 2023 May;27(Suppl 1):S3-12. [doi: [10.5213/inj.2346108.054](https://doi.org/10.5213/inj.2346108.054)] [Medline: [37280754](https://pubmed.ncbi.nlm.nih.gov/37280754/)]
8. Tan JW, Chong DKS, Ng KB, Car LT, Mogali SR. Rehearsal-based digital serious boardgame versus a game-free e-learning tool for anatomical education: quasi-randomized controlled trial. *Anat Sci Educ* 2023;16(5):830-842. [doi: [10.1002/ase.2286](https://doi.org/10.1002/ase.2286)] [Medline: [37166085](https://pubmed.ncbi.nlm.nih.gov/37166085/)]
9. Haruna H. A comparative analysis of gamified and non-gamified interventions' effectiveness in teaching adolescent students about healthy sexual practices in Tanzania. *Sex Res Soc Policy* 2025 Mar;22(1):435-450. [doi: [10.1007/s13178-024-00961-0](https://doi.org/10.1007/s13178-024-00961-0)]
10. Kirsch J, Spreckelsen C. Caution with competitive gamification in medical education: unexpected results of a randomised cross-over study. *BMC Med Educ* 2023 Apr 19;23(1):259. [doi: [10.1186/s12909-023-04258-5](https://doi.org/10.1186/s12909-023-04258-5)] [Medline: [37072842](https://pubmed.ncbi.nlm.nih.gov/37072842/)]
11. Romero-Alemán MDM. A pilot study on gamified and conventional e-quizzes reinforcing human histology among first-year medical and physiotherapy students. *Anat Sci Educ* 2024 Jun;17(4):844-854. [doi: [10.1002/ase.2406](https://doi.org/10.1002/ase.2406)] [Medline: [38491766](https://pubmed.ncbi.nlm.nih.gov/38491766/)]
12. Sestino A, Bernardo A, Rizzo C, Bresciani S. An explorative analysis of the antecedents and consequents of gamification in the digital therapeutic context. *Eur J Innov Manag* 2025 Apr 22;28(4):1315-1341. [doi: [10.1108/EJIM-08-2023-0633](https://doi.org/10.1108/EJIM-08-2023-0633)]
13. Bansal G, Rajgopal K, Chamola V, Xiong Z, Niyato D. Healthcare in metaverse: a survey on current metaverse applications in healthcare. *IEEE Access* 2022;10:119914-119946. [doi: [10.1109/ACCESS.2022.3219845](https://doi.org/10.1109/ACCESS.2022.3219845)]
14. Hu Y, Peng H, Su G, et al. Effect of a metaverse multimodal rehabilitation intervention on quality of life and fear of recurrence in patients with colorectal cancer survivors: a randomized controlled study protocol. *Digit Health* 2024;10:20552076241295542. [doi: [10.1177/20552076241295542](https://doi.org/10.1177/20552076241295542)] [Medline: [39564355](https://pubmed.ncbi.nlm.nih.gov/39564355/)]
15. Murala DK, Panda SK, Dash SP. MedMetaverse: medical care of chronic disease patients and managing data using artificial intelligence, blockchain, and wearable devices state-of-the-art methodology. *IEEE Access* 2023;11:138954-138985. [doi: [10.1109/ACCESS.2023.3340791](https://doi.org/10.1109/ACCESS.2023.3340791)]
16. Liu N, Li P, Wang J, et al. Factors influencing breast cancer awareness: a cross-sectional study in China. *J Comp Eff Res* 2020 Jul;9(10):679-689. [doi: [10.2217/cer-2020-0037](https://doi.org/10.2217/cer-2020-0037)] [Medline: [32648473](https://pubmed.ncbi.nlm.nih.gov/32648473/)]
17. Gouveia PF, Luna R, Fontes F, et al. Augmented reality in breast surgery education. *Breast Care (Basel)* 2023 Jun;18(3):182-186. [doi: [10.1159/000529587](https://doi.org/10.1159/000529587)] [Medline: [37529369](https://pubmed.ncbi.nlm.nih.gov/37529369/)]
18. Moztarzadeh O, Jamshidi MB, Sargolzaei S, et al. Metaverse and healthcare: machine learning-enabled digital twins of cancer. *Bioengineering (Basel)* 2023 Apr 7;10(4):455. [doi: [10.3390/bioengineering10040455](https://doi.org/10.3390/bioengineering10040455)] [Medline: [37106642](https://pubmed.ncbi.nlm.nih.gov/37106642/)]
19. Yoo B, Kim A, Moon HS, et al. Evaluation of group genetic counseling sessions via a metaverse-based application. *Ann Lab Med* 2024 Jan 1;44(1):82-91. [doi: [10.3343/alm.2024.44.1.82](https://doi.org/10.3343/alm.2024.44.1.82)] [Medline: [37665289](https://pubmed.ncbi.nlm.nih.gov/37665289/)]
20. Mansoory MS, Khazaei MR, Azizi SM, Niromand E. Comparison of the effectiveness of lecture instruction and virtual reality-based serious gaming instruction on the medical students' learning outcome about approach to coma. *BMC Med Educ* 2021 Jun 15;21(1):347. [doi: [10.1186/s12909-021-02771-z](https://doi.org/10.1186/s12909-021-02771-z)] [Medline: [34130676](https://pubmed.ncbi.nlm.nih.gov/34130676/)]
21. Adhikari R, Kydonaki C, Lawrie J, et al. A mixed-methods feasibility study to assess the acceptability and applicability of immersive virtual reality sepsis game as an adjunct to nursing education. *Nurse Educ Today* 2021 Aug;103:104944. [doi: [10.1016/j.nedt.2021.104944](https://doi.org/10.1016/j.nedt.2021.104944)] [Medline: [34015677](https://pubmed.ncbi.nlm.nih.gov/34015677/)]
22. Grübel J, Chatain J, Schmid C, et al. Applying augmented reality to convey medical knowledge on Osteoclasts to users of a serious game: vignette experiment. *JMIR Serious Games* 2025 Jun 16;13(1):e64751. [doi: [10.2196/64751](https://doi.org/10.2196/64751)] [Medline: [40523213](https://pubmed.ncbi.nlm.nih.gov/40523213/)]
23. Li R, Wong SY, Duan B, Kahn SM. A gamified metaverse platform (Mammaverse) for breast health education: a mixed-methods study. *Digit Health* 2025;11:20552076251390466. [doi: [10.1177/20552076251390466](https://doi.org/10.1177/20552076251390466)] [Medline: [41170177](https://pubmed.ncbi.nlm.nih.gov/41170177/)]

24. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. *BMJ* 2025 Apr 14;389:e081123. [doi: [10.1136/bmj-2024-081123](https://doi.org/10.1136/bmj-2024-081123)] [Medline: [40228833](https://pubmed.ncbi.nlm.nih.gov/40228833/)]
25. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. *J Med Internet Res* 2011 Dec 31;13(4):e126. [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
26. De Fino M, Cassano F, Bernardini G, Quagliarini E, Fatiguso F. On the user-based assessments of virtual reality for public safety training in urban open spaces depending on immersion levels. *Saf Sci* 2025 May;185:106803. [doi: [10.1016/j.ssci.2025.106803](https://doi.org/10.1016/j.ssci.2025.106803)]
27. Urbaniak GC, Plous S. Research Randomizer (version 40). URL: <http://www.randomizer.org/> [accessed 2026-01-16]
28. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007 May;39(2):175-191. [doi: [10.3758/bf03193146](https://doi.org/10.3758/bf03193146)] [Medline: [17695343](https://pubmed.ncbi.nlm.nih.gov/17695343/)]
29. Demircan B, Kaya H. The effect of using digital games on learning satisfaction and self-confidence in providing nursing students with subcutaneous injection skills: a randomized controlled trial. *Nurse Educ Pract* 2025 May;85:104341. [doi: [10.1016/j.nepr.2025.104341](https://doi.org/10.1016/j.nepr.2025.104341)]
30. McCance KL, Mooney KH, Smith KR, Field R. Validity and reliability of a breast cancer knowledge test. *Am J Prev Med* 1990;6(2):93-98. [doi: [10.1016/S0749-3797\(18\)31030-4](https://doi.org/10.1016/S0749-3797(18)31030-4)] [Medline: [2363955](https://pubmed.ncbi.nlm.nih.gov/2363955/)]
31. Tlili A, Huang R, Shehata B, et al. Is Metaverse in education a blessing or a curse: a combined content and bibliometric analysis. *Smart Learn Environ* 2022 Dec;9(1):24. [doi: [10.1186/s40561-022-00205-x](https://doi.org/10.1186/s40561-022-00205-x)]
32. Huang L, Deng C, Hoffman J, Hadi Mogavi R, Kim JJ, Hui P. Long-term gamification: a survey. Presented at: International Conference on Human-Computer Interaction (HCI in Games: 6th International Conference, HCI-Games 2024); Jun 29 to Jul 4, 2024. [doi: [10.1007/978-3-031-60692-2_3](https://doi.org/10.1007/978-3-031-60692-2_3)]
33. Rutledge C, Walsh CM, Swinger N, et al. Gamification in action: theoretical and practical considerations for medical educators. *Acad Med* 2018 Jul;93(7):1014-1020. [doi: [10.1097/ACM.0000000000002183](https://doi.org/10.1097/ACM.0000000000002183)] [Medline: [29465450](https://pubmed.ncbi.nlm.nih.gov/29465450/)]
34. Wang T, Fan L, Zheng X, et al. The impact of gamification-induced users' feelings on the continued use of mHealth apps: a structural equation model with the self-determination theory approach. *J Med Internet Res* 2021 Aug 12;23(8):e24546. [doi: [10.2196/24546](https://doi.org/10.2196/24546)] [Medline: [34387550](https://pubmed.ncbi.nlm.nih.gov/34387550/)]
35. Sailer M, Hommer L. The gamification of learning: a meta-analysis. *Educ Psychol Rev* 2020 Mar;32(1):77-112. [doi: [10.1007/s10648-019-09498-w](https://doi.org/10.1007/s10648-019-09498-w)]
36. Shrivastava S, Shrivastava P. Gamification in medical education: an approach to enhance active engagement of students. *J Sci Soc* 2023;50(1):10-12. [doi: [10.4103/jss.jss_113_21](https://doi.org/10.4103/jss.jss_113_21)]
37. Anderson T, Prue G, McDowell G, et al. Co-design and evaluation of a digital serious game to promote public awareness about pancreatic cancer. *BMC Public Health* 2024 Feb 22;24(1):570. [doi: [10.1186/s12889-024-18050-7](https://doi.org/10.1186/s12889-024-18050-7)] [Medline: [38388909](https://pubmed.ncbi.nlm.nih.gov/38388909/)]
38. Marcos R, Gomes A, Santos M, Coelho A. Histopoly: a serious game for teaching histology to 1st year veterinary students. *Anat Sci Educ* 2025 Mar;18(3):229-240. [doi: [10.1002/ase.2545](https://doi.org/10.1002/ase.2545)] [Medline: [39782021](https://pubmed.ncbi.nlm.nih.gov/39782021/)]
39. Wang YF, Hsu YF, Fang KT, Kuo LT. Gamification in medical education: identifying and prioritizing key elements through Delphi method. *Med Educ Online* 2024 Dec 31;29(1):2302231. [doi: [10.1080/10872981.2024.2302231](https://doi.org/10.1080/10872981.2024.2302231)] [Medline: [38194415](https://pubmed.ncbi.nlm.nih.gov/38194415/)]
40. Andrew Tong JW, Hee MQ. The impact of an online educational game on breast cancer awareness among university female students, Malaysia: a pilot study. *BMC Cancer* 2023 Oct 6;23(1):947. [doi: [10.1186/s12885-023-11427-8](https://doi.org/10.1186/s12885-023-11427-8)] [Medline: [37803291](https://pubmed.ncbi.nlm.nih.gov/37803291/)]
41. Georgiou Y, Ioannou A. Embodied learning in a digital world: a systematic review of empirical research in k-12 education. In: *Learning in a Digital World: Perspective on Interactive Technologies for Formal and Informal Education*: Springer; 2019:155-177. [doi: [10.1007/978-981-13-8265-9_8](https://doi.org/10.1007/978-981-13-8265-9_8)]
42. Srdanović P, Skala T, Maričević M. Neurocognitive foundations of memory retention in AR and VR cultural heritage experiences. *Electronics (Basel)* 2025;14(15):2920. [doi: [10.3390/electronics14152920](https://doi.org/10.3390/electronics14152920)]
43. Çelik F, Baturay MH. The effect of metaverse on L2 vocabulary learning, retention, student engagement, presence, and community feeling. *BMC Psychol* 2024 Feb 2;12(1):58. [doi: [10.1186/s40359-024-01549-4](https://doi.org/10.1186/s40359-024-01549-4)] [Medline: [38308350](https://pubmed.ncbi.nlm.nih.gov/38308350/)]
44. Rodrigues L, Pereira FD, Toda AM, et al. Gamification suffers from the novelty effect but benefits from the familiarization effect: findings from a longitudinal study. *Int J Educ Technol High Educ* 2022 Dec;19(1):13. [doi: [10.1186/s41239-021-00314-6](https://doi.org/10.1186/s41239-021-00314-6)]
45. Xi N, Chen J, Gama F, Riar M, Hamari J. The challenges of entering the metaverse: an experiment on the effect of extended reality on workload. *Inf Syst Front* 2023;25(2):659-680. [doi: [10.1007/s10796-022-10244-x](https://doi.org/10.1007/s10796-022-10244-x)] [Medline: [35194390](https://pubmed.ncbi.nlm.nih.gov/35194390/)]
46. Thomas TH, Bender C, Donovan HS, et al. The feasibility, acceptability, and preliminary efficacy of a self-advocacy serious game for women with advanced breast or gynecologic cancer. *Cancer* 2023 Oct 1;129(19):3034-3043. [doi: [10.1002/cncr.34887](https://doi.org/10.1002/cncr.34887)] [Medline: [37243943](https://pubmed.ncbi.nlm.nih.gov/37243943/)]
47. Thomas TH, Bender C, Rosenzweig M, et al. Testing the effects of the STRONG together self-advocacy serious game among women with advanced cancer: protocol for the STRONG randomized clinical trial. *Contemp Clin Trials* 2023 Jan;124:107003. [doi: [10.1016/j.cct.2022.107003](https://doi.org/10.1016/j.cct.2022.107003)] [Medline: [36379436](https://pubmed.ncbi.nlm.nih.gov/36379436/)]

48. Von Ah D, Crouch A, Storey S. Acceptability of computerized cognitive training and global cognitive stimulating-based games delivered remotely: results from a randomized controlled trial to address cancer and cancer-related cognitive impairment in breast cancer survivors. *Cancer Med* 2023 Jun;12(11):12717-12727. [doi: [10.1002/cam4.5904](https://doi.org/10.1002/cam4.5904)] [Medline: [37039279](https://pubmed.ncbi.nlm.nih.gov/37039279/)]
49. Fanaroff AC, Orr JA, Anucha C, et al. A randomized controlled trial of gamification to increase physical activity among Black and Hispanic breast and prostate cancer survivors: rationale and design of the ALLSTAR clinical trial. *Am Heart J* 2025 Feb;280:42-51. [doi: [10.1016/j.ahj.2024.10.021](https://doi.org/10.1016/j.ahj.2024.10.021)] [Medline: [39505121](https://pubmed.ncbi.nlm.nih.gov/39505121/)]
50. Barbosa RG, Rodrigues MAF. Breast cancer diagnosis through serious gaming: clinical reasoning, AI-driven character morphing, and emotional engagement. *Entertain Comput* 2025 Jan;52:100863. [doi: [10.1016/j.entcom.2024.100863](https://doi.org/10.1016/j.entcom.2024.100863)]
51. Ahn J, Lee KE. Experiences of peer support activities and the need for a metaverse-based program in young women with breast cancer: a qualitative study. *Asia Pac J Oncol Nurs* 2023 Jul;10(7):100253. [doi: [10.1016/j.apjon.2023.100253](https://doi.org/10.1016/j.apjon.2023.100253)] [Medline: [37448532](https://pubmed.ncbi.nlm.nih.gov/37448532/)]
52. Letafati M, Otoum S. Digital healthcare in the metaverse: insights into privacy and security. *IEEE Consumer Electron Mag* 2024;13(3):80-89. [doi: [10.1109/MCE.2023.3334494](https://doi.org/10.1109/MCE.2023.3334494)]
53. Huang KT, Ma Z, Yao L. Media-induced and psychological factors that Foster empathy through virtual reality in nursing education: 2x2 between-subjects experimental study. *JMIR Med Educ* 2025 Mar 31;11:e59083. [doi: [10.2196/59083](https://doi.org/10.2196/59083)] [Medline: [40163855](https://pubmed.ncbi.nlm.nih.gov/40163855/)]
54. Kadri M, Boubakri FE, Kaghat FZ, Azough A, Zidani KA. IVAL: Immersive Virtual Anatomy Laboratory for enhancing medical education based on virtual reality and serious games, design, implementation, and evaluation. *Entertain Comput* 2024 Mar;49:100624. [doi: [10.1016/j.entcom.2023.100624](https://doi.org/10.1016/j.entcom.2023.100624)]
55. Chang YY, Chao LF, Chang W, et al. Impact of an immersive virtual reality simulator education program on nursing students' intravenous injection administration: a mixed methods study. *Nurse Educ Today* 2024 Jan;132:106002. [doi: [10.1016/j.nedt.2023.106002](https://doi.org/10.1016/j.nedt.2023.106002)] [Medline: [37922767](https://pubmed.ncbi.nlm.nih.gov/37922767/)]
56. Othman SY, Ghallab E, Eltaybani S, Mohamed AM. Effect of using gamification and augmented reality in mechanical ventilation unit of critical care nursing on nurse students' knowledge, motivation, and self-efficacy: a randomized controlled trial. *Nurse Educ Today* 2024 Nov;142:106329. [doi: [10.1016/j.nedt.2024.106329](https://doi.org/10.1016/j.nedt.2024.106329)] [Medline: [39116661](https://pubmed.ncbi.nlm.nih.gov/39116661/)]
57. Willinger L, Schweizer F, Böhm B, et al. Evaluation of the gamified application KIJANI to promote physical activity in children and adolescents: a multimethod study. *Digit Health* 2024;10:20552076241271861. [doi: [10.1177/20552076241271861](https://doi.org/10.1177/20552076241271861)] [Medline: [39161345](https://pubmed.ncbi.nlm.nih.gov/39161345/)]
58. Rudolphi-Solero T, Lorenzo-Álvarez R, Domínguez-Pinos D, Ruiz-Gómez MJ, Sendra-Portero F. An interuniversity competition for medical students to learn radiology in the second life metaverse. *J Am Coll Radiol* 2024 May;21(5):812-821. [doi: [10.1016/j.jacr.2023.09.012](https://doi.org/10.1016/j.jacr.2023.09.012)] [Medline: [37922968](https://pubmed.ncbi.nlm.nih.gov/37922968/)]
59. Joy S, Rumsey K, Ford M. The Treasures of Renal Isle: evaluating a gamified module while enhancing student learning. *SAGE Open Nurs* 2024;10:23779608241290683. [doi: [10.1177/23779608241290683](https://doi.org/10.1177/23779608241290683)] [Medline: [39429545](https://pubmed.ncbi.nlm.nih.gov/39429545/)]
60. De Luca V, Qbilat M, Cuomo A, et al. Virtual reality solution to promote adapted physical activity in older adults: outcomes from VR2Care project exploratory study. *Front Public Health* 2025;13:1584406. [doi: [10.3389/fpubh.2025.1584406](https://doi.org/10.3389/fpubh.2025.1584406)] [Medline: [40433492](https://pubmed.ncbi.nlm.nih.gov/40433492/)]
61. Kim HJ, Oh J, Lee S. Effect of virtual game-based integrated clinical practice simulation program on undergraduate nursing students' attitude toward learning. *Comput Inform Nurs* 2024 Mar 1;42(3):218-225. [doi: [10.1097/CIN.0000000000001105](https://doi.org/10.1097/CIN.0000000000001105)] [Medline: [38261420](https://pubmed.ncbi.nlm.nih.gov/38261420/)]
62. Rickenbacher-Frey S, Adam S, Exadaktylos AK, Müller M, Sauter TC, Birrenbach T. Development and evaluation of a virtual reality training for emergency treatment of shortness of breath based on frameworks for serious games. *GMS J Med Educ* 2023;40(2):Doc16. [doi: [10.3205/zma001598](https://doi.org/10.3205/zma001598)] [Medline: [37361243](https://pubmed.ncbi.nlm.nih.gov/37361243/)]
63. Xiao X, Xiao Y, Dumlao MF. Evaluation of gamification on surgical nursing course using immersive virtual reality: a comparative study. Presented at: ISAIMS '23: Proceedings of the 2023 4th International Symposium on Artificial Intelligence for Medicine Science; Oct 20-22, 2023. [doi: [10.1145/3644116.3644217](https://doi.org/10.1145/3644116.3644217)]
64. Larsen JD, Jensen RO, Pietersen PI, et al. Education in focused lung ultrasound using gamified immersive virtual reality: a randomized controlled study. *Ultrasound Med Biol* 2023 Mar;49(3):841-852. [doi: [10.1016/j.ultrasmedbio.2022.11.011](https://doi.org/10.1016/j.ultrasmedbio.2022.11.011)] [Medline: [36535832](https://pubmed.ncbi.nlm.nih.gov/36535832/)]

Abbreviations

BSE: breast self-examination

CONSORT: Consolidated Standards of Reporting Trials

CONSORT-eHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

EMM: estimated marginal mean

GM: gamified metaverse

LMM: linear mixed model

NGM: nongamified metaverse

PC: personal computer

RCT: randomized controlled trial

Edited by S Brini; submitted 01.Apr.2025; peer-reviewed by A Edalatpour, C Zou, K Alomari; accepted 06.Jan.2026; published 13.Feb.2026.

Please cite as:

Li R, Kahn SM, Duan B, Wong SY

Gamified Versus Nongamified Metaverse Learning for Breast Health Knowledge in Women: Randomized Controlled Trial

JMIR Serious Games 2026;14:e75318

URL: <https://games.jmir.org/2026/1/e75318>

doi: [10.2196/75318](https://doi.org/10.2196/75318)

© Rui Li, Sabzali Musa Kahn, Bingyu Duan, Seng Yue Wong. Originally published in JMIR Serious Games (<https://games.jmir.org>), 13.Feb.2026. 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.

Balance Assessment Using Gamified Digital Technology in Community-Dwelling Older Adults: Mixed Methods Validation Study and Randomized Controlled Trial

Jianan Zhao¹, PhD; Yaqin Xia¹, PhD; Yahui Zhang², MD; Jihong Yu¹, PhD; Chen Chu², PhD

¹College of Fashion and Design, Donghua University, Yanan Road #1448, Shanghai, China

²School of Design, Shanghai Jiao Tong University, Dongchuan Road 800#, Shanghai, China

Corresponding Author:

Chen Chu, PhD

School of Design, Shanghai Jiao Tong University, Dongchuan Road 800#, Shanghai, China

Abstract

Background: Falls cause injury and mortality among older adults, necessitating reliable, scalable, engaging balance assessment tools to use in community settings. Traditional clinician-administered assessments like the Brief Balance Evaluation Systems Test (Brief-BESTest) are limited by subjectivity and accessibility constraints. Computer vision-based digitalization combined with gamification may address these limitations; yet, validation evidence remains limited.

Objective: We (1) digitalized and validated a computer vision-based Brief-BESTest against clinician scoring and (2) investigated whether a gamified interface improves older adults' user experience during balance assessment, without compromising assessment performance.

Methods: This mixed methods study comprised (1) a concurrent validity substudy in a convenience subsample (n=10) and (2) a parallel-group randomized controlled trial (RCT; n=30) with 1:1 allocation. Participants were community-dwelling older adults aged ≥ 60 years recruited from Hongqi Community, Shanghai, through community announcements and health care worker referrals. Phase 1 (n=10; mean age 64.9, SD 2.76 years) evaluated concurrent validity of a computer vision-based digitalized Brief-BESTest using OpenPose skeletal tracking (Carnegie Mellon University Perceptual Computing Lab; Logitech Brio 4K webcam, 27-inch touchscreen) against the clinician-administered version. Phase 2 (n=30; mean age 66.7, SD 3.93 years) used a parallel-group RCT with 1:1 coin-flip allocation. Primary outcome measures include perceived exertion (Borg Rating of Perceived Exertion scale 6 - 20), intrinsic motivation (Intrinsic Motivation Inventory 7-point Likert, including interest and enjoyment, perceived competence, and pressure and tension subscales), and intention to continue use (7-point Likert scale). Semistructured interviews (mean 4.8 minutes) assessed engagement factors. Data collection occurred in a controlled indoor setting with safety railings.

Results: Phase 1 demonstrated excellent intrasession reliability (intraclass correlation coefficient=0.89 - 0.92) and strong concurrent validity (Spearman $\rho=0.91$; 95% CI 0.68 - 0.98; $P<.01$), with no significant mean difference (MD; paired t test: MD 0.23; $P=.77$; $d=-0.07$). In phase 2, Gamified Digital Balance Assessment (GDBA) users reported significantly lower perceived exertion (Mann-Whitney U : MD -2.67; 95% CI -4.60 to -0.74; $P=.01$; $d=-1.08$), higher enjoyment (MD 1.53; $P=.009$; $d=1.17$), higher perceived competence (MD 1.14; $P=.02$; $d=0.89$), and higher intention to continue use (MD 1.66; $P=.001$; $d=1.25$). Pressure and tension ($P=.09$; $d=0.63$) showed no significant difference. Thematic analysis (Cohen $\kappa=0.68$) identified 2 themes: motivational rewards (80% cited real-time feedback) and perceived usability (87% emphasized avatar demonstrations).

Conclusions: This study validated a computer vision-based digital Brief-BESTest and experimentally tested a gamified interface for balance assessment in community-dwelling older adults. Unlike prior work focused largely on single-task digital tests or nongamified interfaces, the GDBA integrates comprehensive, clinically grounded balance assessment with evidence-based gamification tailored to older users. These findings advance digital geriatric assessment by demonstrating that gamified designs can enhance motivation, perceived competence, and tolerability of testing without sacrificing measurement quality. If replicated in longitudinal and real-world settings, such systems could provide scalable, low-cost tools for routine fall-risk screening, self-monitoring, and targeted preventive interventions in community and primary-care environments.

Trial Registration: ClinicalTrials.gov NCT06958653; <https://clinicaltrials.gov/study/NCT06958653>

(*JMIR Serious Games* 2026;14:e80747) doi:[10.2196/80747](https://doi.org/10.2196/80747)

KEYWORDS

digital health; gamification; older adults; balance testing; user engagement

Introduction

Falls are a leading cause of injury, disability, and mortality among older adults, with approximately one-third of community-dwelling individuals aged 65 years or older experiencing a fall each year [1]. A recent comprehensive systematic review and meta-analysis estimated the global prevalence of falls in older adults at 26.5%, underscoring that roughly one in four older adults falls annually [2]. In China, the burden is amplified by a rapidly aging population, which is projected to exceed 400 million people aged 60 years and older by 2040 [3]. This demographic trend increases both absolute numbers of fall events and the potential demand for community-based prevention and monitoring services [4]. Effective fall prevention strategies are urgently needed to address the rising health care demands of older populations, particularly in community settings where access to regular clinical assessment is limited [5,6].

Routine balance assessment is a cornerstone of fall prevention, enabling early identification of postural instability and guiding timely intervention. However, conventional balance assessments—such as the Brief Balance Evaluation Systems Test (Brief-BESTest)—are typically conducted in clinical environments by trained professionals, limiting accessibility for independently living older adults, especially in underserved or rural areas [7]. Moreover, clinic-based assessments are time- and resource-intensive, and many older adults receive only episodic screening; this gap motivates interest in remote and automated assessments that preserve clinical validity while increasing reach [8]. Advances in computer vision technologies present an opportunity to overcome these limitations by facilitating contactless, automated assessments that can be delivered outside clinical settings [9].

Existing digital balance testing systems frequently rely on assessments such as the Timed Up and Go (TUG) Test [10] or the Sit-to-Stand Test [11], which primarily evaluate isolated or simple movements. While these tools are useful, they may lack the multidimensional scope needed for a comprehensive evaluation of balance [12]. In contrast, the Brief-BESTest is a validated and concise clinical tool that assesses 6 key domains of postural control—biomechanical constraints, stability limits, anticipatory adjustments, reactive postural responses, sensory orientation, and stability in gait—making it well-suited for a more holistic assessment of balance in older adults [13,14]. Validation studies support the construct validity of the BESTest family (BESTest, mini-BESTest, and Brief-BESTest) in community-dwelling older adults and demonstrate that these tools discriminate fallers from nonfallers and correlate with established functional measures, which argues for translating the Brief-BESTest's multidimensional structure into a digital format rather than relying solely on single-metric tests [15]. Nevertheless, there are relatively few peer-reviewed implementations that fully translate a multidomain instrument like the Brief-BESTest into a markerless computer-vision pipeline suitable for unsupervised community use; existing digital systems more commonly implement TUG, sit-to-stand, or gait-only assessments [16].

In addition to technical feasibility, sustaining engagement among older users is essential. Gamification—the integration of game design elements such as feedback, rewards, and interactive visuals—has been shown to enhance motivation and adherence in digital health applications [17,18]. A systematic review of gamification for older adults found overall positive effects on engagement, enjoyment, and adherence when gameful elements are appropriately tailored; however, the review also highlights heterogeneity in design choices and a need for stronger evidence linking gamification to objective clinical outcomes in older populations [19]. Recent evidence suggests that well-designed gamified interventions can reduce perceived effort and increase autonomous motivation in older adults performing physical assessments [20]. For older adults, interactive systems may reduce anxiety, improve perceived competence, and encourage repeated use of assessment tools [21,22]. Specifically, real-time visual feedback and avatar-based demonstrations have been identified as key design elements that support self-efficacy and reduce cognitive load during technology-mediated assessments [23]. Yet, little is known about the impact of gamifying comprehensive balance assessments like the Brief-BESTest on user experience and continued use intentions [24].

This study addresses two critical gaps: (1) the absence of a valid and reliable, computer vision-based version of the Brief-BESTest suitable for community deployment, and (2) limited understanding of the role of gamification in enhancing older adults' engagement with digital balance assessments. We present the development and evaluation of the Gamified Digital Balance Assessment (GDBA), which integrates automated Brief-BESTest scoring with gamified interface design. Through a 2-phase mixed methods study, we tested the following objectives and hypotheses:

1. To validate the digitalized Brief-BESTest against the clinician-administered version, hypothesizing strong concurrent validity and reliability.
2. To evaluate the impact of gamification on user experience, hypothesizing that the GDBA would result in lower perceived exertion and higher intrinsic motivation compared to the nongamified version.

Methods

Patient and Public Involvement

No patients or members of the public were formally involved in setting the research questions, outcome measures, intervention design, or recruitment procedures for this trial. A total of 30 community-dwelling older adults participated in the randomized controlled trial. A convenience subsample of 10 participants contributed to the concurrent validity substudy. Participants served only as study participants and were not involved in study governance or dissemination planning.

Trial Design

This 2-phase study was conducted at the Hongqi Community Center in Shanghai in April-May 2025. Participant enrollment for the dataset reported here took place between April 20 and May 20, 2025. It used a sequential 2-phase mixed methods design. Phase 1 used a within-subjects concurrent validity design.

with repeated measures. Phase 2 used a parallel-group randomized controlled trial design with 1:1 allocation, supplemented by qualitative semistructured interviews that carefully follow the CONSORT (Consolidated Standards of Reporting Trials) guideline (Checklist 1) [25]. The Phase 2 randomized controlled trial was retrospectively registered at ClinicalTrials.gov (NCT06958653) after enrollment began (registration date: July 17, 2025). The study team initially categorized the Phase 2 component as a low-risk formative evaluation focused on user experience of a digital assessment interface and did not recognize that any randomized allocation of participants requires prospective trial registration under International Committee of Medical Journal Editors (ICMJE)/JMIR policy. Once this requirement was identified during manuscript preparation, we registered the trial and are updating the registry record to transparently reflect the outcomes and their timing as collected and reported in this manuscript. This research collected written informed consent from all participants and was reported following the Journal Article Reporting Standards (JARS) for mixed methods research [26].

Participant Characteristics, Sampling Procedures, and Selection Criteria

Community-dwelling older adults were recruited offline from the Hongqi Community in Shanghai through community announcements, poster displays at the community center, and referrals from community health care workers. Interested individuals were screened for eligibility via telephone by a trained research assistant using a standardized eligibility checklist. Eligible participants were then invited for an in-person informed consent session prior to enrollment. Inclusion criteria for both phases were (1) age ≥ 60 years, (2) ability to stand unassisted for 30 seconds and walk 10 m independently, and (3) no acute illness at the time of testing. Exclusion criteria included (1) diagnosed neurological conditions affecting balance, (2) uncorrected visual or auditory impairments, (3) recent fall requiring medical attention, and (4) use of medications known to affect postural stability.

Sample Size, Power, and Precision

For Phase 1, a sample of 10 participants was selected for reliability studies. For Phase 2, sample size was determined a priori using G*Power 3.1 software (Heinrich Heine University Düsseldorf). Assuming a 2-tailed independent t test, $\alpha=.05$, power $(1-\beta)=.80$, and a medium-to-large effect size (Cohen $d=0.55$) based on prior gamification studies in older adults [21], the required sample size was calculated as $n=28$. We recruited 30 participants (15 per group) to account for potential dropouts, though none occurred.

Systematic assessment of harms was not a primary objective of this study given the low-risk nature of balance testing in a controlled environment with safety measures. However, we monitored for adverse events during all testing sessions. Potential harms were defined as (1) falls or near-falls during balance tasks and (2) musculoskeletal discomfort or pain exacerbation. All testing was conducted with safety railings present and a trained researcher monitoring participants. No adverse events occurred during the study period. Participants were instructed to stop immediately if they experienced any

discomfort, and emergency protocols were established in consultation with local health care providers.

Product Development

Digitalized Brief-BESTest Design

The digitalized Brief-BESTest was designed to digitize and automate the Brief-BESTest. While the traditional clinician-administered Brief-BESTest relies on subjective scoring, the digitalized Brief-BESTest enables self-guided assessments with automated, objective scoring, improving accessibility in community and home settings.

The system uses OpenPose, an open-source pose estimation framework validated for human movement analysis, to capture skeletal data via a standard 2D RGB camera. OpenPose was selected based on evidence demonstrating its accuracy in tracking older adult movements in clinical contexts [27] and its suitability for noninvasive, contactless assessment without wearable sensors. The system tracks 17 anatomical landmarks (eg, nose, neck, shoulders, elbows, wrists, hips, knees, and ankles) at 30 frames per second.

Ten joint angles relevant to static and dynamic postural tasks (eg, hip flexion, knee flexion, ankle dorsiflexion, and trunk inclination) are computed from the landmark coordinates. The torso vector is defined from the neck landmark to the midpoint between the left and right hip landmarks, serving as a reference axis for postural alignment calculations.

To convert pixel-based coordinates into metric units (centimeters), the system uses participants' self-reported height adjusted by anthropometric correction factors derived from ISO 7250 - 1 international standards. These correction factors (10.77 cm for men, 10.06 cm for women) represent the average vertical distance from the top of the head to the C7 vertebra (neck landmark detected by OpenPose), as established in large-scale anthropometric databases. These values have been validated for use in Chinese adult populations [28]. The corrected height is then used to normalize skeletal dimensions, enabling calculation of joint angles and postural parameters in standardized metric units.

The digitalized Brief-BESTest replicates all 8 tasks from the validated Brief-BESTest: (1) lift leg to the side, (2) lift 2 arms, (3) stand on one leg-left, (4) stand on one leg-right, (5) take a step to the left, (6) take a step to the right, (7) stance with eyes closed, and (8) up and go. These 8 tasks cover all 6 domains of the original Brief-BESTest, including biomechanical constraints, stability limits/verticality, anticipatory postural adjustments, postural responses, sensory orientation, and stability in gait.

Key kinematic parameters include (1) torso angle deviation from vertical (quantifying postural sway), (2) joint angles (hip, knee, ankle flexion/extension, and abduction/adduction), (3) center of mass displacement (estimated via pelvis midpoint), (4) temporal metrics (stance duration and time to stabilization), and (5) spatial metrics (reach distance, step length, and base of support width).

Representative scoring examples include the following: for single-leg stance (Items 3 - 4), participants scoring 3 points maintain stance ≥ 20 seconds with postural sway $< 3^\circ$ and hands

on hips, while those scoring 0 points cannot maintain stance ≥ 2 seconds. For functional reach forward (Item 2), score 3 requires ≥ 32 cm forward reach without heel lift or trunk rotation, while score 0 indicates < 3 cm displacement or loss of balance. For the TUG test (Item 8), score 3 requires completion in ≤ 11 seconds with trunk sway $< 5^\circ$ during gait, while score 0 indicates > 11 seconds with significant imbalance (trunk sway $> 8^\circ$, stumbles, or requiring assistance).

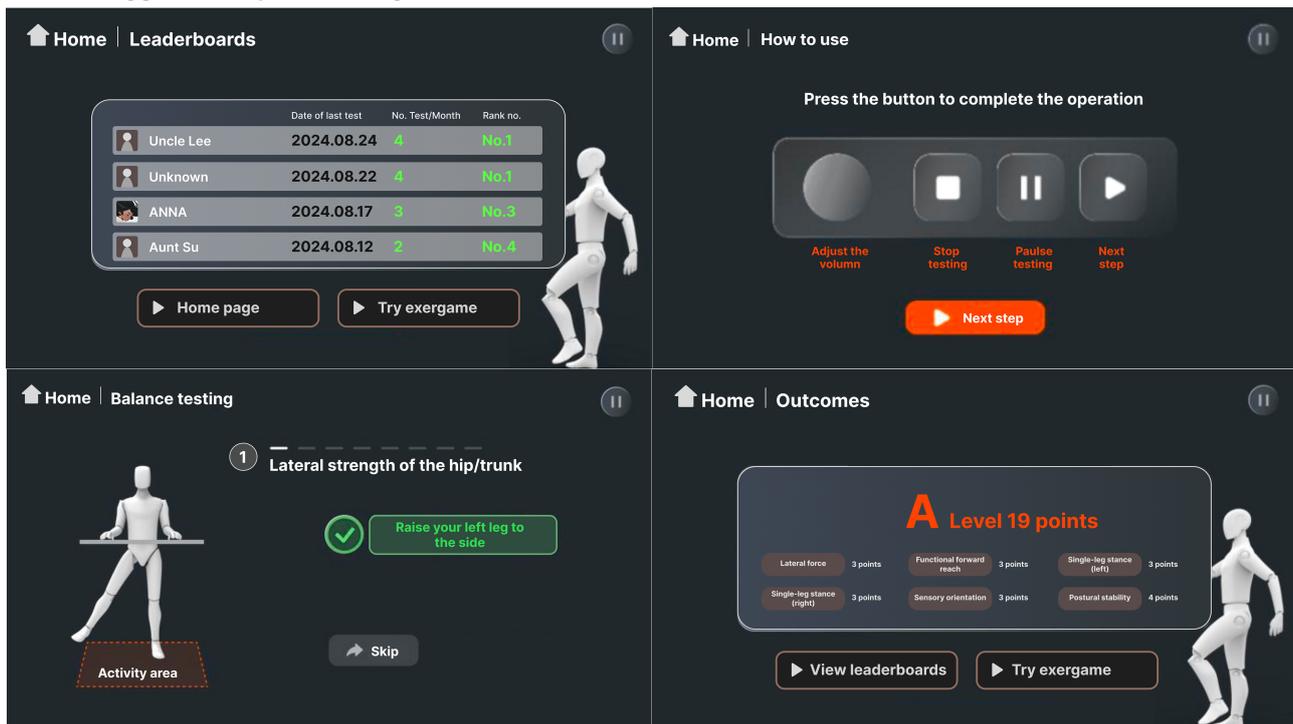
The total balance score (0 - 24) is calculated as the sum across all eight items, with established cutoffs for fall risk interpretation: 0 - 12 (high risk), 13 - 18 (moderate risk), and 19 - 24 (low risk).

GDBA Design

The GDBA builds upon the digitalized Brief-BESTest by incorporating evidence-based gamification elements designed to enhance motivation and engagement among older adults. The gamification design was guided by self-determination theory [29], which posits that autonomy, competence, and relatedness are key drivers of intrinsic motivation, and by recent systematic reviews on gamification for older adult health interventions [23,30]. The examples of the interface are presented in Figure 1. Several core game mechanics were implemented.

- **Points and scoring:** each balance task awards 0 - 3 points based on performance, with real-time point accumulation displayed prominently. Achieving higher scores unlocks encouraging messages (eg, "Excellent balance!").
- **Animated avatars:** a 3D avatar demonstrates each task before the participant performs it, showing ideal posture and movement. Avatar design was built based on the shapely human figure without specific age references to enhance relatedness.
- **Progress visualization:** a dynamic progress bar shows overall completion status, and a colored time count timer helps participants pace themselves during timed tasks.
- **Performance summary report:** upon completion, participants receive a comprehensive report including total score with fall risk category, domain-specific subscores, task-level feedback, and personalized training recommendations.
- **Optional leaderboard:** an anonymized leaderboard displays recent user rankings (initials and scores only) with emphasis on personal progress over social comparison. Participants can opt out during initial setup.
- **Multimodal accessibility design:** the interface features high-contrast colors (black background, orange, and green highlights), large fonts (≥ 24 point), voice narration in Mandarin (120 words per min), and simplified navigation with large physical buttons ($\geq 100 \times 100$ pixels) to accommodate age-related sensory and motor changes.
- **Difficulty levels and progression (under development):** although not implemented in this study, the system is designed to support adaptive difficulty training sessions (eg, unlocking advanced balance exercises after achieving baseline competence). These features are planned for future versions to support long-term engagement.
- **Hardware setup:** the assessment environment comprises a 1 m \times 3 m zone with safety railings (height: 90 cm) on all sides and a centrally placed 1 m \times 1 m ethylene-vinyl acetate (EVA) foam pad (10 cm thick, 35D density) for sensory orientation tasks. Hardware includes a Logitech Brio 4K webcam mounted 2.5 m from the assessment center; a 27-inch LCD touchscreen display positioned at eye level (adjustable 1.2 - 1.7 m); an embedded computer (Intel i7, 16 GB RAM, NVIDIA GTX 1650 GPU) running custom Python software (Python Software Foundation; OpenPose for tracking, NumPy/SciPy for kinematics, Figma for interface); and a detachable ergonomic control console with 3 large tactile buttons designed per Chinese anthropometric standards (GB 10000 - 88). All components are optimized for older adult users aged > 60 years.

Figure 1. User interface examples of the Gamified Digital Balance Assessment (GDBA) system showing the task instruction screen with animated avatar demonstration, real-time performance feedback during balance task execution, and postassessment summary with score visualization and leaderboard. Hongqi Community Center, Shanghai, China, 2025.



Study Procedure and Data Diagnostics

The study was conducted in 2 sequential phases to evaluate the reliability and user experience of a GDBA tailored for community-dwelling older adults.

Phase 1: Reliability of Digitalized Brief-BESTest Assessment

In the first phase, participants performed a single balance assessment session, during which both the clinician-administered Brief-BESTest and the digitalized Brief-BESTest were scored concurrently. This approach enabled direct comparison between clinical and automated assessments under identical task conditions.

Testing was conducted in a controlled indoor setting featuring a 1 m×1 m, 10-cm-thick EVA foam mat (35D density) and safety handrails on 3 sides. Prior to the assessment, participants completed a baseline questionnaire collecting demographic data (age and sex), anthropometric measurements (height and weight), and fall history (past 12 months).

During the assessment, a certified physical therapist delivered standardized verbal instructions and rated each task using the validated Brief-BESTest rubric (maximum score=24). Simultaneously, the digitalized Brief-BESTest system recorded participants' movements using a monocular 4K camera and calculated scores via an algorithm that mirrors the original scoring criteria. The torso and joint movements were analyzed in real time, and balance scores were automatically computed.

To evaluate interrater reliability, a second trained clinician independently rated 20% of the sample. Both raters had more than 5 years of clinical experience and completed standardized training on Brief-BESTest administration prior to the study.

This concurrent scoring design ensured consistent task execution while enabling evaluation of intermethod reliability of the automated system's scoring against expert clinician judgment.

Potential Assessor Bias Mitigation: while the primary clinician-administered assessments were conducted by a single certified physiotherapist to ensure consistency in instruction delivery, we acknowledge this may introduce assessor-specific bias (eg, subjective interpretation of borderline performance). To mitigate this limitation, we (1) used a second independent rater for 20% of cases to establish interrater reliability (intraclass correlation coefficient [ICC]=0.94; 95% CI 0.85 - 0.98), (2) ensured the primary rater followed the standardized Brief-BESTest manual verbatim, and (3) video-recorded all sessions to allow for post hoc verification. Future studies should use multiple assessors across sessions to further minimize individual rater effects.

Phase 2: Impact of GDBA on User Experience

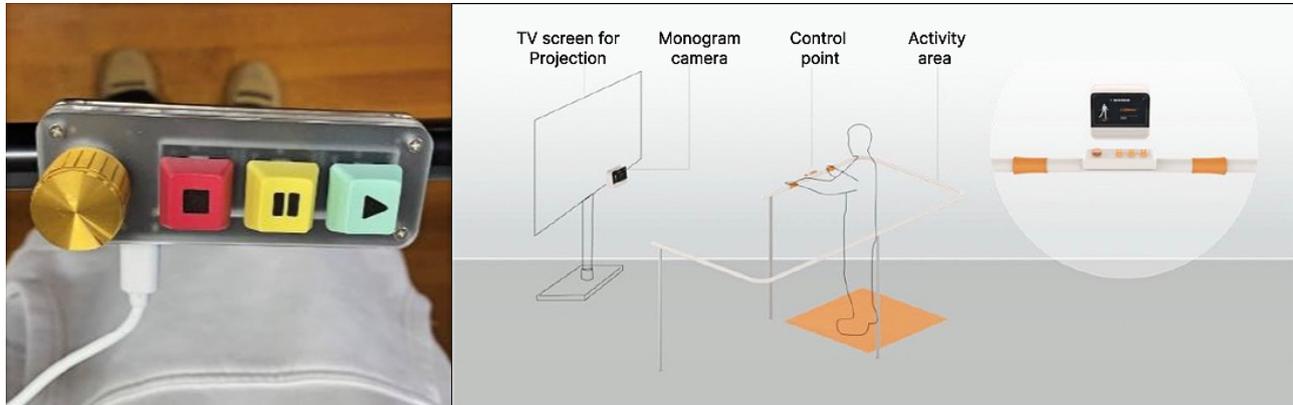
The second phase involved a parallel group randomized controlled trial (RCT) to assess the impact of gamification on user experience.

After informed consent, participants were allocated using a simple coin-flip method (heads=control, tails=experimental) to either the control group (uses digitalized Brief-BESTest) or the experimental group (uses GDBA) through a simple coin-randomization method by a blinded researcher. Questionnaires that assessed their demographic information, fall history, and balance confidence were used. Testing was conducted in a 1 m×3 m evaluation zone equipped with front, side, and rear safety railings and a centrally placed EVA foam pad (identical to Phase 1). The DBTS system included a display screen, a Logitech Brio 4K webcam (30 fps) for motion tracking,

and a built-in speaker for voice prompts. A detachable, ergonomically designed user console—compliant with Chinese

anthropometric standards—was mounted on the front railing for interface navigation (see Figure 2).

Figure 2. Experimental setup for the evaluation of the Gamified Digital Balance Assessment (GDBA) in community-dwelling older adults (N=30, aged >60 years), including a 1 m×3 m assessment zone with safety railings and an ethylene-vinyl acetate (EVA) foam mat, a 4K camera and display system for motion tracking and visual feedback, and an ergonomically designed user control console. Hongqi Community Center, Shanghai, China, 2025.



In the control group, participants performed balance tasks following prerecorded verbal instructions from a certified physical therapist. In the experimental group, tasks were presented via the GDBA interface, which included animated avatars, voice guidance, progress indicators, and real-time performance feedback. Each participant completed one practice trial per task to minimize learning effects, followed by the formal assessment. A 2-minute seated rest period was provided between tasks to reduce fatigue.

Immediately following the assessment, participants completed self-report measures on perceived exertion, intrinsic motivation, and intention for continued use. They then participated in a brief semistructured interview exploring their perceptions of system usability and engagement. All participants (n=30) completed individual semistructured interviews (mean duration=4.8 min, range=3 - 7 min) immediately post assessment. The interviewer was aware of group assignment but blinded to quantitative scores. All interviews were audio-recorded and transcribed verbatim in Mandarin within 48 hours. Transcripts were verified for accuracy by the primary interviewer. Participants received a nominal compensation (US \$10 equivalent) upon study completion.

Each participant attended a single testing session lasting approximately 25 - 35 minutes with no interim analyses, which included informed consent, a baseline questionnaire, a balance assessment, postassessment questionnaires, and qualitative interviews. The participants can quit the experiment whenever they want. This cross-sectional design enabled evaluation of immediate user experience but did not assess longitudinal adherence or learning effects over multiple sessions.

Outcome Measures and Covariates

Baseline Variable

Balance ability was assessed using the Brief-BESTest, a validated clinical tool comprising 8 tasks across 6 domains of postural control, including biomechanical constraints, stability limits/verticality, anticipatory postural adjustments, postural responses, sensory orientation, and stability in gait [13]. Each

task is scored on a 0 - 3 scale (0=inaccurate performance, 3=normal performance), yielding a total score of 0 - 24, with higher scores indicating better balance.

Balance confidence was assessed as a baseline variable using the Activities-specific Balance Confidence (ABC) scale (0% - 100%); for reporting consistency, ABC scores were normalized to a 0 - 1 scale by dividing by 100 and were collected immediately after the assessment session. Higher scores indicate greater balance confidence.

User Experience

Perceived physical exertion was measured using the Borg Rating of Perceived Exertion (RPE) Scale, ranging from 6 (“no exertion”) to 20 (“maximal exertion”) [31]. Participants verbally reported their RPE immediately after completing all balance tasks to reflect overall physical demand and fatigue during the assessment. This measure provided insight into the tolerability and physical burden of the assessment procedures.

Motivational engagement was evaluated using the Intrinsic Motivation Inventory (IMI), a validated tool using a 7-point Likert scale (1=“not at all true,” 7=“very true”) [32]. Three subscales were analyzed, including interest/enjoyment (assesses task engagement and inherent enjoyment of the activity), perceived competence (measures self-perceived ability and confidence in performing the tasks), and pressure/tension (evaluates task-related stress and anxiety). Subscale scores were calculated as the mean of item responses, with higher scores indicating greater enjoyment, competence, or pressure, respectively. The IMI has demonstrated good internal consistency and construct validity in older adult populations [33].

Intention to continue use was assessed through both quantitative and qualitative methods. Quantitatively, participants rated their likelihood of using the system again on a single-item 7-point Likert scale (1=“very unlikely,” 7=“very likely”) immediately following the assessment. Qualitatively, semistructured interviews explored factors influencing future use intentions. Interview questions included, “Would you consider using this

system regularly for balance checking? Why or why not?” and “What features would encourage you to use this system more often?” Interviews lasted 3 - 5 minutes, were audio-recorded with permission, transcribed verbatim, and analyzed using the 6-phase thematic analysis framework.

Training and Reliability of Data Collectors

Data collection was conducted by trained research personnel who completed a standardized 4-hour training session on Brief-BESTest administration, equipment operation, and participant interaction protocols. Training included video-based instruction, practice sessions with pilot participants, and competency assessment by the principal investigator. Interrater reliability was established by having 2 certified physiotherapists (both with >3 years' experience in geriatric assessment) independently score 20% (2/10) of Phase 1 participants, achieving excellent agreement. The digitalized Brief-BESTest system underwent technical validation prior to data collection, including calibration testing with known reference objects to verify measurement accuracy. For qualitative interviews, the primary interviewer (JZ) received training in semistructured interview techniques and motivational interviewing principles. Interview fidelity was maintained through use of a standardized interview guide and audio recording of all sessions for quality review.

Psychometrics

Reliability in this sample: internal consistency was assessed for all multi-item scales in this sample (n=30, Phase 2). The IMI subscales demonstrated good internal consistency, including interest and enjoyment (Cronbach $\alpha=.87$; 95% CI 0.76 - 0.93), perceived competence ($\alpha=.84$; 95% CI 0.71 - 0.91), and pressure and tension ($\alpha=.79$; 95% CI 0.63 - 0.88). The Brief-BESTest total score showed excellent internal consistency in this current sample ($\alpha=.91$; 95% CI 0.83 - 0.95). Test-retest reliability was not assessed in this study due to the single-session design. However, the digitalized version demonstrated excellent intrasession reliability (ICC=0.89 - 0.92) across multiple automated scoring iterations of the same video recordings.

Convergent validity: convergent validity of the digitalized Brief-BESTest was supported by strong correlation with the clinician-administered version (Spearman $\rho=0.91$; $P<.01$), suggesting both methods measure the same construct.

Statistical Methods

All quantitative analyses were conducted using SPSS (IBM Corp). To assess the validity of the digitalized Brief-BESTest, balance scores from the clinician-administered and algorithm-generated assessments were compared using paired *t* tests or Wilcoxon signed-rank tests, depending on normality. Effect sizes were calculated using Cohen *d*. Interrater reliability for the clinician-administered Brief-BESTest was assessed on 20% of cases scored by 2 clinicians using the intraclass correlation coefficient (ICC, 2-way mixed model, absolute agreement). ICC values >0.75 were considered excellent, 0.40 - 0.75 moderate to good, and <0.40 poor. Convergent validity was examined via Spearman correlation (ρ).

In the second phase, nonparametric tests (Mann–Whitney *U*) were used to compare user experience outcomes—perceived exertion, intrinsic motivation, and intention to continue use—between the experimental group and the control group. IMI subscales (interest/enjoyment, perceived competence, and pressure/tension) were analyzed separately. Significant effects were reported with Cohen *d*. Qualitative interview data were analyzed thematically [34] by the primary interviewer (JZ, female, PhD in digital health; CC, male, MSc in health informatics). The coding team included 2 independent analysts. Neither had prior relationships with participants. Reflexivity measures included maintaining research journals and weekly debriefing sessions during the 3-week analysis period. Analysis followed six phases: (1) familiarization, both coders independently read all 30 transcripts twice; (2) generating initial codes, line-by-line coding in NVivo 14 (Lumivero; Release 1.7.1) produced 156 initial codes; (3) searching for themes, related codes were grouped into 8 candidate themes through iterative discussion; (4) reviewing themes, candidate themes were reviewed against the dataset for internal coherence and external distinction; (5) defining themes, final themes were refined with clear operational definitions; and (6) producing the report, exemplar quotes were selected to illustrate each theme. Inter-coder agreement was moderate (Cohen $\kappa=0.68$); all discrepancies were resolved through consensus discussion. Data saturation was assessed iteratively; no new themes emerged after interview 25 (experimental group: interview 13; control group: interview 12), suggesting thematic saturation was achieved. Member checking was not conducted due to the brief time frame and immediate postassessment interview design; however, findings were triangulated with quantitative data to enhance credibility.

Ethical Considerations

This study was approved by the Institutional Review Board of Shanghai Jiao Tong University (approval number: H20240601I, approved on January 12, 2024). The randomized controlled trial was retrospectively registered at ClinicalTrials.gov (NCT06958653) on July 17, 2025, after enrollment began on April 20, 2025. The delay occurred because the study team initially categorized the randomized component as a low-risk formative evaluation of user experience for a digital assessment interface and did not recognize that randomized allocation requires prospective registration under ICMJE/JMIR policy; once identified during manuscript preparation, the ClinicalTrials.gov record has been corrected/updated to reflect the RCT sample size (n=30) and the substudy sample size (n=10), to designate user experience as the primary outcome domain, and to list the ABC scale as a baseline characteristic rather than an outcome, with the primary outcomes being user experience measures (perceived exertion, intrinsic motivation, and intention to continue use).

All procedures were conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent after receiving detailed verbal and written explanations of the study purpose, procedures, potential risks, and benefits. All participant data were deidentified using unique numeric codes. Video recordings captured during balance assessments contained only skeletal joint position data and did

not record identifiable facial features or other personal characteristics. All images included in the manuscript are representative schematics or composite visualizations created from anonymized skeletal data; no actual participant photographs are shown. The interface screenshots and environmental setup photos were taken in empty testing environments without participants present, ensuring no individual identification is possible. All participants were informed that no identifiable images would be captured or published. Participants received 70 Chinese Yuan (approximately US \$10) as compensation for their time upon completion of the study session. To minimize fall risk during balance assessments, all testing was conducted in a designated area equipped with safety railings on 3 sides.

Results

Phase 1

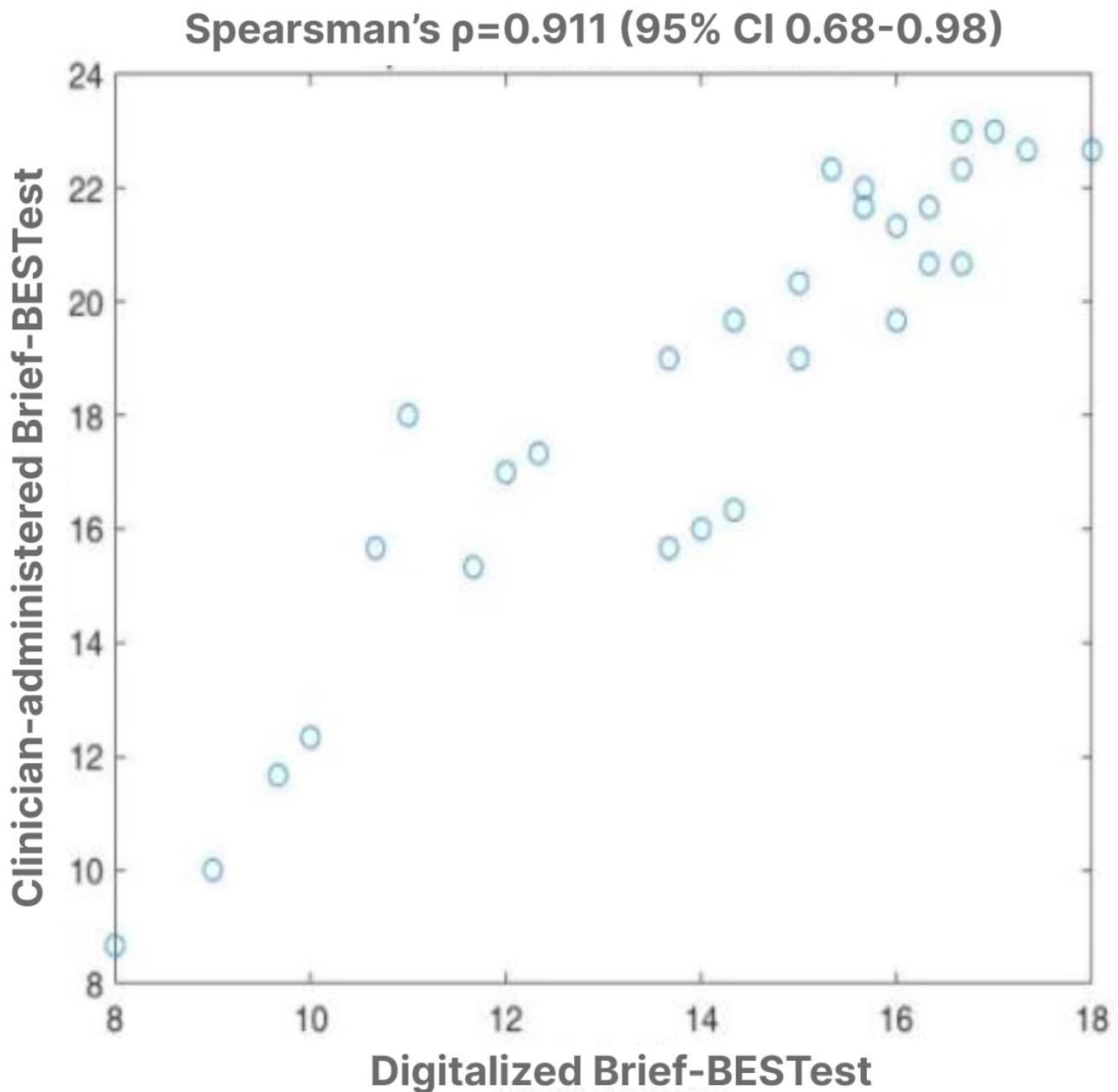
In the first phase, 10 older adults (mean age 64.9, SD 2.76 years; 5 men, 5 women) participated; 5 were excluded due to recent falls requiring medical attention ($n=4$) and declined to participate ($n=1$). Of these, 30% (3/10) reported a prior history of falls. The mean BMI was 26.7 kg/m² (SD 3.1).

Balance performance scores from the clinician-administered Brief-BESTest and the digitalized Brief-BESTest were 14.3 (SD 3.24) and 14.53 (SD 3.03), respectively. A paired-sample t test revealed no significant difference between the 2 assessments ($P=.77$; Cohen $d=-0.07$), indicating strong similarity in 2 sets of results.

Absolute reliability, assessed via the SE of measurement, demonstrated low variability in scores. Relative reliability was reasonable across all 6 Brief-BESTest domains for both assessment methods. ICCs for the clinician-administered version ranged from 0.84 to 0.96 (ICC=0.92; 95% CI 0.84 - 0.96). For the digitalized version, ICCs ranged from 0.78 to 0.94 (ICC=0.89; 95% CI 0.78 - 0.94). Mean differences (MDs) between the 2 test scores ranged from 0.2 to 0.4 points, indicating high consistency.

Convergent validity analysis, as shown in [Figure 3](#), showed a strong and significant positive correlation between the 2 assessment methods (Spearman $\rho=0.91$; $P<.01$), confirming that the digitalized Brief-BESTest provides results comparable to the clinician-administered Brief-BESTest.

Figure 3. Correlation between balance scores obtained from the clinician-administered Brief Balance Evaluation Systems Test (Brief-BESTest) and the digitalized Brief-BESTest in community-dwelling older adults (n=10; mean age 64.9, SD 2.76 years; 50% female; Shanghai, China, April 2025). Spearman $\rho=0.91$ (95% CI 0.68–0.98).



Phase 2

A total of 30 older adults (age mean 66.7, SD 3.93) were recruited for the study and randomized into 15 in the experimental group and 15 in the control group. The recruited participants consisted of 12 men and 18 women, with 43.34% reporting a history of falls. The BMI (kg/m^2) is a mean of 26.73 (SD 3.11), with no significant difference between the experimental group and the control group ($P=.45$). All older adults successfully and independently completed the balance tests, with none of the participants missing in the second round

of testing. The baseline information is specified in [Table 1](#). Meanwhile, the balance ability for the control group (mean 8.67, SD 4.12) and the experimental group (mean 19.4, SD 3.60; $P=.61$) as well as the balance confidence for the control group (mean 0.72, SD 0.18) and the experimental group (mean 0.84, SD 0.17; $P=.08$) were assessed. Indicating similar balance ability and balance confidence between 2 groups of participants. No harm was documented in the experiment. The primary outcomes were user experience measures (perceived exertion, intrinsic motivation, and intention to continue use). The experiment results are specified in [Table 2](#). The CONSORT flow diagram is specified in [Figure 4](#).

Table . Baseline demographic and clinical characteristics of community-dwelling older adults participating in Phase 2 (n=30) randomized controlled trial comparing digitalized Brief Balance Evaluation Systems Test (Brief-BESTest; control) versus Gamified Digital Balance Assessment (experimental). Hongqi Community, Shanghai, China, April-May 2025 (N=30).

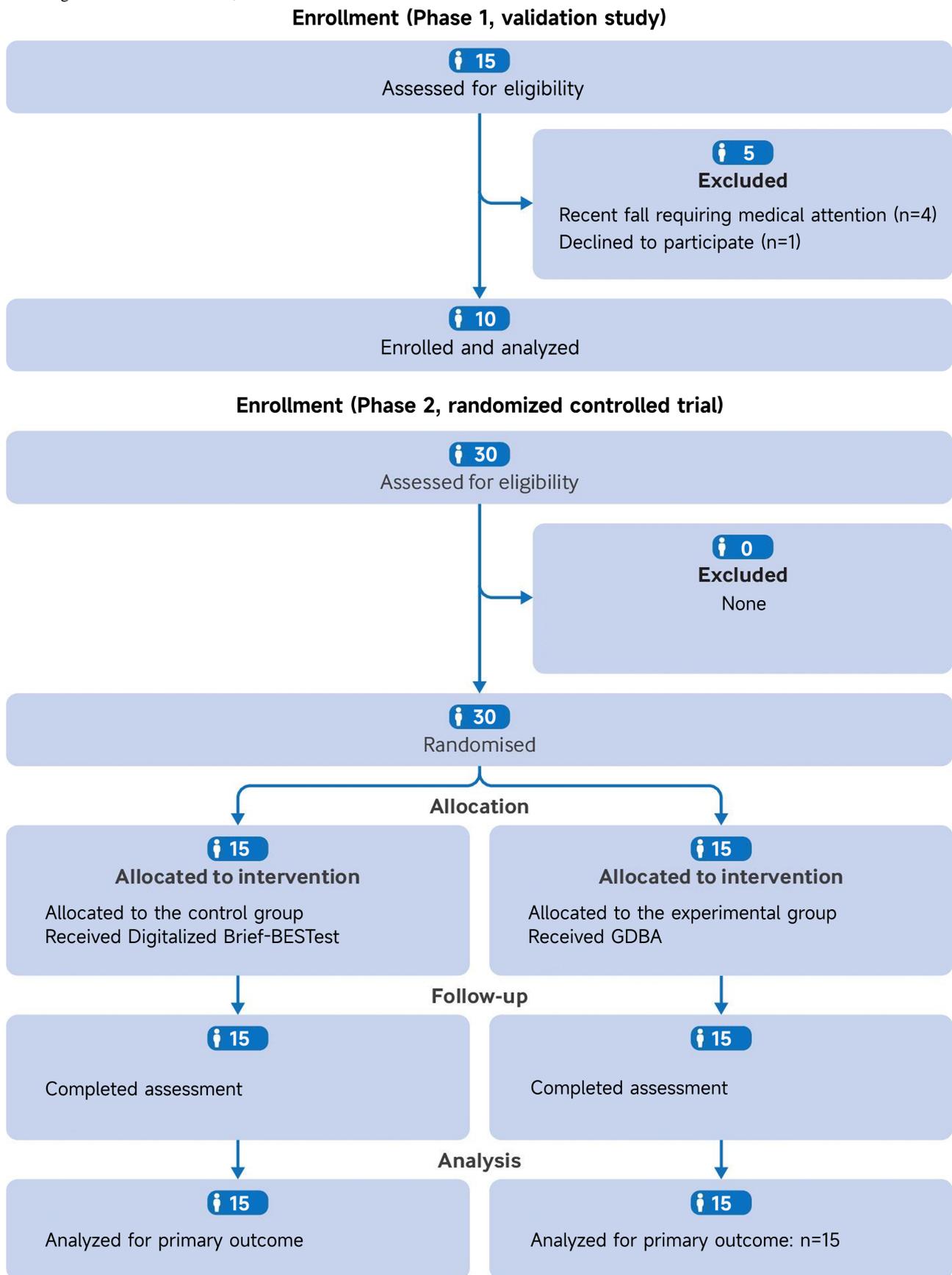
Characteristic	Total (n=30)	Control group (n=15)	Experimental group (n=15)	P value
Age (years), mean (SD)	66.7 (3.93)	65.8 (3.32)	67.6 (4.39)	.21
Gender (female), n(%)	18 (60)	10 (66.67)	8 (53.33)	— ^a
History of falls, n (%)	13 (43.34)	7 (46.67%)	6 (40)	—
BMI (kg/m ²), mean (SD)	26.75 (3.11)	26.58 (3.19)	26.92 (3.03)	.45
Balance ability, mean (SD)	19.04 (3.86)	18.67 (4.12)	19.4 (3.60)	.61
Balance confidence, mean (SD)	0.78 (0.18)	0.72 (0.18)	0.84 (0.17)	.08

^aNot applicable.

Table . Comparison of balance performance and user experience outcomes between the control group (digitalized Brief Balance Evaluation Systems Test [Brief-BESTest], n=15) and the experimental group (Gamified Digital Balance Assessment, n=15) in the Phase 2 randomized controlled trial among community-dwelling older adults. Hongqi Community, Shanghai, China, April-May 2025 (N=30).

Outcome measure	Control group (n=15), mean (SD)	Experimental group (n=15), mean (SD)	Mean difference (95% CI)	Cohen <i>d</i>	P value
Perceived exertion	12.60 (2.75)	9.93 (2.12)	-2.67 (-4.60 to -0.74)	-1.08	.01
Enjoyment and interest	2.73 (1.28)	4.26 (1.33)	1.53 (0.44 to 2.62)	1.17	.009
Perceived competence	3.13 (1.19)	4.27 (1.33)	1.14 (0.19 to 2.09)	0.89	.02
Pressure and tension	3.40 (1.12)	4.13 (1.19)	-0.73 (-1.60 to 0.14)	0.63	.09
Intention to continue use	3.07 (1.28)	4.73 (1.39)	1.66 (0.70 to 2.62)	1.25	<.01

Figure 4. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. Brief-BESTest: Brief Balance Evaluation Systems Test; GDBA: Gamified Digital Balance Assessment;



Perceived Exertion

The study examined the perceived exertion level of participants using the Borg RPE scale. The control group reported a mean perceived exertion of 12.60 (SD 2.75; 95% CI 11.1 - 14.1), while the experimental group reported 9.93 (SD 2.12; 95% CI 8.8 - 11.1). Participants using the GDBA had significantly lower perceived exertion levels compared to those using the digitalized Brief-BESTest (MD -2.67; 95% CI -4.60 to -0.74; $P=.01$; Cohen $d=-1.08$), representing a large effect size.

Intrinsic Motivation

The IMI subscales were used to measure participants' enjoyment and interest, perceived competence, and pressure and tension during the testing process.

For perceived enjoyment and interest, the experimental group reported a significantly higher level (mean 4.26, SD 1.33; 95% CI 3.5 - 5.0) compared to the control group (mean 2.73, SD 1.28, 95% CI 2.0 - 3.5; MD 1.53; 95% CI 0.44-2.62; $P=.009$; Cohen $d=1.17$), indicating a large effect size.

In terms of perceived competence, the experimental group scored higher (mean 4.27, SD 1.33; 95% CI 3.5 - 5.0) than the control group (mean 3.13, SD 1.19; 95% CI 2.5 - 3.8). The use of the GDBA resulted in a significant improvement in perceived competence (MD 1.14; 95% CI 0.19-2.09; $P=.02$; Cohen $d=0.89$), representing a large effect size.

In the dimension of perceived pressure and tension, the experimental group (mean 4.13, SD 1.19; 95% CI 3.5 - 4.8) reported a lower sense of pressure and tension compared with the control group (mean 3.40, SD 1.12; 95% CI 2.8 - 4.0), though the difference was not statistically significant (MD -0.73; 95% CI -1.60 to 0.14; $P=.09$; Cohen $d=0.63$).

Intention to Continue Use

Overview

Quantitative ratings revealed that participants in the experimental group reported a significantly higher intention to continue use (mean 4.73, SD 1.39; 95% CI 3.9 - 5.5) compared with those in the control group (mean 3.07, SD 1.28; 95% CI 2.3 - 3.8). An independent-samples t test confirmed this difference was statistically significant (MD 1.66; 95% CI 0.70-2.62; $P<.01$; Cohen $d=1.25$), indicating a large effect size.

Thematic analysis of the interview transcripts identified 2 overarching themes influencing continued use (E=Experimental, C=Control).

Theme 1: Motivational Rewards

Twelve of 15 GDBA users (80%) cited real-time feedback as motivating, "Seeing my balance score appear right after each task felt rewarding... It kept me coming back." (E04, female, 68).

Seven participants noted game-like features triggered competitive motivation, "When I saw 'Excellent Balance!' with a gold star, I smiled... I wanted that on every task." (E09, female, 71).

In contrast, control participants noted the absence of feedback as demotivating, "I had no idea how I performed until the end. It felt empty." (C03, female, 67).

Theme 2: Perceived Usability

Thirteen GDBA participants (87%) emphasized avatar demonstrations reduced cognitive load, "The 3D character showing exactly what to do was so helpful... I just copied the avatar." (E01, female, 72).

Ten control participants (67%) cited lack of variety as a barrier, "It's functional, but boring... There's nothing interesting about it." (C05, male, 66).

Integration of Quantitative and Qualitative Findings

Qualitative interview data provided convergent evidence supporting the quantitative findings. The significantly higher enjoyment scores in the GDBA group (quantitative) were corroborated by participants' qualitative descriptions of the experience as "fun," "engaging," and "motivating." Specifically, 12 of 15 (80%) GDBA participants mentioned real-time visual feedback as a key source of enjoyment, stating it made the assessment feel "like a game" rather than a clinical test. In contrast, control group participants described the digitalized Brief-BESTest as "straightforward" or "clinical," with 10 of 15 (67%) stating that the lack of interactive elements made the experience feel "monotonous."

The quantitative finding of lower perceived exertion in the GDBA group was explained qualitatively by participants' reports that the visual tutorial and feedback helped them "learning faster on what they should do" and "forget about the effort," suggesting reduced attention to physical strain. Several GDBA participants explicitly stated, "I didn't feel tired because I was concentrating on understanding each task."

The significantly higher intention to continue use (quantitative) was illuminated by qualitative themes emphasizing motivational rewards and perceived usability. As one participant noted, "If I could check my balance regularly and see improvements over time, I would definitely use it every few months." Conversely, control group participants identified specific barriers to continued use, most frequently citing "hard to understand," "lack of variety," and "minimal feedback" as reasons they would not return. This triangulation of quantitative and qualitative data strengthens confidence in the conclusion that gamification elements meaningfully enhance engagement and motivation for digital balance assessment among older adults.

Discussion

Principal Findings

This study pursued two prespecified objectives: (1) to validate a computer vision-based, digitalized Brief-BESTest against a clinician-administered reference standard and (2) to test whether a gamified interface improves user experience during balance assessment among community-dwelling older adults without compromising assessment performance. Both hypotheses were supported. Phase 1 demonstrated that automated scoring by the digitalized system was statistically equivalent to expert clinician scoring and showed strong concurrent validity, consistent with

the hypothesis of adequate reliability and validity. Phase 2 showed that participants assigned to the GDBA reported more favorable user experience across all primary outcomes—including lower perceived exertion, higher intrinsic motivation, and greater intention to continue use—while balance performance was comparable between the groups.

The validation findings are notable because much of the existing literature on camera-based or remote balance testing has concentrated on single-task paradigms, such as the TUG or sit-to-stand tests [10,16], which provide valuable but limited information about postural control. In contrast, the Brief-BESTest assesses 6 theoretically distinct domains of balance [35,36] and its construct validity and fall discrimination ability have been supported across community and institutionalized populations [37]. Transforming such a multidomain instrument into an automated pipeline provides a more clinically comprehensive approach than single-task paradigms [37]. The observed agreement between automated and clinician scoring is consistent with that OpenPose-based skeletal tracking can produce meaningful kinematic measurements in older adults [28]. While our findings should be interpreted as preliminary, they suggest that under standardized conditions, pose estimation algorithms can capture task-relevant movement features aligned with established scoring rubrics [28]. If replicated in larger samples and varied environments, this approach could help extend standardized balance screening beyond hospital, responding to calls for scalable digital fall risk assessment [8].

The user-experience results of gamification align with previous research that indicates gamified elements can improve engagement, enjoyment, and adherence when appropriately tailored to seniors' preferences and capabilities [19,20]. Systematic reviews have concluded that gamification can enhance engagement, enjoyment, and adherence when appropriately tailored to older users [38]. Our findings extend this body of work by focusing on assessment rather than intervention [39]. This distinction matters because the real-world value of screening and monitoring systems depends on repeated voluntary engagement over time, not only on momentary performance [40]. In assessment settings, older adults may also experience evaluative threat or worry about functional decline, which can reduce engagement and distort user experience [41,42]. The observed improvements in enjoyment, perceived competence, and intention to continue use align with self-determination theory, which posits that supporting competence and autonomy enhances intrinsic motivation [29]. Real-time feedback and avatar demonstrations likely reduced uncertainty and clarified task expectations [43], consistent with senior technology acceptance models emphasizing perceived ease of use and perceived usefulness as determinants of adoption [44]. Qualitative data reinforced these interpretations: participants frequently described the feedback as “rewarding” and the avatar guidance as cognitively supportive, echoing findings that interactive and feedback-rich systems can improve engagement among older adults [21]. The reduction in perceived exertion in the gamified group warrants consideration. Theories of attentional focus suggest that engaging, structured feedback may redirect attention away from bodily strain signals, thereby

altering subjective effort perception [38]. Yet, we did not observe significant decreased pressure or tension associated with gamified features, addressing concerns that competitive or evaluative elements may heighten anxiety in older populations [21,45].

Taken together, these findings have practical implications for community screening. Measurement validity is necessary but not sufficient for real-world impact; if older adults experience assessments as confusing or effortful, uptake and repeat use will be limited. By integrating a clinically grounded instrument with older adult-tailored interaction design, systems like GDBA may increase willingness to engage in periodic screening, supporting earlier risk identification and more timely connection to fall prevention services.

Limitations and Future Directions

Despite these promising outcomes, this study has several limitations that also suggest directions for future research. First, the single-session design precludes conclusions about long-term adherence, learning effects, and sustained motivation; longitudinal randomized trials are needed to determine whether early engagement with GDBA leads to habitual use and clinically meaningful outcomes such as reduced falls, improved independence, and lower health care use. Second, participants were relatively healthy, community-dwelling older adults from one urban district in Shanghai, limiting generalizability to frailer, rural, or digitally inexperienced populations. Future work should recruit more diverse samples and incorporate caregiver-assisted or adaptive onboarding to enhance accessibility. Third, validation occurred under optimal indoor conditions, which may not reflect home environments; subsequent studies should test system robustness under variable lighting, space, and camera angles and develop adaptive algorithms to improve accuracy. Fourth, the small Phase 1 sample and possible observer effects from the physiotherapist's presence warrant replication in larger, more naturalistic settings. The use of a single primary physiotherapist may also have introduced assessor bias, despite strong interrater reliability on 20% of cases; future studies should use multiple assessors to verify scoring generalizability. Fifth, qualitative limitations include brief interviews, lack of interviewer blinding, and no member checking; future work should use longitudinal interviews, blinded interviewers, and participant validation. Finally, the study did not examine clinical end points or engagement mechanisms; mixed methods research should explore how gamification influences motivation and behavior change and assess cost-effectiveness, feasibility, and adaptive designs to sustain engagement and expand applicability. In addition, although the study protocol was approved by the institutional review board before recruitment, the trial was registered retrospectively at ClinicalTrials.gov after participant enrollment had begun. This may raise concerns about selective reporting. We mitigated this limitation by ensuring that the registered information, the institutional review board-approved protocol, and the outcomes and analyses reported in this paper are fully aligned, and we commit to prospective registration for all future trials of this system.

Conclusions

This study demonstrates the feasibility of integrating digital assessment with human-centered motivational design in a community-oriented geriatric context. The key contribution to the field lies in providing empirical evidence that evidence-based gamification design can significantly enhance older adults' intrinsic motivation, reduce perceived physical effort, and increase intention for continued use without inducing performance anxiety or cognitive burden, thereby validating theoretical frameworks for practical application in aging populations. More broadly, the findings suggest that validity and engagement should be treated as co-equal design targets in digital health assessment systems, and gamification serves as an effective tool in enhancing user experience and engagement.

In terms of real-world implications, the GDBA represents a scalable, cost-effective technology with the potential to extend accessible balance screening to underserved populations with limited clinical access (particularly in rural or resource-limited areas) and reduce health care use and fall-related morbidity through proactive community-based monitoring. If replicated in larger, prospectively registered, multi-environment studies with longitudinal follow-up, such systems could enable scalable community screening, earlier detection of balance decline, and more proactive fall-prevention referral pathways. Conceptually, this work proposes a translational framework: preserve the clinical assessment criteria in a digital format, validate against expert standards, and optimize experiential design to reduce burden and sustain voluntary engagement among older adults.

Acknowledgments

Special thanks are given to all participants of this research.

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing. The GAI tool used was ChatGPT-5.1, OpenAI. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted through collective responsibility.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy and ethical restrictions (participants did not consent to public data sharing), but anonymized data are available from the corresponding author on reasonable request for researchers who meet the criteria for access to confidential data. Requests should include a detailed research proposal and data use agreement.

Conflicts of Interest

None declared.

Editorial Notice

This randomized study was only retrospectively registered because the study team initially categorized the randomized component as a low-risk formative evaluation focused on user experience of a digital assessment interface and did not recognize that randomized allocation requires prospective trial registration under International Committee of Medical Journal Editors (ICMJE)/JMIR policy. Once this requirement was identified during manuscript preparation, the authors registered the RCT and updated the registry record for transparency and consistency. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low and because the authors provided an official letter from the IRB confirming that the submitted study design is identical to the protocol assessed and approved before trial initiation. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Checklist 1

CONSORT-eHEALTH checklist (V 1.6.1).

[[PDF File, 1199 KB](#) - [games_v14i1e80747_app1.pdf](#)]

References

1. Falls. World Health Organization. 2021. URL: <https://www.who.int/news-room/fact-sheets/detail/falls> [accessed 2026-03-23]
2. Salari N, Darvishi N, Ahmadipanah M, Shohaimi S, Mohammadi M. Global prevalence of falls in the older adults: a comprehensive systematic review and meta-analysis. *J Orthop Surg Res* 2022 Jun 28;17(1):334. [doi: [10.1186/s13018-022-03222-1](https://doi.org/10.1186/s13018-022-03222-1)] [Medline: [35765037](https://pubmed.ncbi.nlm.nih.gov/35765037/)]

3. United Nations. World population ageing 2019 highlights. In: United Nations, Department of Economic and Social Affairs 2019:1-38 URL: <https://digitallibrary.un.org/record/3846855/files/WorldPopulationAgeing2019-Highlights.pdf> [accessed 2026-04-02]
4. Yu S, Master F. China seeks a silver lining as it tells companies to target older consumers. Reuters. 2025. URL: <https://www.reuters.com/world/china/china-seeks-silver-lining-it-tells-companies-target-older-consumers-2025-08-26/> [accessed 2026-04-02]
5. Montero-Odasso MM, Kamkar N, Pieruccini-Faria F, et al. Evaluation of clinical practice guidelines on fall prevention and management for older adults: a systematic review. *JAMA Netw Open* 2021 Dec 1;4(12):e2138911. [doi: [10.1001/jamanetworkopen.2021.38911](https://doi.org/10.1001/jamanetworkopen.2021.38911)] [Medline: [34910151](https://pubmed.ncbi.nlm.nih.gov/34910151/)]
6. Zhu D, Zhao J, Wu T, Zhu B, Wang M, Han T. Effects of a computer vision-based exercise application for people with knee osteoarthritis: randomized controlled trial. *JMIR Mhealth Uhealth* 2025 May 12;13(1):e63022. [doi: [10.2196/63022](https://doi.org/10.2196/63022)] [Medline: [40354624](https://pubmed.ncbi.nlm.nih.gov/40354624/)]
7. Chen B, Chen C, Hu J, et al. Computer vision and machine learning-based gait pattern recognition for flat fall prediction. *Sensors (Basel)* 2022 Oct 19;22(20):7960. [doi: [10.3390/s22207960](https://doi.org/10.3390/s22207960)] [Medline: [36298311](https://pubmed.ncbi.nlm.nih.gov/36298311/)]
8. González-Castro A, Leirós-Rodríguez R, Prada-García C, Benítez-Andrades JA. The applications of artificial intelligence for assessing fall risk: systematic review. *J Med Internet Res* 2024 Apr 29;26:e54934. [doi: [10.2196/54934](https://doi.org/10.2196/54934)] [Medline: [38684088](https://pubmed.ncbi.nlm.nih.gov/38684088/)]
9. Sánchez San Blas H, Filipe Sales Mendes A, Pérez Robledo F, Villarrubia González G, de Paz Santana JF. A multi-agent system approach for balance disorder treatment: integrating computer vision and gamification. *IEEE Access* 2024;12:118649-118664. [doi: [10.1109/ACCESS.2024.3448216](https://doi.org/10.1109/ACCESS.2024.3448216)]
10. Savoie P, Cameron JAD, Kaye ME, Scheme EJ. Automation of the timed-up-and-go test using a conventional video camera. *IEEE J Biomed Health Inform* 2020 Apr;24(4):1196-1205. [doi: [10.1109/JBHI.2019.2934342](https://doi.org/10.1109/JBHI.2019.2934342)] [Medline: [31403450](https://pubmed.ncbi.nlm.nih.gov/31403450/)]
11. Zhao Z, Yang T, Qin C, et al. Exploring the potential of the sit-to-stand test for self-assessment of physical condition in advanced knee osteoarthritis patients using computer vision. *Front Public Health* 2024;12(12):1348236. [doi: [10.3389/fpubh.2024.1348236](https://doi.org/10.3389/fpubh.2024.1348236)] [Medline: [38384889](https://pubmed.ncbi.nlm.nih.gov/38384889/)]
12. Nightingale CJ, Mitchell SN, Butterfield SA. Validation of the timed up and go test for assessing balance variables in adults aged 65 and older. *J Aging Phys Act* 2019 Apr 1;27(2):230-233. [doi: [10.1123/japa.2018-0049](https://doi.org/10.1123/japa.2018-0049)] [Medline: [30117359](https://pubmed.ncbi.nlm.nih.gov/30117359/)]
13. Padgett PK, Jacobs JV, Kasser SL. Is the BESTest at its best? A suggested brief version based on interrater reliability, validity, internal consistency, and theoretical construct. *Phys Ther* 2012 Sep;92(9):1197-1207. [doi: [10.2522/ptj.20120056](https://doi.org/10.2522/ptj.20120056)] [Medline: [22677295](https://pubmed.ncbi.nlm.nih.gov/22677295/)]
14. Shinohara T, Saida K, Miyata K. Ability of the brief-balance evaluation systems test to evaluate balance deficits in community-dwelling older adults: a cross-sectional study. *Physiother Theory Pract* 2022 Oct;38(10):1381-1388. [doi: [10.1080/09593985.2020.1840682](https://doi.org/10.1080/09593985.2020.1840682)] [Medline: [33289587](https://pubmed.ncbi.nlm.nih.gov/33289587/)]
15. O'Hoski S, Sibley KM, Brooks D, Beauchamp MK. Construct validity of the BESTest, mini-BESTest and briefBESTest in adults aged 50 years and older. *Gait Posture* 2015 Sep;42(3):301-305. [doi: [10.1016/j.gaitpost.2015.06.006](https://doi.org/10.1016/j.gaitpost.2015.06.006)] [Medline: [26183191](https://pubmed.ncbi.nlm.nih.gov/26183191/)]
16. Saunders S, Reid J, Mehdipour A, et al. Virtual Assessment Protocols for Timed-Up-and-Go, 5-Repetition Chair Rise, and 1-Minute Sit-to-Stand Tests in community-dwelling older adults: Investigating feasibility, reliability, and validity. *Arch Phys Med Rehabil* 2025 Jun;106(6):926-933. [doi: [10.1016/j.apmr.2024.11.011](https://doi.org/10.1016/j.apmr.2024.11.011)] [Medline: [39615825](https://pubmed.ncbi.nlm.nih.gov/39615825/)]
17. Huotari K, Hamari J. A definition for gamification: anchoring gamification in the service marketing literature. *Electron Markets* 2017 Feb;27(1):21-31. [doi: [10.1007/s12525-015-0212-z](https://doi.org/10.1007/s12525-015-0212-z)]
18. Bosch-Barceló P, Martínez-Navarro O, Masbarnat-Almenara M, Tersa-Miralles C, Pakarinen A, Fernández-Lago H. Gamification integration in technological devices for motor rehabilitation in Parkinson disease: scoping review. *JMIR Serious Games* 2025 Jul 4;13(1):e69433. [doi: [10.2196/69433](https://doi.org/10.2196/69433)] [Medline: [40614259](https://pubmed.ncbi.nlm.nih.gov/40614259/)]
19. Koivisto J, Malik A. Gamification for older adults: a systematic literature review. *Gerontologist* 2021 Sep 13;61(7):e360-e372. [doi: [10.1093/geront/gnaa047](https://doi.org/10.1093/geront/gnaa047)] [Medline: [32530026](https://pubmed.ncbi.nlm.nih.gov/32530026/)]
20. Elzeky MEH, Ramadan El-Shaboury RH, Salameh B, Samaan E, F.M.Shahine N, Abdelhamid A. Effects of gamified versus pedometer-based walking intervention on physical activity, fatigue, and sleep quality among hemodialysis patients: a quasi-experimental study. *BMC Nephrol* 2025;26(1):108. [doi: [10.1186/s12882-025-04012-7](https://doi.org/10.1186/s12882-025-04012-7)]
21. White BK, Martin A, White J. Gamification and older adults: opportunities for gamification to support health promotion initiatives for older adults in the context of COVID-19. *Lancet Reg Health West Pac* 2023 Jun;35(35):100528. [doi: [10.1016/j.lanwpc.2022.100528](https://doi.org/10.1016/j.lanwpc.2022.100528)] [Medline: [35815240](https://pubmed.ncbi.nlm.nih.gov/35815240/)]
22. Bravo-Aparicio J, Trillo-Charlín I, Avendaño-Coy J, Beltran-Alacreu H. Effectiveness of gamification on enjoyment and satisfaction in older adults: systematic review and meta-analysis. *JMIR Aging* 2024;8:e72559-e72559. [doi: [10.2196/72559](https://doi.org/10.2196/72559)]
23. Li A, Qiang W, Li J, Geng Y, Qiang Y, Zhao J. Retracted: evaluating the clinical efficacy of an exergame-based training program for enhancing physical and cognitive functions in older adults with mild cognitive impairment and dementia residing in rural long-term care facilities: randomized controlled trial. *J Med Internet Res* 2025 Feb 19;27:e69109. [doi: [10.2196/69109](https://doi.org/10.2196/69109)] [Medline: [39969990](https://pubmed.ncbi.nlm.nih.gov/39969990/)]

24. Zhang G, Xia Y, Li X, Zhang Y, Xu X, Sun T. A gamification mHealth intervention to enhance adherence to personalized exercise for older adults with chronic diseases: a randomized controlled trial protocol. *Patient Prefer Adherence* 2025;19:3145-3157. [doi: [10.2147/PPA.S547361](https://doi.org/10.2147/PPA.S547361)] [Medline: [41104108](https://pubmed.ncbi.nlm.nih.gov/41104108/)]
25. Moher D, Chan A, Collins G, Hróbjartsson A, Moher D, Schulz K. The CONSORT reporting checklist. The EQUATOR network reporting guideline platform. 2025. URL: <https://resources.equator-network.org/reporting-guidelines/consort/consort-checklist.docx> [accessed 2026-03-23]
26. Mixed methods article reporting standards (MMARS) information recommended for inclusion in manuscripts. American Psychological Association. 2024. URL: <https://apastyle.apa.org/jars/mixed-table-1.pdf> [accessed 2026-03-23]
27. Ali MM, Medhat Hassan M, Zaki M. Human pose estimation for clinical analysis of gait pathologies. *Bioinform Biol Insights* 2024;18:11779322241231108. [doi: [10.1177/11779322241231108](https://doi.org/10.1177/11779322241231108)] [Medline: [38757143](https://pubmed.ncbi.nlm.nih.gov/38757143/)]
28. Ge F, Wu C, Ge F, Xu S, Xiao J. Reliability and validity of OpenPose for measuring HKA angle in dynamic walking videos in patients with knee osteoarthritis. *Sci Rep* 2025 Jul 7;15(1):24286. [doi: [10.1038/s41598-025-09627-2](https://doi.org/10.1038/s41598-025-09627-2)] [Medline: [40624241](https://pubmed.ncbi.nlm.nih.gov/40624241/)]
29. Deci EL, Olafsen AH, Ryan RM. Self-determination theory in work organizations: the state of a science. *Annu Rev Organ Psychol Organ Behav* 2017 Mar 21;4(1):19-43 [FREE Full text] [doi: [10.1146/annurev-orgpsych-032516-113108](https://doi.org/10.1146/annurev-orgpsych-032516-113108)]
30. Chen X, Wu L, Feng H, et al. Comparison of exergames versus conventional exercises on the health benefits of older adults: systematic review with meta-analysis of randomized controlled trials. *JMIR Serious Games* 2023 Jun 22;11(1):e42374. [doi: [10.2196/42374](https://doi.org/10.2196/42374)] [Medline: [37347534](https://pubmed.ncbi.nlm.nih.gov/37347534/)]
31. Borg GAV. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 1982 May;14(5):377. [doi: [10.1249/00005768-198205000-00012](https://doi.org/10.1249/00005768-198205000-00012)]
32. Markland D, Hardy L. On the factorial and construct validity of the intrinsic motivation inventory: conceptual and operational concerns. *Res Q Exerc Sport* 1997 Mar;68(1):20-32. [doi: [10.1080/02701367.1997.10608863](https://doi.org/10.1080/02701367.1997.10608863)] [Medline: [9094760](https://pubmed.ncbi.nlm.nih.gov/9094760/)]
33. Dong Y, Liu X, Tang M, et al. Age-related differences in upper limb motor performance and intrinsic motivation during a virtual reality task. *BMC Geriatr* 2023 Apr 27;23(1):251. [doi: [10.1186/s12877-023-03970-7](https://doi.org/10.1186/s12877-023-03970-7)] [Medline: [37106330](https://pubmed.ncbi.nlm.nih.gov/37106330/)]
34. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp0630a](https://doi.org/10.1191/1478088706qp0630a)]
35. Marques A, Almeida S, Carvalho J, Cruz J, Oliveira A, Jácome C. Reliability, validity, and ability to identify fall status of the balance evaluation systems test, mini-balance evaluation systems test, and brief-balance evaluation systems test in older people living in the community. *Arch Phys Med Rehabil* 2016 Dec;97(12):2166-2173. [doi: [10.1016/j.apmr.2016.07.011](https://doi.org/10.1016/j.apmr.2016.07.011)] [Medline: [27497826](https://pubmed.ncbi.nlm.nih.gov/27497826/)]
36. Viveiro LAP, Gomes GCV, Bacha JMR, et al. Reliability, validity, and ability to identify fall status of the Berg Balance Scale, Balance Evaluation Systems Test (BESTest), Mini-BESTest, and Brief-BESTest in older adults who live in nursing homes. *J Geriatr Phys Ther* 2019;42(4):E45-E54. [doi: [10.1519/JPT.0000000000000215](https://doi.org/10.1519/JPT.0000000000000215)] [Medline: [30407272](https://pubmed.ncbi.nlm.nih.gov/30407272/)]
37. Meseguer-Henarejos AB, López-García JJ, López-Pina JA, Martínez-González-Moro I, Martínez-Carrasco Á. The balance evaluation systems test (BESTest), mini-BESTest and brief-BESTest as clinical tools to assess balance control across different populations: a reliability generalization meta-analysis. *PLoS ONE* 2024;20(4):e0318302. [doi: [10.1371/journal.pone.0318302](https://doi.org/10.1371/journal.pone.0318302)]
38. Hydari MZ, Adjerid I, Striegel AD. Health wearables, gamification, and healthful activity. *Manage Sci* 2023 Jul;69(7):3920-3938. [doi: [10.1287/mnsc.2022.4581](https://doi.org/10.1287/mnsc.2022.4581)] [Medline: [37547027](https://pubmed.ncbi.nlm.nih.gov/37547027/)]
39. 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(65):570-581. [doi: [10.1016/j.apergo.2016.10.013](https://doi.org/10.1016/j.apergo.2016.10.013)] [Medline: [27825723](https://pubmed.ncbi.nlm.nih.gov/27825723/)]
40. Benthien KS, Rasmussen K, Nielsen CP, Kidholm K, Grønkjær M, Toft U. Proactive health support: a randomised controlled trial of telephone-based self-management support for persons at risk of hospital admission. *Age Ageing* 2022 Oct 6;51(10):afac212. [doi: [10.1093/ageing/afac212](https://doi.org/10.1093/ageing/afac212)] [Medline: [36201327](https://pubmed.ncbi.nlm.nih.gov/36201327/)]
41. Holgersson J, Söderström E. Bridging the gap - exploring elderly citizens' perceptions of digital exclusion. Presented at: ECIS 2019 Proceedings 2019; Jun 8-14, 2019 URL: <https://www.diva-portal.org/smash/get/diva2:1317500/FULLTEXT01.pdf> [accessed 2026-04-07]
42. Zhao J, Zhu D, Yang Y, et al. Understanding the effectiveness of an online support group in enhancing recovery identity and health-related quality of life of middle-aged and older ovarian cancer patients: a randomized controlled trial. *Psychol Trauma* 2025 Jun;17(Suppl 1):S196-S205. [doi: [10.1037/tra0001929](https://doi.org/10.1037/tra0001929)] [Medline: [40569778](https://pubmed.ncbi.nlm.nih.gov/40569778/)]
43. Zhao J, Zhu D, Chang F, Han T. Rehab-diary: enhancing recovery identity with an online support group for middle aged and older ovarian cancer patients. *Proc ACM Hum-Comput Interact* 2024 Sep 24;8(MHCI):1-25. [doi: [10.1145/3676520](https://doi.org/10.1145/3676520)]
44. Shin HR, Um SR, Yoon HJ, et al. Comprehensive senior technology acceptance model of daily living assistive technology for older adults with frailty: cross-sectional study. *J Med Internet Res* 2023 Apr 10;25:e41935. [doi: [10.2196/41935](https://doi.org/10.2196/41935)] [Medline: [37036760](https://pubmed.ncbi.nlm.nih.gov/37036760/)]
45. Zhang Y, Xi N, Liang C, Hamari J. The relationships between different forms of gamification and user experience: a study in the context of elderly well-being applications. 2024 Presented at: Hawaii International Conference on System Sciences. [doi: [10.24251/HICSS.2024.410](https://doi.org/10.24251/HICSS.2024.410)]

Abbreviations

ABC: Activities-specific Balance Confidence scale
Brief-BESTest: Brief Balance Evaluation Systems Test
CONSORT: Consolidated Standards of Reporting Trials
EVA: ethylene-vinyl acetate
GDBA: Gamified Digital Balance Assessment
ICC: intraclass correlation coefficient
ICMJE: International Committee of Medical Journal Editors
IMI: Intrinsic Motivation Inventory
JARS: Journal Article Reporting Standards
MD: mean difference
RCT: randomized controlled trial
RPE: Borg Rating of Perceived Exertion scale
TUG: Timed Up and Go test

Edited by S Brini; submitted 16.Jul.2025; peer-reviewed by DE Patil, F Chang, G Amabili; revised version received 22.Feb.2026; accepted 23.Feb.2026; published 15.Apr.2026.

Please cite as:

Zhao J, Xia Y, Zhang Y, Yu J, Chu C

Balance Assessment Using Gamified Digital Technology in Community-Dwelling Older Adults: Mixed Methods Validation Study and Randomized Controlled Trial

JMIR Serious Games 2026;14:e80747

URL: <https://games.jmir.org/2026/1/e80747>

doi: [10.2196/80747](https://doi.org/10.2196/80747)

© Jianan Zhao, Yaqin Xia, Yahui Zhang, Jihong Yu, Chen Chu. Originally published in JMIR Serious Games (<https://games.jmir.org>), 15.Apr.2026. 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.

A Novel Virtual Reality Intervention Combining Movement Exercises and Body Illusions for the Treatment of Chronic Back Pain: Prospective Feasibility Study

Isabel Neumann^{1*}, PhD; Stefan Lindner^{2*}, MSc; Yevgeniya Nedilko³, BA; Ralitsa Zhivova³, MD; Michael Götde⁴, MA; Tobias Tischer-Zeitz², MD; Heike L Rittner², MD; Ivo Käthner^{1,5,6}, PhD

¹Department of Clinical Psychology and Psychotherapy, Institute of Psychology, University of Würzburg, Marcusstraße 9-11, Würzburg, Bavaria, Germany

²Department of Anaesthesiology, Intensive Care, Emergency and Pain Medicine, Centre for Interdisciplinary Pain Medicine, University Hospital Würzburg, Würzburg, Bavaria, Germany

³Biomim Non-Profit UG (Ltd), Frankfurt, Germany

⁴videoreality GmbH, Frankfurt, Germany

⁵Department of Medicine, HMU Health and Medical University, Potsdam, Brandenburg, Germany

⁶Department of Physiological Psychology, University of Bamberg, Bamberg, Bavaria, Germany

*these authors contributed equally

Corresponding Author:

Isabel Neumann, PhD

Department of Clinical Psychology and Psychotherapy, Institute of Psychology, University of Würzburg, Marcusstraße 9-11, Würzburg, Bavaria, Germany

Abstract

Background: Virtual reality (VR) has proven effective in delivering nonpharmacological interventions to reduce acute and chronic pain. For the treatment of nonspecific chronic low back pain (CLBP), it offers benefits over traditional treatment options, such as the possibility of gamified movement exercises with real-time performance feedback and virtual embodiment. We implemented a novel immersive VR intervention (a serious game) that combined these elements.

Objective: This study evaluated the feasibility, tolerability, and initial clinical efficacy of the gamified VR intervention.

Methods: Patients with CLBP ($n=20$; mean age 47, SD 14 years; symptom duration >3 months to ≤ 5 years; convenience sample) took part in a prospective, single-arm, and preregistered trial over 9 weeks. The VR therapy phase lasted 3 weeks, and there were 2 VR sessions per week conducted at the University Hospital Würzburg (Germany). Before the therapy phase, there was a 2-week baseline phase, and the posttherapy phase lasted 4 weeks. During the sessions, patients wore a head-mounted display. In VR, they embodied a virtual avatar and performed gamified movement exercises. Participants were immersed in a virtual toy factory, and they had the task of teaching 5 different toys how to move. They received real-time feedback on performance through a hologram overlaying their avatar. Based on performance, movements to be performed became gradually more difficult (graded exposure). Primary outcome measures were adherence and side effects for assessing feasibility and tolerability (Simulator Sickness Questionnaire), and pain intensity ratings (numerical rating scale 0 - 10) were used for assessing initial clinical efficacy. Secondary outcomes included back- and task-specific functioning and questionnaires to further test initial clinical efficacy and assess fear-avoidance beliefs.

Results: Adherence was high (18/20, 90%). Participants indicated lower pain in the posttherapy phase compared with baseline levels (mean difference 0.73, CI 0.27-1.19; $t_{16}=3.38$; $P=.004$; $d=0.82$). There were only few and minor side effects. Task- and back-specific functioning improved (ie, performing daily life activities; Back Performance Scale: $F_{2,34}=4.53$; $P=.02$; $\eta^2_g=0.04$; Roland-Morris Disability Questionnaire: $F_{2,26}=4.73$; $P=.02$; $\eta^2_g=0.08$), and movement restrictions decreased ($F_{2,32}=10.82$; $P<.001$; $\eta^2_g=0.06$). There were no changes in the psychological outcome measures (eg, fear avoidance beliefs). Across all VR sessions, study participants reported high levels of fun (mean 8.07, SD 1.99).

Conclusions: We implemented a gamified immersive VR intervention for the treatment of CLBP. The combination of gamified movement exercises and virtual body illusions is unique, and for the first time, it included real-time feedback via a hologram overlaying the virtual avatar of the user. The study demonstrated the feasibility and safety of the intervention. Initial tests of the clinical efficacy revealed positive effects on pain, physical functioning, and daily activities. However, these did not reach the thresholds of clinical importance. A randomized controlled trial is needed to test the specificity of the effects.

Trial Registration: German Clinical Trials Register DRKS00031535; <https://drks.de/search/en/trial/DRKS00031535>

(*JMIR Serious Games* 2026;14:e81051) doi:[10.2196/81051](https://doi.org/10.2196/81051)

KEYWORDS

immersive virtual reality; chronic low back pain; gamification; kinesiphobia; feasibility

Introduction

Background

About 1 in 5 people experience chronic pain, with chronic back pain being the most common disorder [1,2]. Chronic pain can be a debilitating condition negatively affecting quality of life and mental health for the individual and leading to high societal costs [3-5]. Pharmacological treatments are often not effective and are associated with substantial side effects and risks, such as sedation, gonadal dysfunction, sleep disturbances, and abuse of opioid treatments [6]. Hence, there is a need for nonpharmacological treatments, especially as interdisciplinary multimodal pain therapy programs are considered the gold standard for day patients [7]. Virtual reality (VR) is a powerful technology to deliver nonpharmacological treatments for acute and chronic pain due to its immersive properties [8-10].

Prior Work

Early VR studies focused on distraction and induction of positive affect for the treatment of acute pain [11-15]. More recent VR studies for chronic pain explored the feasibility of implementing a variety of treatments that go beyond these mechanisms, including pain education and pain relief skills training [16,17], embodiment [18-21], physical exercises [22-24], and mindfulness-based interventions [25]. In short, VR allows the implementation of traditional treatment approaches while employing additional, VR-specific mechanisms (refer to [9] for a review).

For nonspecific chronic low back pain (CLBP), pain-related fear and subsequent avoidance behavior are major factors for pain chronification and maintenance [26-28]. The fear-avoidance model of CLBP offers a cognitive-behavioral explanation for the development and maintenance of chronic pain based on approach and avoidance behaviors [27,28]. While protective behavior and movement avoidance can be useful for a limited period of time after an injury, perceiving pain as highly threatening can lead to catastrophizing, pain-related fear, and consequently avoidance through negative reinforcement. If maintained and generalized to novel situations, avoidance negatively affects valued activities and can lead to pain hypervigilance and depressive symptoms, which in turn negatively affect pain, resulting in a vicious circle. Accordingly, pain-related fear, catastrophizing, pain hypervigilance, and avoidance are predictors for chronic pain development (eg, [29-35]) and for experienced pain intensity and disability associated with pain [29,35-37], and they are associated with the status of pain recovery at a later time point [38,39]. By systematically reducing pain-related fear and avoidance behaviors, graded exposure in vivo can be a strategy to interrupt the vicious circle [40].

Physical exercise is a recommended first-line treatment for CLBP [41,42]. Systematic reviews suggest that treatment adherence is key to its efficacy in reducing pain and disability [43,44] but that adherence outside the clinic is low [45]. For CLBP, VR offers several advantages over traditional treatment approaches [46]. It can increase user motivation and engagement through gamification [23,47], offers the possibility of precise real-time feedback of movements, and holds the promise of standardized therapies that mitigate costs associated with supervision necessary during standard care. In addition, VR therapy has an immediate analgesic effect through distraction [48], which might help persons with kinesiphobia to exercise despite pain-catastrophizing thoughts and pain-related fear.

To reduce acute pain, previous VR interventions often immersed participants in pleasant virtual scenarios to heighten the mood and, thereby, achieved significant pain reduction [13-15]. For chronic pain and nonspecific CLBP in particular, regular exercise can be a key to recovery. Studies from the field of sports science investigated how affect during exercise influences future physical activity behavior [49]. In line with the hedonic theory, initial evidence suggests that positive affect during exercise increases the likelihood of future physical activity behavior [50] and negative affect reduces the likelihood [51] (see [49] for a review). VR appears to be a particularly suited method to induce a pleasurable exercise experience through high levels of embodiment, presence, and interactivity [52-54], and, thereby, promote future exercise. Gamified exercises could be a means to increase fun and could lead to higher adherence for patients with CLBP, as suggested by a previous article [46].

Furthermore, virtual embodiment could help to correct distorted body perceptions or maladaptive beliefs about the body (eg, fragile back) that are common in CLBP [55-58]. Although the benefits of VR for the treatment of CLBP are promising, there are very few interventions in immersive VR to date [10].

Goal of This Study

For this study, we implemented a novel immersive VR therapy that combines gamified movement exercises designed as graded exposure therapy with innovative real-time feedback and virtual body illusions. Specifically, participants were immersed in a virtual toy factory presented through a head-mounted display (HMD). Within the virtual toy factory, study participants had the task of teaching 5 different toys how to move. Meanwhile, they received real-time feedback on performance through a hologram overlaying their avatar. Based on performance, movements to be performed became gradually more difficult (graded exposure). According to the guidelines for clinical VR studies in health care proposed by a previous article [59], our study was a VR phase 2 trial that needed to be conducted prior to a randomized controlled trial (VR phase 3). We focused the evaluation of the intervention on feasibility and tolerability, and

gathered data to assess initial clinical efficacy, with a focus on pain reduction as the main clinical outcome.

Methods

Study Design

The feasibility, tolerability, and initial efficacy of a novel VR intervention for CLBP were tested in a prospective, single-arm, proof-of-concept study. The study protocol was approved by the Ethical Review Board of the Faculty of Medicine of the University of Würzburg (190/22-am) and preregistered at the German Clinical Trials Register (ID DRKS00031535). We adhered to the procedures described in our study registration. The study took place from May 2023 to February 2024.

Ethical Considerations

Ethics approval to conduct this research was obtained from the Ethics Committee of the University of Würzburg prior to its start (IRB 190/22-am; November 2022). All study procedures were conducted in accordance with the Declaration of Helsinki. In particular, the study protocol, including the intervention and data collection methods, was considered to comply with the ethical standards for research involving human participants, and data were collected and stored in line with the European Union data protection law (General Data Protection Regulation [GDPR]). All participants provided written informed consent before participation. The consent form outlined the purpose, procedures, risks, benefits, and voluntary nature of participating in the study in plain terms. Participants were informed of their right to withdraw at any time without penalty. Research data were deidentified to ensure the privacy of all participants. Personal identifiers were not linked to the study data. No identifiable information will be shared with external parties or institutions. Electronic records were stored securely on password-protected devices. Participants provided consent that their anonymized data can be stored in a public repository to allow for secondary data analysis. Study participants received €60 (approximately US \$70.75) for their participation.

Participants

We recruited 20 study participants with CLBP via convenience sampling through internal postings at the University Hospital Würzburg, Germany, and newspaper advertisements. The ideal sample size was calculated a priori with G*Power (version 3.1.9.2 [60]) based on an expected medium to large effect size ($d=0.77$) for pain reduction, with a power of 0.8 and an α value of .05. This yielded 16 study participants, and to account for dropouts, we planned on recruiting 20 participants. The inclusion criteria were as follows: age 18 - 65 years; diagnosis of nonspecific CLBP (symptom duration >3 months to ≤ 5 years and current pain intensity of $\geq 4/10$ on a numerical rating scale [NRS]); and sufficient vision and language comprehension for the intervention. The exclusion criteria were as follows: specific back pain (eg, spinal canal stenosis, disc protrusion, and carcinoma); other chronic pain conditions (eg, fibromyalgia); current neurological or mental disorders; and other current diseases, including infectious, metabolic, endocrine, and severe internal secondary diseases. Participants were also excluded if they were receiving antidepressant medication, they were

pregnant, or their glasses did not fit under the HMD. The inclusion and exclusion criteria were checked with structured questionnaires in case no current patient record was available. The mean age of the included participants was 46.85 (SD 14.25; range 24 - 65) years, and the mean duration of back pain was 3.00 (SD 1.49; range 1 - 5) years. Out of the 20 included participants, 13 were female.

Outcome Measures

Primary Outcome Measures

Feasibility and Tolerability

To assess feasibility (primary outcome), we assessed adherence and side effects. We defined it as a success if $\geq 80\%$ of participants completed the intervention, no serious side effects occurred (eg, injuries caused by falls), and $\leq 20\%$ of participants experienced cybersickness (assessed with the Simulator Sickness Questionnaire [SSQ]) [61].

Pain Intensity

The co-primary outcome measure was pain intensity. The participants were asked to rate their pain intensity with the Patient-Reported Outcomes Measurement Information System 29 (PROMIS 29) scale from 0 (no pain) to 10 (worst pain imaginable) each day at a fixed time (when they usually experience high pain). We compared the mean ratings of the baseline period (baseline pain levels) prior to the intervention with the mean ratings assessed in the posttreatment phase (posttreatment pain levels).

Secondary Outcome Measures

The secondary outcome measures described in the subsections below relate to the real-life impact on physiological outcomes (Range of Motion and Other Behavioral Measures subsection) and to the fear-avoidance model of chronic back pain that influenced the design of the intervention (Psychological Measures subsection). We investigated the initial impact on movement restrictions and daily life activities. The measures described below are a combination of subjective, patient-reported measures and objective tests.

Range of Motion and Other Behavioral Measures

As a secondary measure, we analyzed range of motion (ROM). A physiotherapist measured ROM in a standard physical assessment and classified ROM loss in 4 clinically important categories (nil, minimal, moderate, and major), as generally used in the McKenzie Method of Mechanical Diagnosis and Therapy (MDT) [62]. The reliability of the MDT has been accepted, and this method uses a practicable way to assess clinically important differences in ROM change [63]. The possible restrictions in ROM were assessed at the 3 clinical assessments (prestudy, poststudy, and follow-up). To quantify the classification of movement restrictions, we assigned a numerical value to every class (nil=0, minimal=1, moderate=2, and major=3). An overall score was calculated for each person and each time point for an objective comparison of ROM changes.

For measuring individual patient-related activities, we asked the participants to fill in the Patient-Specific Functional Scale

(PSFS) [64]. In this questionnaire, the participants are required to state 3 to a maximum of 5 activities that they are unable to perform or only perform with difficulty due to back pain and that they would like to improve. They rate each of the named activities on a Likert scale from 0 (unable to perform the activity) to 10 (can perform the activity at the same level as before the injury or problem), with a higher score indicating an improvement in functionality. The minimal clinically important difference (MCID) for the mean value of the 3 rated activities is 2 points [65]. We administered the German version of the Roland-Morris Disability Questionnaire (RMDQ) [66,67] to measure how the participants' back pain restricts specific activities. The participants were required to tick 1 (yes) or 0 (no) for restrictions in certain activities, resulting in a score ranging from 0 to 24. The minimal detectable change (MDC) is 2 - 3 points [68]. At the end of the last session and at the end of the poststudy phase, we collected responses to the modified Patient Global Impression of Change (PGIC) [69]. The participants were asked to specify the changes related to their painful condition on a Likert scale from 1 (no change or the condition has gotten worse) to 7 (a great deal better and a considerable improvement that has made all the difference). Together with a physiotherapist, the participants also completed the Back Performance Scale (BPS) [70] for the evaluation and assessment of everyday functions of the lumbar spine on a Likert scale from 0 (no restriction) to 3 (massive restrictions). The BPS consists of 5 functional tests to measure mobility and load capacity in the most common sagittal plane of the spine (sock test, long-sitting test, finger-floor test, lift-up test, and lifting test), resulting in scores of 0 - 15 points. The MCID is 3.6 points [71].

Psychological Measures

We assessed beliefs about the influence of physical activity and work on pain and psychological distress, and the extent to which pain is perceived as threatening. To this end, the participants were asked to complete the German version of the Fear Avoidance Beliefs Questionnaire (FABQ) [72,73] on a Likert scale from 0 (completely disagree) to 6 (completely agree). The score range is 0 - 24 for the subscale physical activity and 0 - 42 for the subscale work, with higher values indicating more fear-avoidance beliefs. Regarding chronic nonspecific low back pain, the MDC is 3.69 points for the subscale physical activity and 5.95 points for the subscale work [74]. The participants were also asked to fill in the PROMIS 29 [75,76] on a Likert scale from 1 (not at all) to 5 (a lot). The score range is 1 - 5 for each of the 7 domains (physical functioning, anxiety, depression, fatigue, sleep disturbance, limitations due to pain, and participation in social roles and activities) to measure physical, mental, and social health. The MCID is between 3 and 5 points [77]. According to the fear-avoidance model, if individuals perceive pain as highly threatening, it can lead to catastrophizing. Therefore, we assessed pain catastrophizing via the German version of the Pain Catastrophizing Scale (PCS) [78,79] on a scale from 0 (not at all) to 4 (always). The score range is 0 - 52, with an MDC of 7.73 points [80]. We used the German 11-item version of the Tampa Scale for Kinesiophobia (TSK) [64,65], which requires the participants to indicate their agreement with statements related to painful movement-related

anxiety on a Likert scale from 1 (strongly disagree) to 4 (strongly agree). The score range is 11 - 44, with higher scores indicating stronger kinesiophobia, and the MDC is 5.6 points [66].

VR Interface and Hardware

During the VR intervention, the participants wore an HMD (Oculus Meta Quest 2) and held controllers (Oculus Meta Quest 2 tracking sensors) in both hands while standing. Their movements were tracked by an Azure Kinect DK system (Microsoft Corp) positioned in front of them.

The virtual environment was created with Unity (Unity Software Inc). For the intervention, we implemented gamified movement exercises, for which the participants were immersed in a virtual toy factory. Within the virtual toy factory, the participants had the task of teaching 5 different toys how to move.

We implemented movement exercises that were selected by a physiotherapist to be compatible with the goals and constraints of the planned VR intervention for the treatment of CLBP. It is unclear which exercise therapy is most effective for CLBP [81,82]. However, clinical practice guidelines on CLBP strongly recommend exercise therapy [42]. General activity of the participants appears to be more important than the type of exercise [83], and correctly performed exercise therapy is safe [84]. We selected standing movements in one place for safety reasons, as the real world is hidden from view, and so that the participants can always perceive the feedback provided by the hologram. The following movements, which are often restricted in patients with CLBP [65,70], were selected and performed repetitively: (1) lateral flexion of the lower spine (participants held their hips with their hands and moved the hips to the left and right sides), (2) squat (with a slightly wider stance and movement of the arms in flexion while squatting deeper for balance control), and (3) trunk rotation (in 2 different styles: one involving global rotation with arm movements like an archer, and the other with the hands at the hip and turning with the shoulder line) [62]. Prior to each movement exercise, the participants received detailed instructions in VR via a virtual screen. During the movement exercises, the participants saw a virtual hologram overlaying their virtual body and were asked to follow the movements of the hologram.

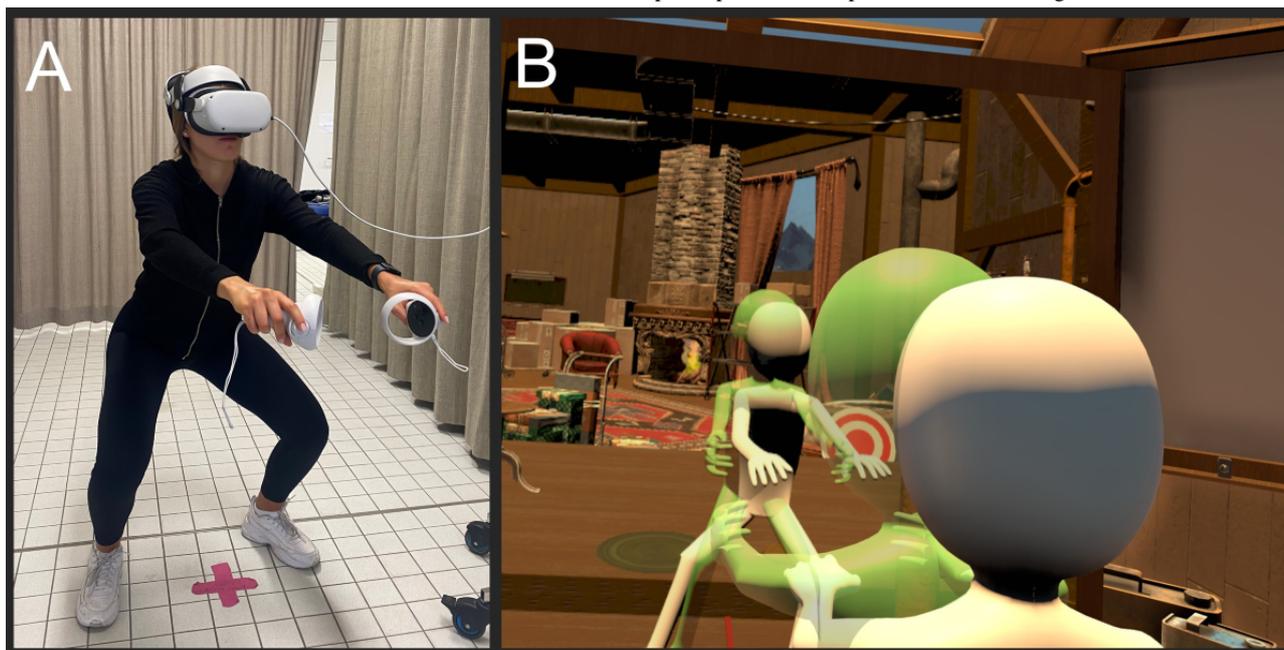
To determine the movement range to be performed, we conducted a usability study with 20 healthy participants prior to the clinical feasibility study. The mean value of the maximum ROM of the second to fourth best subjects was determined as the highest level of difficulty (level 3) for the participants with CLBP. We omitted the best value due to potential measurement errors or statistical outliers. The other 2 levels were calculated in equal parts from the difference between the mean threshold and the starting position (Multimedia Appendix 1).

The hologram provided the participants with real-time visual feedback on their movement accuracy (Figure 1). In addition, they received a performance score and auditory feedback (eg, "well done" after they finished an exercise). The movement performance score was based on an overlap with the ideal movement pattern determined in the study with healthy participants. While the participants performed the movement,

the hologram overlaid their virtual body (based on the ideal movement pattern). The score was based on a match (within a margin of error) between the movement performed (depicted by the avatar of the user) and the “ideal” movement (displayed through the hologram). A higher score was given if there was a closer match between both movements. Each movement exercise consisted of 3 difficulty levels that could be selected (ROM gradually increased with each level). An algorithm was used to increase the level of difficulty. All participants started

at the first level for all exercises, and in each session, the participants performed all 4 movements at least once. If the pain intensity increased by more than 1 point on the NRS, the difficulty level of that exercise was not increased. The participants were allowed to try 2 of the other exercises at an increased level in the same session. In each subsequent session, the participants started at the previously performed difficulty level. The aim was to achieve the greatest ROM within the VR sessions without provoking a clinically relevant increase in pain.

Figure 1. Illustration of movement exercises in virtual reality (VR). (A) A female person in VR equipped with a head-mounted display and controllers, who is executing a movement exercise. (B) View of the person in VR seeing her embodied body from a third-person perspective and a hologram for feedback of the movement exercise. Consent has been obtained from the depicted person for the publication of this image.



Embodiment of the Virtual Character

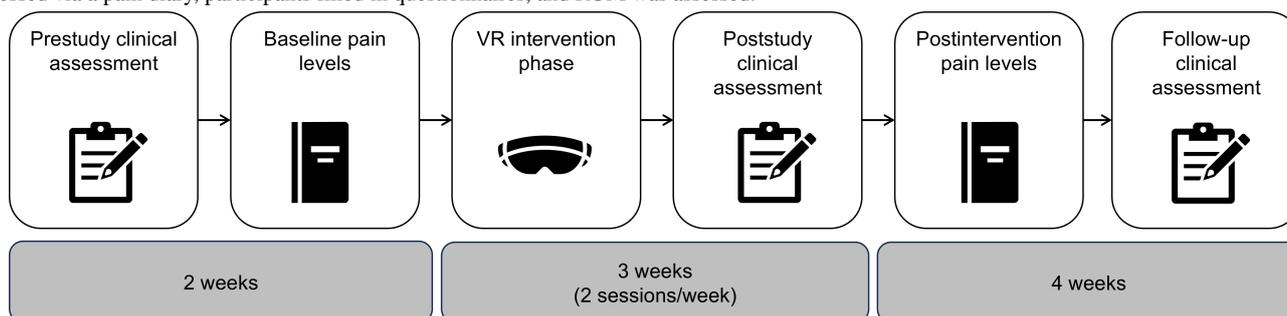
At the beginning of the VR session, upon entering the toy factory (after the calibration phase and pain ratings), the participants were told to accustom themselves with the virtual body. They saw the virtual body from a first-person perspective and were asked to perform movements. The virtual body moved according to their real movements that were tracked via the handheld controllers and the Azure Kinect sensor. This procedure was previously demonstrated to induce a full-body illusion (eg, [85,86]; see Multimedia Appendix 1). A virtual mirror was placed in front of the avatar for the participant to

observe the movements of the body. In the second step, a virtual teleportation procedure was displayed (visual effect), and the avatar’s position was moved such that subsequently the participant observed it from a third-person perspective while keeping full agency over it.

Procedure

The prospective feasibility study lasted 9 weeks for each patient (Figure 2). Prior to the prestudy clinical assessment, the participants were screened according to the inclusion and exclusion criteria via a phone call.

Figure 2. During the preintervention phase (2 weeks), study participants with chronic low back pain (n=20) filled in questionnaires, range of motion (ROM) was assessed, and baseline pain levels were assessed via a pain diary. In the intervention phase (3 weeks), the virtual reality (VR) intervention was conducted, participants filled in questionnaires, and ROM was assessed. In the postintervention phase (4 weeks), postintervention pain levels were assessed via a pain diary, participants filled in questionnaires, and ROM was assessed.

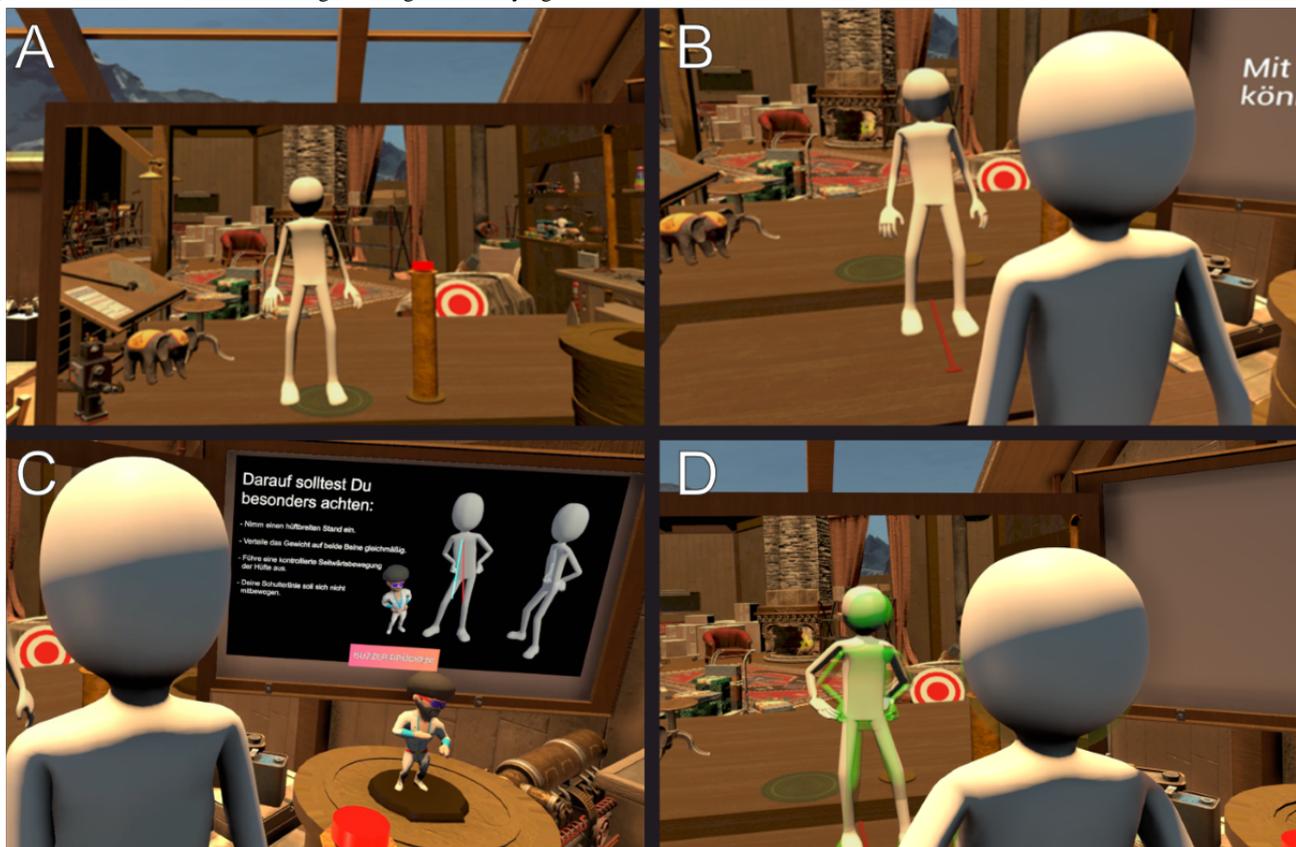


In the preintervention phase, the participants underwent a medical examination, including an assessment of their ROM and a query of current medications (after providing informed consent). If eligibility was confirmed, the participants were asked to fill in the questionnaires that served as secondary outcome measures. In the following 2 weeks, the participants were asked to fill in a pain diary at the time of the expected highest pain intensity. They were also asked to estimate the maximum pain intensity at an individual time point at the end of each day on an NRS (ranging from 0 to 10). Furthermore, they were asked to indicate the duration and intensity of pain exacerbations and the possible use of pain medication.

The intervention phase lasted 3 weeks and included 2 VR sessions per week (on separate days with at least 1 day off between sessions) with a duration of 30 minutes per session

(Figure 3). The number of sessions per week was chosen to ensure regular exercise and make study participation manageable in daily life. The exercise duration was similar to the duration of a regular physiotherapy session. In each session, there were 6 exercises, and the participants practiced each exercise. The order of exercises could be chosen by the participants. The difficulty level of the movement exercises was gradually increased over the course of treatment. At the beginning of each VR session and after each movement exercise, the participants were asked to rate their current pain intensity. Furthermore, after each movement exercise (while being immersed in VR), the participants received feedback from the VR application about their performance. At the end of each VR session, the participants were asked to indicate their levels of presence (how much they felt as if they were in the virtual environment) and fun during the application on an NRS (ranging from 0 to 10).

Figure 3. Overview of the virtual reality session procedure. (A) Upon entering the virtual toy factory, participants observed their avatar from a first-person perspective. Their movements were tracked, and the virtual body moved accordingly. They were standing in front of a virtual mirror, enabling them to see their whole body. (B) In the second step, participants observed their avatar from a third-person perspective. (C) Prior to each exercise, the required movements were demonstrated on a virtual screen. In the depicted case, lateral flexion of the lower spine was required to teach the toy how to dance (study participants held their hips with their hands and moved the hips to the left and right sides). (D) While participants performed the movements, they received real-time feedback through a hologram overlaying their avatar.



After the last VR session, the participants were physiotherapeutically assessed again. ROM was measured, and they were asked to fill in the questionnaires that served as secondary outcome measures. Furthermore, they received instructions and information about movement exercises to be practiced at home at their own discretion during the poststudy phase. For this, they were sent 3 videos by email so that they could practice each exercise at least once a day. These 3 exercises (squat, lateral flexion, and rotation) in the standing

position were the same as executed in VR and with the same number of repetitions.

In the postintervention phase, the pain diary was continuously filled in daily. The participants were asked to state in a retrospective verbal question how often they had practiced the movement exercises, to which all participants responded in the affirmative. A follow-up assessment took place 4 weeks after the last VR session, in which ROM was assessed and the secondary outcome questionnaires were filled in.

Statistical Analysis

Primary outcomes in this study were pain intensity ratings, adherence, and side effects. Secondary outcomes in this study were changes in ROM, TSK, FABQ, PGIC, PSFS, BPS, RMDQ, PROMIS 29, and PCS scores from the prestudy clinical assessment to the poststudy clinical assessment.

The primary outcome measures were analyzed via frequencies (adherence and side effects) and 2-sided *t* tests (pain ratings). All secondary outcome measures were analyzed separately with repeated measures ANOVA (rmANOVA) with the within-subjects factor time (prestudy, poststudy, and follow-up). In case of significant main effects, 2-sided post hoc *t* tests were conducted. Regarding missing data, we used the Little MCAR (missing completely at random) test; however, the findings did not reach significance (all $P > .05$; see Table S1 in [Multimedia Appendix 2](#) for the results of the Little MCAR test and the number of missing data per outcome variable). We have reported means and SDs. The significance level was set at $P < .05$.

Statistical analyses were conducted in R 4.1.0 [87]. The rstatix package [88] was used for the rmANOVA, the lsr package [89] was used for *t* tests, the psych package was used for descriptives [90], and the ggplot2 package [91] was used for data visualization.

Results

Primary Outcome Measures: Feasibility, Tolerability, and Pain Relief

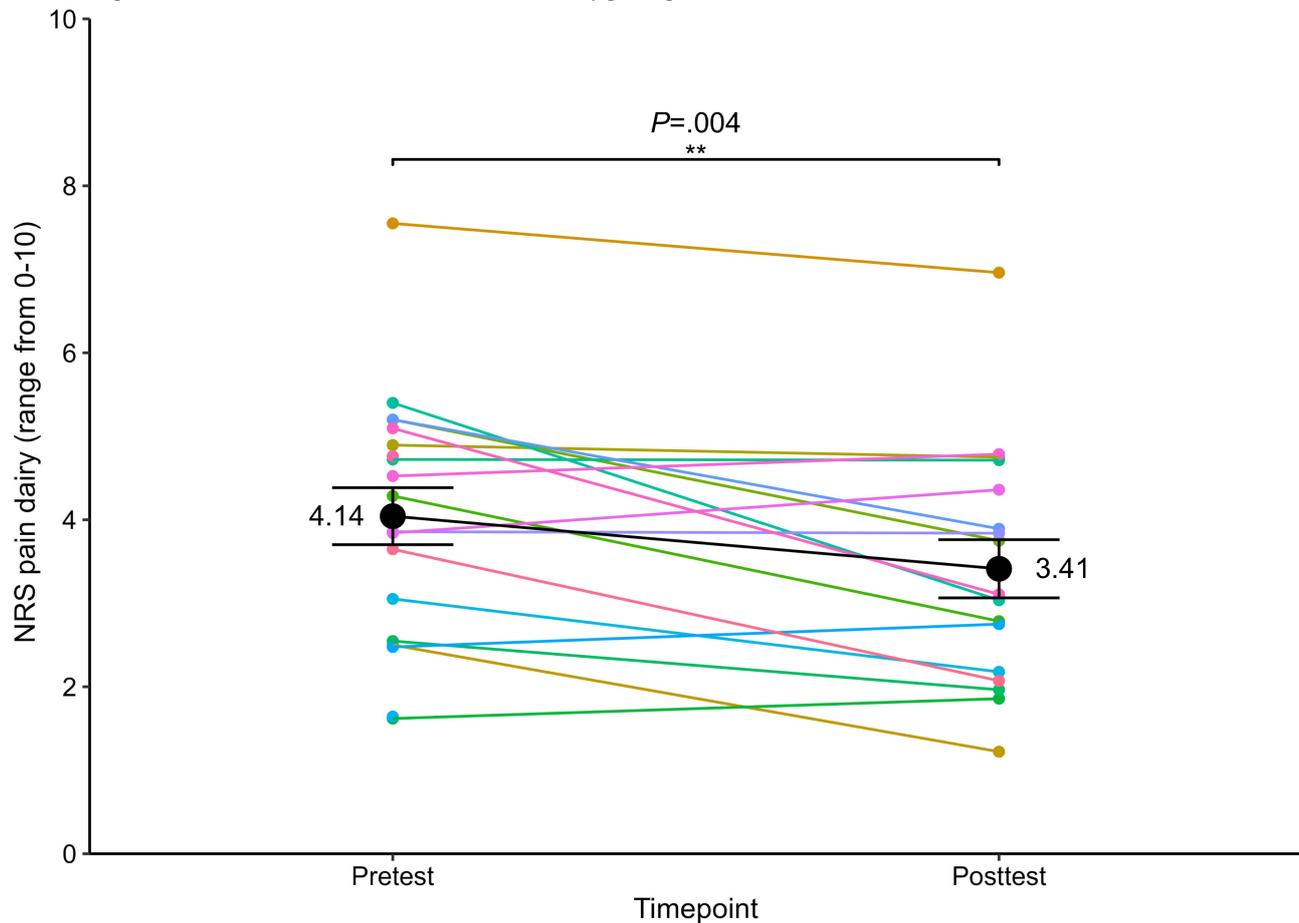
The participant flowchart is presented in [Figure 4](#). Out of the 20 study participants included, 18 (90%) completed the study, and no serious side effects occurred. Two study participants decided to stop participating (one participant after the first session, and the other participant after the third session in VR). One participant indicated not having fun and did not see it as relevant to an individual's pain symptoms, and another participant perceived the toy factory scenario as not serious enough for treating the pain condition.

Figure 4. Participant flow diagram. Interested participants with chronic low back pain were screened via a phone call based on the inclusion criteria. They were then invited to a prestudy clinical assessment, and if they met the inclusion criteria, they were included in the study. All included participants experienced the virtual reality intervention in the intervention phase. BPS: Back Performance Scale; NRS: numerical rating scale; ROM: range of motion. *Phone screening: experiencing chronic back pain for >5 years (n=7), experiencing other pain (eg, thrombosis and chronic bone pain; n=2), being too old (n=1), and taking opioids (n=1); Clinical assessment: experiencing chronic back pain for >5 years (n=2), experiencing specific back pain (n=2), and localization of pain not fitting (thoracic and cervical spine; n=2).

Regarding potential symptoms of cybersickness, there were only few and minor side effects assessed with the SSQ, with the most frequent relating to sweating, strained eyes, and difficulties in focusing with the HMD (see Table S1 in [Multimedia Appendix 1](#) for the number of participants indicating SSQ symptoms per session and Figure S1 in [Multimedia Appendix 1](#) for mean SSQ scores per session).

We identified a significant pain reduction from baseline pain levels (mean 4.14, SD 1.44) to postintervention pain levels (mean 3.41, SD 1.44), with a mean difference of 0.73 (CI 0.27-1.19; $t_{16}=3.38$; $P=.004$; $d=0.82$) ([Figure 5](#)). Baseline pain levels did not differ from day 1 to day 14 ($t_{14}=0.23$; $P=.82$; $d=0.06$), demonstrating baseline stability. Among the participants who experienced pain reduction ($n=13$), 6 experienced pain reduction of ≥ 1 point and 1 experienced pain reduction of ≥ 2 points on the NRS.

Figure 5. Pain relief from preintervention to postintervention. Means and SEs of baseline and posttreatment pain (black dots and line, measured via a numerical rating scale [NRS; 0 - 10]), and means of individual study participants (colored dots and lines; $n=18$; 2-sided t test). * $P\leq.01$.

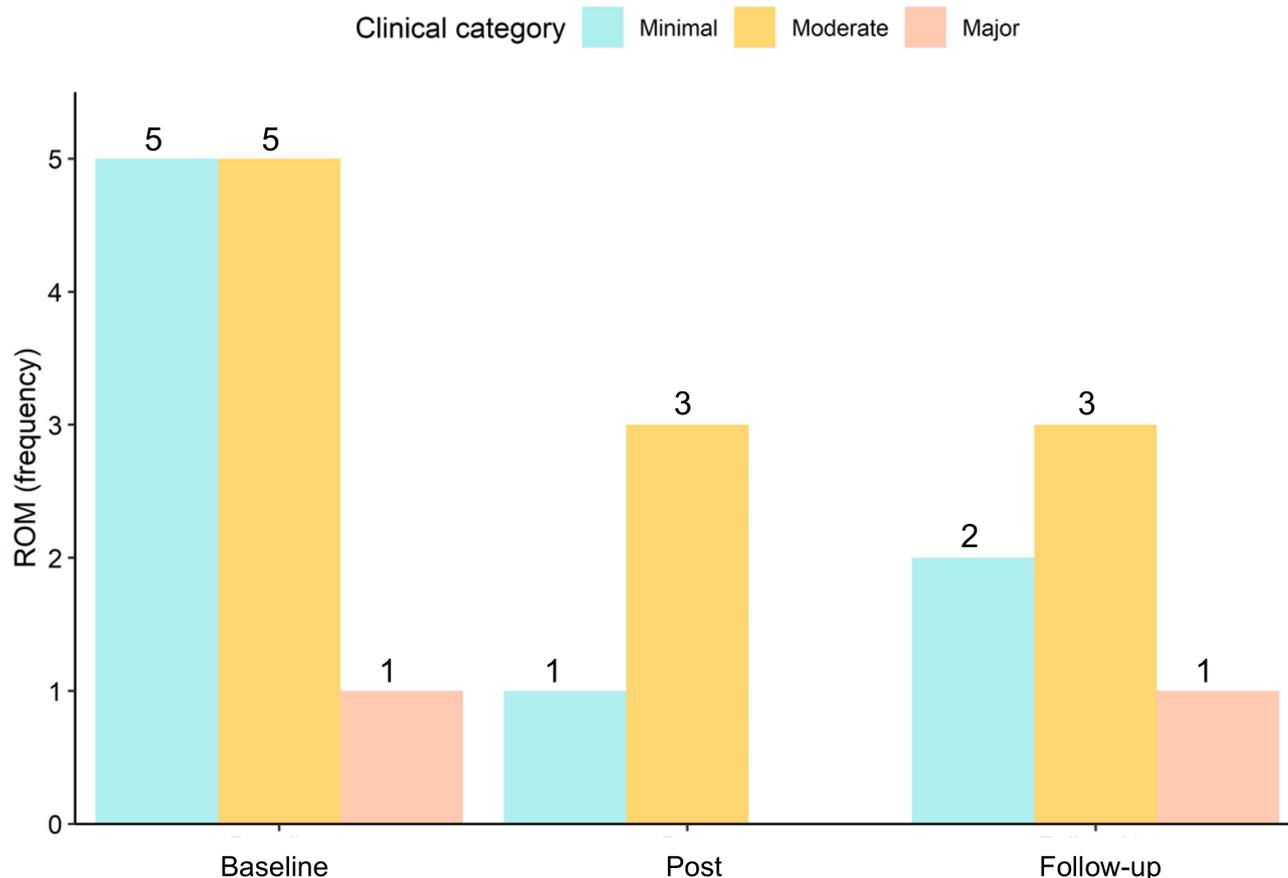


There was no change in pain from immediately before the session (mean 3.94, SD 1.06) to after the session (mean 4.23, SD 1.32), with a mean difference of 0.30 (CI -0.90 to 0.30 ; $t_{17}=-1.04$; $P=.31$; $d=0.15$) (see [Figure S2](#) in [Multimedia Appendix 1](#) for pain ratings before and after each VR session).

Secondary Outcome Measures: ROM and Behavioral and Psychological Variables

There was a significant main effect of time on movement restrictions ($F_{2,32}=10.82$; $P<.001$; $\eta^2_g=0.06$). Movement

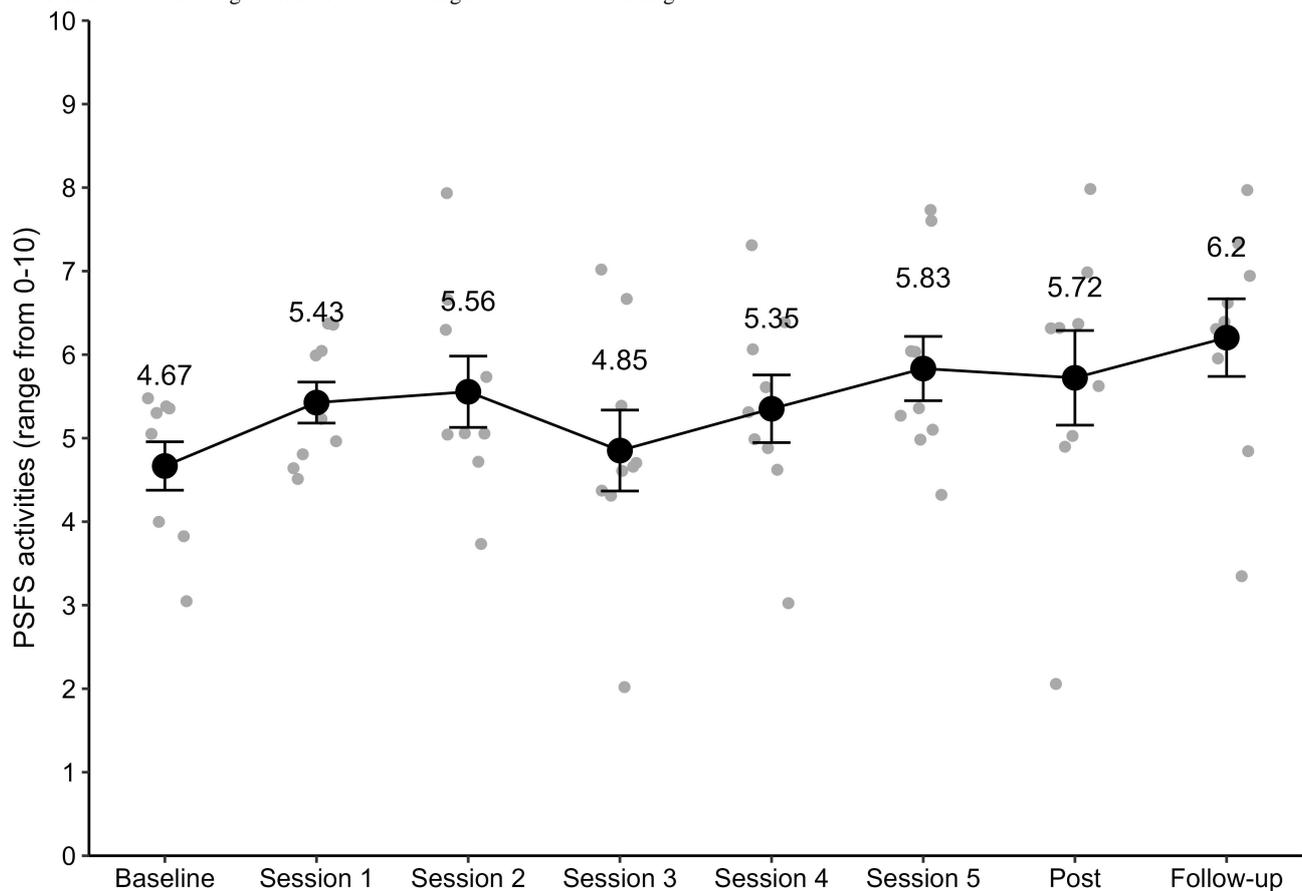
restrictions reduced from preintervention (sum score of 55) to postintervention (sum score of 22) ($t_{17}=3.70$; $P<.001$; CI 0.45-1.66) and from preintervention to the follow-up assessment (sum score of 22) ($t_{17}=3.98$; $P<.001$; $d=0.94$; CI 0.83-2.72). There were no significant differences from postintervention to the follow-up assessment ($t_{17}=1.79$; $P=.09$; $d=0.42$; CI -0.13 to 1.57) ([Figure 6](#)).

Figure 6. Reduced frequency of movement restrictions from preintervention (baseline) to postintervention and follow-up. ROM: range of motion.

Regarding the questionnaires addressing behavioral variables, for the BPS, there was a significant main effect of time ($F_{2,34}=4.53$; $P=.02$; $\eta^2_g=0.04$). BPS scores reduced from preintervention (mean 2.06, SD 2.15) to postintervention (mean 1.56, SD 1.98) ($t_{17}=2.15$; $P=.046$; $d=0.51$; CI 0.01-0.99) and from preintervention to the follow-up assessment (mean 1.17, SD 1.62) ($t_{17}=2.56$; $P=.02$; $d=0.61$; CI 0.17-1.61). There were no significant differences from postintervention to the follow-up assessment ($t_{17}=1.28$; $P=.29$; $d=0.30$; CI -0.25 to 1.03). For the RMDQ, a significant main effect of time was noted ($F_{2,26}=4.73$; $P=.02$; $\eta^2_g=0.08$). RMDQ scores decreased from preintervention (mean 5.0, SD 3.11) to the follow-up assessment (mean 3.57,

SD 3.01) ($t_{13}=2.59$; $P=.02$; $d=0.69$; CI 0.24-2.62) and from postintervention (mean 5.86, SD 3.61) to the follow-up assessment ($t_{13}=2.43$; $P=.03$; $d=0.65$; CI 0.25-4.32). There were no significant differences from preintervention to postintervention ($t_{13}=-1.21$; $P=.25$; $d=0.32$; CI -2.39 to 0.68). Furthermore, there was a trend toward significance for the PSFS ($F_{2,16}=2.68$; $P=.10$; $\eta^2_g=0.21$) (see Figure 7 for courses of functional improvement and Table S2 in Multimedia Appendix 1 for the frequency and improvements of the most common activities listed). Lifting, carrying, walking, and standing (Table S2 in Multimedia Appendix 1) showed clinically significant improvements from preintervention to the follow-up assessment (≥ 2 points) [65].

Figure 7. Improvement of function. Means and SEs (black) and individual scores (gray) over all 3 Patient-Specific Functional Scale (PSFS) activities in each session. Note that higher scores indicate a higher level of functioning.

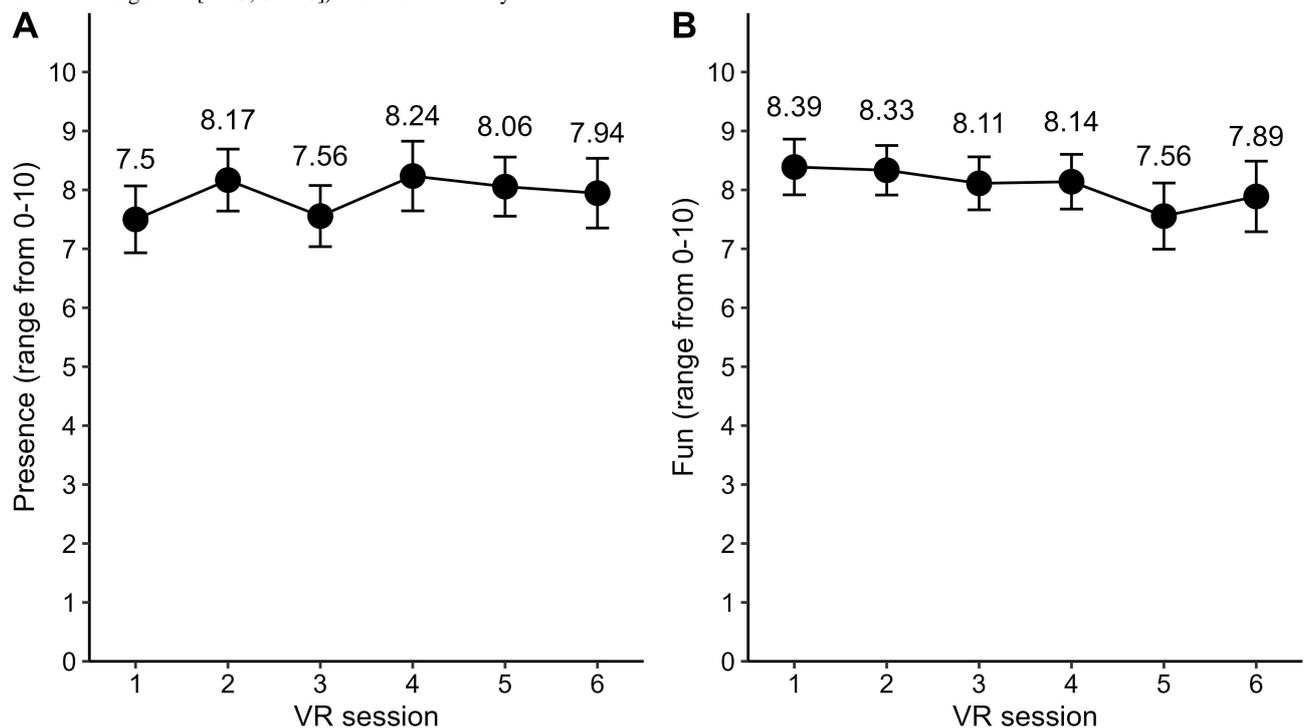


All other questionnaires (behavioral variables: PGIC; psychological variables: FABQ, PROMIS 29, PCS, and TSK) did not yield significant differences (all $P \geq .79$). The results of all questionnaires are presented in Table S3 in [Multimedia Appendix 1](#).

Positive Evaluation of VR Application Measures

Across all VR sessions, the participants reported high feelings of presence (mean 7.88, SD 2.01) and high levels of fun (mean 8.07, SD 1.99) ([Figure 8](#)).

Figure 8. Levels of presence and fun. Means and SEs of the reported feeling of presence (A) and level of fun (B) in each session (both measured via a numerical rating scale [NRS; 0 - 10]). VR: virtual reality.



Discussion

Principal Results

Our novel immersive VR intervention for chronic back pain follows an innovative approach in which users embody a virtual avatar, perform gamified movement exercises, and receive real-time feedback about their performance. We demonstrated the feasibility and tolerability of the intervention as 90% of participants completed the 3-week intervention phase, and there were no adverse events posing a danger to the health of the participants or indicating cybersickness. Importantly, as an initial test of the efficacy of the intervention, we focused on pain as the main outcome and assessed physical functioning and the influence on psychological variables in relation to the fear-avoidance model of chronic pain. In terms of effect size, the study revealed a large and statistically significant pain reduction from the 2-week baseline phase to the 4-week poststudy phase. However, this does not constitute a clinically meaningful reduction in pain over the treatment interval [92-94]. Notably, the intervention improved functioning (ie, performing daily life activities [PSFS and BPS]) and decreased movement restrictions, as demonstrated in ROM and RMDQ scores. However, the mean changes were below commonly cited thresholds for clinical importance. There were no effects on the scales assessing psychological outcomes. The results are promising, but a randomized controlled trial is needed to assess the specificity of the effects. The participants generally liked the intervention, as evident from the high fun rating in each session. Only 1 participant did not like the gamified task and dropped out after session 2. While gamified VR is suitable for the majority of participants, it might not be suited for everyone, as evident from the dropout of another study participant who had the impression that the intervention was not serious enough.

Comparison With Prior Work

There are a few studies investigating the effects of interventions delivered via immersive VR to treat CLBP [10,95,96]. Our study extends the findings of a meta-analysis that showed positive effects of VR on back pain for non- and moderate-immersive virtual environments [10]. The only fully immersive VR study included in the meta-analysis revealed detrimental effects on pain intensity [97]. However, that study included mindfulness and cognitive behavioral therapy-based content, whereas our study investigated the effect of exercise. Using a subgroup analysis, the meta-analysis by Bordeleau et al [10] showed that the association between VR interventions and back pain intensity remained statistically significant only for VR exercises. While exercises are an effective and recommended first-line treatment for CLBP [41,42,98,99], adherence outside the clinic is often low [45]. It was previously demonstrated that gamified tasks in VR increase motivation [47]. Moreover, VR appears to be a particularly suited method to induce a fun experience through gamified tasks and high levels of embodiment, presence, and interactivity [46,52-54], and thereby, exercise in VR could serve as a gateway for future exercise [49]. In this study, we demonstrated the feasibility and tolerability of a gamified intervention delivered via VR, including performance feedback and virtual embodiment. We tested only the initial efficacy of the intervention, and when comparing the effects to previously published studies, it is important to keep the relatively low number of exercise sessions in mind (6 sessions).

Study participants with CLBP rated their pain daily before the intervention for 2 weeks and daily after the intervention for 4 weeks [97,100]. Only considering the same number of sessions as in our study, the average reduction in pain intensity was in a similar range in previous VR interventions [17,100,101].

Although the effect in our study was large in terms of effect size, it did not constitute a clinically meaningful effect in terms of an absolute pain reduction for most of the study participants [92]. This is likely due to several reasons. First, the mean pain ratings of the study participants during the baseline period were only slightly above the threshold for inclusion (≥ 4), indicating low to medium pain levels. Previous studies and a current meta-analysis demonstrated larger effects for higher baseline pain levels [8,17,102]. Second, the duration of the intervention was relatively short. Bordeleau et al [10] demonstrated the beneficial effects of VR when more than 12 sessions were performed. A recent study with persons with CLBP who took part in a behavioral skills-based VR program demonstrated a clinically meaningful pain reduction during home-based training [17]. That study was substantially longer than our study (56 daily sessions over the course of 8 weeks) and included different elements (pain education, relaxation/interoception, mindful escape via 360° videos, interactive games for pain distraction, and dynamic breathing).

In addition to preintervention and postintervention pain ratings (baseline vs follow-up phase), pain ratings before and after each VR session (assessed in VR) served to evaluate the immediate effects of the intervention and to document possible pain exacerbation caused by movements. We found no immediate changes in pain ratings in response to the intervention. Further, none of the participants ended the intervention due to pain induced by the required movements. This finding is in line with a usability study that gamified a lumbar flexion task in VR for persons with CLBP [24]. This demonstrates that the VR games distracted from pain otherwise caused by movements and encouraged participants to perform movements in VR that might be avoided in the real world due to fear of pain.

Partially based on the fear-avoidance model [27,28], our intervention was designed as a graded exposure therapy in VR, that is, ROM for the individual movement exercises was gradually increased (3 difficulty levels per exercise). In line with the results by Bordeleau et al [10], the VR intervention yielded measurable improvements in back and task-specific functioning, for example, in daily activities that were difficult or impossible to perform due to back pain, as measured in the PSFS. Participants in this study indicated functional impairments in typical activities [65] like walking, standing, bending, or carrying. Notably, lifting, carrying, walking, and standing showed clinically significant improvements [65], and there were gradual improvements over sessions. Functional tests of the lumbar spine (BPS) were conducted and revealed significant improvements over the course of the study. With the RMDQ, back pain was evaluated in the context of specific activities, and pain reductions over the course of the study were identified. Considering the low initial degree of disability, a difference of 1 point can be considered as an improvement [68]. The fact that the beneficial effects of functional measurements were not revealed immediately after the intervention but at follow-up can probably be attributed to persons with CLBP continuing to exercise after the intervention (as encouraged). This suggests that the VR intervention might serve as a gateway to continued beneficial exercises after the intervention. However, owing to

the lack of a control group, the specificity of the results remains to be demonstrated.

While the intervention was also designed to have beneficial effects on kinesiophobia and fear-avoidance beliefs, effects on these outcomes, as well as on pain catastrophizing, could not be revealed due to the lack of clinically meaningful scores of the study participants at the beginning of the study. Specifically, all study participants with CLBP had low scores on the TSK [103]. Likewise, none of the participants had scores above 30 on the PCS, which would indicate a clinically significant level of pain catastrophizing [104]. For fear-avoidance beliefs assessed with the FABQ, none of the participants had scores above values for fear-avoidance beliefs related to work that are considered high, and for the physical activity subscale, only 1 participant had a high score (>15) that dropped to 3 (postintervention evaluation) [105,106]. However, another study discussed that, besides psychological variables such as catastrophic thinking, physiological adaptations related to the avoidance of certain movements could also play a crucial role in the maintenance of pain, emphasizing the role of treating functional impairments [107,108].

Overall, it is encouraging that there were only few and minor symptoms assessed with the SSQ, with most relating to sweating (probably due to exercising, especially in summertime, and wearing an HMD), followed by eye strain and difficulty focusing with the HMD. Vertigo was not present after any of the sessions, and stomach awareness and nausea were only reported by a single participant after 1 session. Therefore, the intervention is not likely to induce symptoms of cybersickness.

Limitations and Outlook

Most importantly, we demonstrated the feasibility and tolerability of the novel intervention, and tests of the initial clinical efficacy are promising. However, it is important to stress that we conducted a single-arm study, and to demonstrate the specificity of the effects, a controlled study is necessary. The generalizability of the results is further limited due to the rather small sample size. The revealed effects did not exceed previously defined thresholds of clinical efficacy. The duration of the intervention was shorter compared to conventional physiotherapy for CLBP [109]. Hence, a longer duration of the intervention is desirable to increase its effects on clinical outcomes. The intervention is partially based on the fear-avoidance model of chronic pain, and while effective in reducing pain and behavioral outcomes, effects on cognition could not be revealed due to low scores on catastrophizing, fear-avoidance beliefs, and kinesiophobia of the study participants; hence, a larger or more selective sample is needed to shed further light on the potential of the intervention to affect these outcomes [107].

A fully immersive VR system has several advantages in terms of delivering a gamified environment, and higher levels of immersion have positive effects on pain through increased distraction [9,14,46]. Future studies could also explore the potential of VR to not only change the visual appearance of the body of the participants (as done in our study) but also manipulate the visual feedback in relation to movements, and investigate how overstated or understated movements displayed

(compared to actual performed movements) affect exercise performance and effectiveness. In a previous study, an overstated or understated neck rotation in relation to neck pain was investigated [110]. Further, using a virtual cycling exercise, studies investigated how a match or mismatch between actual pedaling resistance and the virtual environment (higher/lower hill gradient) affects exercise performance [111] and perception of exercise effort [112]. Despite its benefits, immersive VR is also associated with a higher risk of collision with objects in the real environment because the HMD blocks visual input from the real environment. Given technological advances, augmented reality could be a promising alternative that utilizes the benefits of VR, such as gamification, while ensuring safe use (eg, showing obstacles in the real world while performing exercises) [113]. The present camera-based body tracking system allowed unrestricted movements of the participants and provided adequate accuracy for real-time feedback with the hologram. In order for it to work, a controlled environment with adequate lighting and high-contrast clothes were needed to properly detect the body. For a home environment or if more precise tracking

is desired, tracking based on sensors attached to the body might be needed. Sensors attached to the body might also be preferable if biofeedback analysis is the relevant outcome measurement, as implemented previously [23].

Conclusion

We demonstrated the feasibility and safety of a novel VR intervention for nonspecific CLBP. It involved embodiment and gamified movement exercises in immersive VR with real-time feedback about performance provided through a hologram overlaying the avatar of the user. Almost all participants liked the intervention (high fun ratings), which is important as it can motivate patients to exercise and potentially increase exercise adherence in the long term. Despite its short duration of 6 sessions over the course of 3 weeks, positive effects on pain, physical functioning, and daily activities were noted. However, these effects fall short of reaching previously defined thresholds of clinical importance, and the study design does not allow to conclude on the specificity of these effects. Hence, a longer intervention is desirable, and a controlled study is needed to test its clinical effects.

Acknowledgments

The authors would like to thank Kristina Ruisz and Gina Lehmann for their help in data collection. The authors would also like to thank Fabian Rucker for his help in creating the software and providing technical support. No generative artificial intelligence tool was used to generate manuscript content.

Funding

The study received funding from the German Federal Ministry for Education and Research (BMBF; 16GDI106). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

Data and analysis scripts are available on OSF [114].

Authors' Contributions

This study was designed by IN, SL, IK, and HLR. The intervention was designed by SL, MG, IN, IK, YN, RZ, and HLR. The software was implemented by MG and videoreality GmbH. The experiments were performed by IN and SL. The data were analyzed by IN, SL, and IK, and the results were critically examined by all authors. IN and IK had a primary role in preparing the manuscript, which was edited by SL, HLR, RZ, and YN. All authors have approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional material to support the study.

[PDF File, 338 KB - [games_v14i1e81051_app1.pdf](#)]

Multimedia Appendix 2

Results of the Little missing completely at random test for each outcome variable.

[PDF File, 143 KB - [games_v14i1e81051_app2.pdf](#)]

References

1. Ferreira ML, de Luca K, Haile LM, GBD 2021 Low Back Pain Collaborators. Global, regional, and national burden of low back pain, 1990–2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of

- Disease Study 2021. *Lancet Rheumatol* 2023 Jun;5(6):e316-e329. [doi: [10.1016/S2665-9913\(23\)00098-X](https://doi.org/10.1016/S2665-9913(23)00098-X)] [Medline: [37273833](https://pubmed.ncbi.nlm.nih.gov/37273833/)]
2. Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. *Eur J Pain* 2006 May;10(4):287-333. [doi: [10.1016/j.ejpain.2005.06.009](https://doi.org/10.1016/j.ejpain.2005.06.009)] [Medline: [16095934](https://pubmed.ncbi.nlm.nih.gov/16095934/)]
 3. Hooten WM. Chronic pain and mental health disorders: shared neural mechanisms, epidemiology, and treatment. *Mayo Clin Proc* 2016 Jul;91(7):955-970. [doi: [10.1016/j.mayocp.2016.04.029](https://doi.org/10.1016/j.mayocp.2016.04.029)] [Medline: [27344405](https://pubmed.ncbi.nlm.nih.gov/27344405/)]
 4. Niv D, Kreitler S. Pain and quality of life. *Pain Pract* 2001 Jun;1(2):150-161. [doi: [10.1046/j.1533-2500.2001.01016.x](https://doi.org/10.1046/j.1533-2500.2001.01016.x)] [Medline: [17129291](https://pubmed.ncbi.nlm.nih.gov/17129291/)]
 5. Stubhaug A, Hansen JL, Hallberg S, Gustavsson A, Eggen AE, Nielsen CS. The costs of chronic pain-long-term estimates. *Eur J Pain* 2024 Jul;28(6):960-977. [doi: [10.1002/ejp.2234](https://doi.org/10.1002/ejp.2234)] [Medline: [38214661](https://pubmed.ncbi.nlm.nih.gov/38214661/)]
 6. Dowell D, Ragan KR, Jones CM, Baldwin GT, Chou R. CDC clinical practice guideline for prescribing opioids for pain - United States, 2022. *MMWR Recomm Rep* 2022 Nov 4;71(3):1-95. [doi: [10.15585/mmwr.r7103a1](https://doi.org/10.15585/mmwr.r7103a1)] [Medline: [36327391](https://pubmed.ncbi.nlm.nih.gov/36327391/)]
 7. Maser D, Müßgens D, Kleine-Borgmann J, et al. Setting the stage for pain relief: how treatment setting impacts interdisciplinary multimodal pain treatment for patients with chronic back pain. *Pain* 2024 Dec 1;165(12):2909-2919. [doi: [10.1097/j.pain.0000000000003318](https://doi.org/10.1097/j.pain.0000000000003318)] [Medline: [39051839](https://pubmed.ncbi.nlm.nih.gov/39051839/)]
 8. Lier EJ, de Vries M, Steggink EM, Ten Broek RPG, van Goor H. Effect modifiers of virtual reality in pain management: a systematic review and meta-regression analysis. *Pain* 2023 Aug 1;164(8):1658-1665. [doi: [10.1097/j.pain.0000000000002883](https://doi.org/10.1097/j.pain.0000000000002883)] [Medline: [36943251](https://pubmed.ncbi.nlm.nih.gov/36943251/)]
 9. Trost Z, France C, Anam M, Shum C. Virtual reality approaches to pain: toward a state of the science. *Pain* 2021 Feb 1;162(2):325-331. [doi: [10.1097/j.pain.0000000000002060](https://doi.org/10.1097/j.pain.0000000000002060)] [Medline: [32868750](https://pubmed.ncbi.nlm.nih.gov/32868750/)]
 10. Bordeleau M, Stamenkovic A, Tardif PA, Thomas J. The use of virtual reality in back pain rehabilitation: a systematic review and meta-analysis. *J Pain* 2022 Feb;23(2):175-195. [doi: [10.1016/j.jpain.2021.08.001](https://doi.org/10.1016/j.jpain.2021.08.001)] [Medline: [34425250](https://pubmed.ncbi.nlm.nih.gov/34425250/)]
 11. 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](https://doi.org/10.1097/00002508-200009000-00010)] [Medline: [11014398](https://pubmed.ncbi.nlm.nih.gov/11014398/)]
 12. Kenney MP, Milling LS. The effectiveness of virtual reality distraction for reducing pain: a meta-analysis. *Psychol Conscious Theory Res Pract* 2016;3(3):199-210. [doi: [10.1037/cns0000084](https://doi.org/10.1037/cns0000084)]
 13. Sharar SR, Alamdari A, Hoffer C, Hoffman HG, Jensen MP, Patterson DR. Circumplex model of affect: a measure of pleasure and arousal during virtual reality distraction analgesia. *Games Health J* 2016 Jun;5(3):197-202. [doi: [10.1089/g4h.2015.0046](https://doi.org/10.1089/g4h.2015.0046)] [Medline: [27171578](https://pubmed.ncbi.nlm.nih.gov/27171578/)]
 14. Triberti S, Repetto C, Riva G. Psychological factors influencing the effectiveness of virtual reality-based analgesia: a systematic review. *Cyberpsychol Behav Soc Netw* 2014 Jun;17(6):335-345. [doi: [10.1089/cyber.2014.0054](https://doi.org/10.1089/cyber.2014.0054)] [Medline: [24892195](https://pubmed.ncbi.nlm.nih.gov/24892195/)]
 15. Gupta A, Scott K, Dukewich M. Innovative technology using virtual reality in the treatment of pain: does it reduce pain via distraction, or is there more to it? *Pain Med* 2018 Jan 1;19(1):151-159. [doi: [10.1093/pm/pnx109](https://doi.org/10.1093/pm/pnx109)] [Medline: [29025113](https://pubmed.ncbi.nlm.nih.gov/29025113/)]
 16. Maddox T, Garcia H, Ffrench K, et al. In-home virtual reality program for chronic low back pain: durability of a randomized, placebo-controlled clinical trial to 18 months post-treatment. *Reg Anesth Pain Med* 2024 May 7;49(5):373-375. [doi: [10.1136/rapm-2022-104093](https://doi.org/10.1136/rapm-2022-104093)] [Medline: [36427904](https://pubmed.ncbi.nlm.nih.gov/36427904/)]
 17. Garcia LM, Birkhead BJ, Krishnamurthy P, et al. An 8-week self-administered at-home behavioral skills-based virtual reality program for chronic low back pain: double-blind, randomized, placebo-controlled trial conducted during COVID-19. *J Med Internet Res* 2021 Feb 22;23(2):e26292. [doi: [10.2196/26292](https://doi.org/10.2196/26292)] [Medline: [33484240](https://pubmed.ncbi.nlm.nih.gov/33484240/)]
 18. Solcà M, Ronchi R, Bello-Ruiz J, et al. Heartbeat-enhanced immersive virtual reality to treat complex regional pain syndrome. *Neurology* 2018 Jul 31;91(5):e479-e489. [doi: [10.1212/WNL.0000000000005905](https://doi.org/10.1212/WNL.0000000000005905)] [Medline: [29980635](https://pubmed.ncbi.nlm.nih.gov/29980635/)]
 19. Matamala-Gomez M, Donegan T, Bottiroli S, Sandrini G, Sanchez-Vives MV, Tassorelli C. Immersive virtual reality and virtual embodiment for pain relief. *Front Hum Neurosci* 2019;13:279. [doi: [10.3389/fnhum.2019.00279](https://doi.org/10.3389/fnhum.2019.00279)] [Medline: [31551731](https://pubmed.ncbi.nlm.nih.gov/31551731/)]
 20. Matamala-Gomez M, Diaz Gonzalez AM, Slater M, Sanchez-Vives MV. Decreasing pain ratings in chronic arm pain through changing a virtual body: different strategies for different pain types. *J Pain* 2019 Jun;20(6):685-697. [doi: [10.1016/j.jpain.2018.12.001](https://doi.org/10.1016/j.jpain.2018.12.001)] [Medline: [30562584](https://pubmed.ncbi.nlm.nih.gov/30562584/)]
 21. Pamment J, Aspell JE. Putting pain out of mind with an “out of body” illusion. *Eur J Pain* 2017 Feb;21(2):334-342. [doi: [10.1002/ejp.927](https://doi.org/10.1002/ejp.927)] [Medline: [27509229](https://pubmed.ncbi.nlm.nih.gov/27509229/)]
 22. Alemanno F, Houdayer E, Emedoli D, et al. Efficacy of virtual reality to reduce chronic low back pain: proof-of-concept of a non-pharmacological approach on pain, quality of life, neuropsychological and functional outcome. *PLoS One* 2019;14(5):e0216858. [doi: [10.1371/journal.pone.0216858](https://doi.org/10.1371/journal.pone.0216858)] [Medline: [31120892](https://pubmed.ncbi.nlm.nih.gov/31120892/)]
 23. France CR, Thomas JS. Virtual immersive gaming to optimize recovery (VIGOR) in low back pain: a phase II randomized controlled trial. *Contemp Clin Trials* 2018 Jun;69:83-91. [doi: [10.1016/j.cct.2018.05.001](https://doi.org/10.1016/j.cct.2018.05.001)] [Medline: [29730393](https://pubmed.ncbi.nlm.nih.gov/29730393/)]
 24. Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski S. Feasibility and safety of a virtual reality dodgeball intervention for chronic low back pain: a randomized clinical trial. *J Pain* 2016 Dec;17(12):1302-1317. [doi: [10.1016/j.jpain.2016.08.011](https://doi.org/10.1016/j.jpain.2016.08.011)] [Medline: [27616607](https://pubmed.ncbi.nlm.nih.gov/27616607/)]

25. Wren AA, Neiman N, Caruso TJ, et al. Mindfulness-based virtual reality intervention for children and young adults with inflammatory bowel disease: a pilot feasibility and acceptability study. *Children (Basel)* 2021 May 5;8(5):368. [doi: [10.3390/children8050368](https://doi.org/10.3390/children8050368)] [Medline: [34063034](https://pubmed.ncbi.nlm.nih.gov/34063034/)]
26. Parsons TD, Trost Z. Virtual reality graded exposure therapy as treatment for pain-related fear and disability in chronic pain. In: Ma M, Jain L, Anderson P, editors. *Virtual, Augmented Reality and Serious Games for Healthcare Intelligent Systems Reference Library*: Springer; 2014, Vol. 68:523-546. [doi: [10.1007/978-3-642-54816-1_25](https://doi.org/10.1007/978-3-642-54816-1_25)]
27. Vlaeyen JWS, Crombez G, Linton SJ. The fear-avoidance model of pain. *Pain* 2016 Aug;157(8):1588-1589. [doi: [10.1097/j.pain.0000000000000574](https://doi.org/10.1097/j.pain.0000000000000574)] [Medline: [27428892](https://pubmed.ncbi.nlm.nih.gov/27428892/)]
28. Vlaeyen JWS, Linton SJ. Fear-avoidance model of chronic musculoskeletal pain: 12 years on. *Pain* 2012 Jun;153(6):1144-1147. [doi: [10.1016/j.pain.2011.12.009](https://doi.org/10.1016/j.pain.2011.12.009)] [Medline: [22321917](https://pubmed.ncbi.nlm.nih.gov/22321917/)]
29. Swinkels-Meewisse IEJ, Roelofs J, Oostendorp RAB, Verbeek ALM, Vlaeyen JWS. Acute low back pain: pain-related fear and pain catastrophizing influence physical performance and perceived disability. *Pain* 2006 Jan;120(1-2):36-43. [doi: [10.1016/j.pain.2005.10.005](https://doi.org/10.1016/j.pain.2005.10.005)] [Medline: [16359797](https://pubmed.ncbi.nlm.nih.gov/16359797/)]
30. Wertli MM, Eugster R, Held U, Steurer J, Kofmehl R, Weiser S. Catastrophizing—a prognostic factor for outcome in patients with low back pain: a systematic review. *Spine J* 2014 Nov;14(11):2639-2657. [doi: [10.1016/j.spinee.2014.03.003](https://doi.org/10.1016/j.spinee.2014.03.003)] [Medline: [24607845](https://pubmed.ncbi.nlm.nih.gov/24607845/)]
31. Haythornthwaite JA, Campbell CM, Edwards RR. When thinking about pain contributes to suffering: the example of pain catastrophizing. *Pain* 2024 Nov 1;165(11S):S68-S75. [doi: [10.1097/j.pain.0000000000003372](https://doi.org/10.1097/j.pain.0000000000003372)] [Medline: [39560417](https://pubmed.ncbi.nlm.nih.gov/39560417/)]
32. Nicholas MK, Linton SJ, Watson PJ, Main CJ. Early identification and management of psychological risk factors (“yellow flags”) in patients with low back pain: a reappraisal. *Phys Ther* 2011 May;91(5):737-753. [doi: [10.2522/ptj.20100224](https://doi.org/10.2522/ptj.20100224)] [Medline: [21451099](https://pubmed.ncbi.nlm.nih.gov/21451099/)]
33. Goubert L, Crombez G, Van Damme S. The role of neuroticism, pain catastrophizing and pain-related fear in vigilance to pain: a structural equations approach. *Pain* 2004 Feb;107(3):234-241. [doi: [10.1016/j.pain.2003.11.005](https://doi.org/10.1016/j.pain.2003.11.005)] [Medline: [14736586](https://pubmed.ncbi.nlm.nih.gov/14736586/)]
34. Picavet HSJ, Vlaeyen JWS, Schouten JS. Pain catastrophizing and kinesiophobia: predictors of chronic low back pain. *Am J Epidemiol* 2002 Dec 1;156(11):1028-1034. [doi: [10.1093/aje/kwf136](https://doi.org/10.1093/aje/kwf136)] [Medline: [12446259](https://pubmed.ncbi.nlm.nih.gov/12446259/)]
35. Shigetoh H. Hypervigilance to pain affects activities of daily living: an examination using the Japanese version of the pain vigilance awareness questionnaire. *J Phys Ther Sci* 2017 Dec;29(12):2094-2096. [doi: [10.1589/jpts.29.2094](https://doi.org/10.1589/jpts.29.2094)] [Medline: [29643581](https://pubmed.ncbi.nlm.nih.gov/29643581/)]
36. Sullivan MJL, Lynch ME, Clark AJ. Dimensions of catastrophic thinking associated with pain experience and disability in patients with neuropathic pain conditions. *Pain* 2005 Feb;113(3):310-315. [doi: [10.1016/j.pain.2004.11.003](https://doi.org/10.1016/j.pain.2004.11.003)] [Medline: [15661438](https://pubmed.ncbi.nlm.nih.gov/15661438/)]
37. Roelofs J, Peters ML, Patijn J, Schouten EGW, Vlaeyen JWS. Electronic diary assessment of pain-related fear, attention to pain, and pain intensity in chronic low back pain patients. *Pain* 2004 Dec;112(3):335-342. [doi: [10.1016/j.pain.2004.09.016](https://doi.org/10.1016/j.pain.2004.09.016)] [Medline: [15561389](https://pubmed.ncbi.nlm.nih.gov/15561389/)]
38. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? *JAMA* 2010 Apr 7;303(13):1295-1302. [doi: [10.1001/jama.2010.344](https://doi.org/10.1001/jama.2010.344)] [Medline: [20371789](https://pubmed.ncbi.nlm.nih.gov/20371789/)]
39. George SZ, Beneciuk JM. Psychological predictors of recovery from low back pain: a prospective study. *BMC Musculoskelet Disord* 2015 Mar 7;16(1):49. [doi: [10.1186/s12891-015-0509-2](https://doi.org/10.1186/s12891-015-0509-2)] [Medline: [25849159](https://pubmed.ncbi.nlm.nih.gov/25849159/)]
40. de Jong JR, Vlaeyen JWS, van Eijsden M, Loo C, Onghena P. Reduction of pain-related fear and increased function and participation in work-related upper extremity pain (WRUEP): effects of exposure in vivo. *Pain* 2012 Oct;153(10):2109-2118. [doi: [10.1016/j.pain.2012.07.001](https://doi.org/10.1016/j.pain.2012.07.001)] [Medline: [22902198](https://pubmed.ncbi.nlm.nih.gov/22902198/)]
41. Airaksinen O, Brox JI, Cedraschi C, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. *Eur Spine J* 2006 Mar;15 Suppl 2(Suppl 2):S192-S300. [doi: [10.1007/s00586-006-1072-1](https://doi.org/10.1007/s00586-006-1072-1)] [Medline: [16550448](https://pubmed.ncbi.nlm.nih.gov/16550448/)]
42. Longtin C, Décarry S, Cook CE, Tousignant-Laflamme Y. What does it take to facilitate the integration of clinical practice guidelines for the management of low back pain into practice? Part 1: a synthesis of recommendation. *Pain Pract* 2021 Nov;21(8):943-954. [doi: [10.1111/papr.13033](https://doi.org/10.1111/papr.13033)] [Medline: [33998769](https://pubmed.ncbi.nlm.nih.gov/33998769/)]
43. Geneen LJ, Moore RA, Clarke C, Martin D, Colvin LA, Smith BH. Physical activity and exercise for chronic pain in adults: an overview of Cochrane reviews. *Cochrane Database Syst Rev* 2017 Apr 24;4(4):CD011279. [doi: [10.1002/14651858.CD011279.pub3](https://doi.org/10.1002/14651858.CD011279.pub3)] [Medline: [28436583](https://pubmed.ncbi.nlm.nih.gov/28436583/)]
44. Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in chronic low back pain. *Ann Intern Med* 2005 May 3;142(9):776-785. [doi: [10.7326/0003-4819-142-9-200505030-00014](https://doi.org/10.7326/0003-4819-142-9-200505030-00014)] [Medline: [15867410](https://pubmed.ncbi.nlm.nih.gov/15867410/)]
45. Martin LR, Williams SL, Haskard KB, Dimatteo MR. The challenge of patient adherence. *Ther Clin Risk Manag* 2005 Sep;1(3):189-199. [doi: [10.2147/tcrm.s12160382](https://doi.org/10.2147/tcrm.s12160382)] [Medline: [18360559](https://pubmed.ncbi.nlm.nih.gov/18360559/)]
46. Tack C. Virtual reality and chronic low back pain. *Disabil Rehabil Assist Technol* 2021 Aug;16(6):637-645. [doi: [10.1080/17483107.2019.1688399](https://doi.org/10.1080/17483107.2019.1688399)] [Medline: [31746250](https://pubmed.ncbi.nlm.nih.gov/31746250/)]
47. Winter C, Kern F, Gall D, Latoschik ME, Pauli P, Käthner I. Immersive virtual reality during gait rehabilitation increases walking speed and motivation: a usability evaluation with healthy participants and patients with multiple sclerosis and stroke. *J Neuroeng Rehabil* 2021 Apr 22;18(1):68. [doi: [10.1186/s12984-021-00848-w](https://doi.org/10.1186/s12984-021-00848-w)] [Medline: [33888148](https://pubmed.ncbi.nlm.nih.gov/33888148/)]

48. Lindner S, Latoschik ME, Rittner H. Use of virtual reality as a component of acute and chronic pain treatment [Article in German]. *Anesthesiol Intensivmed Notfallmed Schmerzther* 2020 Sep;55(9):549-561. [doi: [10.1055/a-1022-3038](https://doi.org/10.1055/a-1022-3038)] [Medline: [32916738](https://pubmed.ncbi.nlm.nih.gov/32916738/)]
49. Rhodes RE, Kates A. Can the affective response to exercise predict future motives and physical activity behavior? A systematic review of published evidence. *Ann Behav Med* 2015 Oct;49(5):715-731. [doi: [10.1007/s12160-015-9704-5](https://doi.org/10.1007/s12160-015-9704-5)] [Medline: [25921307](https://pubmed.ncbi.nlm.nih.gov/25921307/)]
50. Williams DM, Dunsiger S, Jennings EG, Marcus BH. Does affective valence during and immediately following a 10-min walk predict concurrent and future physical activity? *Ann Behav Med* 2012 Aug;44(1):43-51. [doi: [10.1007/s12160-012-9362-9](https://doi.org/10.1007/s12160-012-9362-9)] [Medline: [22532005](https://pubmed.ncbi.nlm.nih.gov/22532005/)]
51. Liao Y, Chou CP, Huh J, Leventhal A, Dunton G. Associations of affective responses during free-living physical activity and future physical activity levels: an ecological momentary assessment study. *Int J Behav Med* 2017 Aug;24(4):513-519. [doi: [10.1007/s12529-016-9626-z](https://doi.org/10.1007/s12529-016-9626-z)] [Medline: [28008556](https://pubmed.ncbi.nlm.nih.gov/28008556/)]
52. Bird JM, Karageorghis CI, Baker SJ, Brookes DA, Nowicky AV. Ready Exerciser One: effects of music and virtual reality on cycle ergometer exercise. *Br J Health Psychol* 2021 Feb;26(1):15-32. [doi: [10.1111/bjhp.12445](https://doi.org/10.1111/bjhp.12445)] [Medline: [32538512](https://pubmed.ncbi.nlm.nih.gov/32538512/)]
53. Bird JM, Karageorghis CI, Jones L, Harris DJ, Alharbi M, Vine SJ. Beyond Rubik: the Embodiment-Presence-Interactivity Cube applied to exercise. *Psychol Sport Exerc* 2024 Sep;74:102684. [doi: [10.1016/j.psychsport.2024.102684](https://doi.org/10.1016/j.psychsport.2024.102684)] [Medline: [38830499](https://pubmed.ncbi.nlm.nih.gov/38830499/)]
54. Jones L, Zenko Z. A systematic narrative review of extrinsic strategies to improve affective responses to exercise. *Front Sports Act Living* 2023;5(5):1186986. [doi: [10.3389/fspor.2023.1186986](https://doi.org/10.3389/fspor.2023.1186986)] [Medline: [37496882](https://pubmed.ncbi.nlm.nih.gov/37496882/)]
55. Darlow B, Dean S, Perry M, Mathieson F, Baxter GD, Dowell A. Easy to harm, hard to heal: patient views about the back. *Spine (Phila Pa 1976)* 2015 Jun 1;40(11):842-850. [doi: [10.1097/BRS.0000000000000901](https://doi.org/10.1097/BRS.0000000000000901)] [Medline: [25811262](https://pubmed.ncbi.nlm.nih.gov/25811262/)]
56. Harvie DS, Rio E, Smith RT, Olthof N, Coppieters MW. Virtual reality body image training for chronic low back pain: a single case report. *Front Virtual Real* 2020;1. [doi: [10.3389/frvir.2020.00013](https://doi.org/10.3389/frvir.2020.00013)]
57. Lotze M, Moseley GL. Role of distorted body image in pain. *Curr Rheumatol Rep* 2007 Dec;9(6):488-496. [doi: [10.1007/s11926-007-0079-x](https://doi.org/10.1007/s11926-007-0079-x)] [Medline: [18177603](https://pubmed.ncbi.nlm.nih.gov/18177603/)]
58. Nishigami T, Wand BM, Newport R, et al. Embodying the illusion of a strong, fit back in people with chronic low back pain. a pilot proof-of-concept study. *Musculoskelet Sci Pract* 2019 Feb;39:178-183. [doi: [10.1016/j.msksp.2018.07.002](https://doi.org/10.1016/j.msksp.2018.07.002)] [Medline: [30049618](https://pubmed.ncbi.nlm.nih.gov/30049618/)]
59. Birkhead B, Khalil C, Liu X, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative Study. *JMIR Ment Health* 2019 Jan 31;6(1):e11973. [doi: [10.2196/11973](https://doi.org/10.2196/11973)] [Medline: [30702436](https://pubmed.ncbi.nlm.nih.gov/30702436/)]
60. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007 May;39(2):175-191. [doi: [10.3758/bf03193146](https://doi.org/10.3758/bf03193146)] [Medline: [17695343](https://pubmed.ncbi.nlm.nih.gov/17695343/)]
61. 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](https://doi.org/10.1207/s15327108ijap0303_3)]
62. McKenzie R, May S. *The Lumbar Spine Mechanical Diagnosis & Therapy: Spinal Publications; 2003, Vol. 2.*
63. Garcia AN, Costa LDC, de Souza FS, et al. Reliability of the mechanical diagnosis and therapy system in patients with spinal pain: a systematic review. *J Orthop Sports Phys Ther* 2018 Dec;48(12):923-933. [doi: [10.2519/jospt.2018.7876](https://doi.org/10.2519/jospt.2018.7876)] [Medline: [29932871](https://pubmed.ncbi.nlm.nih.gov/29932871/)]
64. Stratford P, Gill C, Westaway M, Binkley J. Assessing disability and change on individual patients: a report of a patient specific measure. *Physiotherapy Canada* 1995 Oct;47(4):258-263. [doi: [10.3138/ptc.47.4.258](https://doi.org/10.3138/ptc.47.4.258)]
65. Heldmann P, Schöttker-Königer T, Schäfer A. Cross-cultural adaption and validity of the "Patient Specific Functional Scale". *Int J Health Profess* 2015;2(1):73-82. [doi: [10.1515/ijhp-2015-0002](https://doi.org/10.1515/ijhp-2015-0002)]
66. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. *Spine (Phila Pa 1976)* 1983 Mar;8(2):141-144. [doi: [10.1097/00007632-198303000-00004](https://doi.org/10.1097/00007632-198303000-00004)] [Medline: [6222486](https://pubmed.ncbi.nlm.nih.gov/6222486/)]
67. Wiesinger GF, Nuhr M, Quittan M, Ebenbichler G, Wöfl G, Fialka-Moser V. Cross-cultural adaptation of the Roland-Morris questionnaire for German-speaking patients with low back pain. *Spine (Phila Pa 1976)* 1999 Jun 1;24(11):1099-1103. [doi: [10.1097/00007632-199906010-00009](https://doi.org/10.1097/00007632-199906010-00009)] [Medline: [10361659](https://pubmed.ncbi.nlm.nih.gov/10361659/)]
68. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa 1976)* 2000 Dec 15;25(24):3115-3124. [doi: [10.1097/00007632-200012150-00006](https://doi.org/10.1097/00007632-200012150-00006)] [Medline: [11124727](https://pubmed.ncbi.nlm.nih.gov/11124727/)]
69. Hurst H, Bolton J. Assessing the clinical significance of change scores recorded on subjective outcome measures. *J Manipulative Physiol Ther* 2004 Jan;27(1):26-35. [doi: [10.1016/j.jmpt.2003.11.003](https://doi.org/10.1016/j.jmpt.2003.11.003)] [Medline: [14739871](https://pubmed.ncbi.nlm.nih.gov/14739871/)]
70. Strand LI, Moe-Nilssen R, Ljunggren AE. Back Performance Scale for the assessment of mobility-related activities in people with back pain. *Phys Ther* 2002 Dec;82(12):1213-1223. [doi: [10.1093/ptj/82.12.1213](https://doi.org/10.1093/ptj/82.12.1213)] [Medline: [12444880](https://pubmed.ncbi.nlm.nih.gov/12444880/)]
71. Verra M. Funktioneller Leistungstest für den Rücken [Article in German]. *Physiopraxis* 2009 Nov;7(11/12):50-51. [doi: [10.1055/s-0032-1308248](https://doi.org/10.1055/s-0032-1308248)]

72. Pfingsten M, Kröner-Herwig B, Leibing E, Kronshage U, Hildebrandt J. Validation of the German version of the Fear-Avoidance Beliefs Questionnaire (FABQ). *Eur J Pain* 2000;4(3):259-266. [doi: [10.1053/eujp.2000.0178](https://doi.org/10.1053/eujp.2000.0178)] [Medline: [10985869](https://pubmed.ncbi.nlm.nih.gov/10985869/)]
73. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 1993 Feb;52(2):157-168. [doi: [10.1016/0304-3959\(93\)90127-B](https://doi.org/10.1016/0304-3959(93)90127-B)] [Medline: [8455963](https://pubmed.ncbi.nlm.nih.gov/8455963/)]
74. Monticone M, Frigau L, Vernon H, et al. Reliability, responsiveness and minimal clinically important difference of the two Fear Avoidance and Beliefs Questionnaire scales in Italian subjects with chronic low back pain undergoing multidisciplinary rehabilitation. *Eur J Phys Rehabil Med* 2020 Oct;56(5):600-606. [doi: [10.23736/S1973-9087.20.06158-4](https://doi.org/10.23736/S1973-9087.20.06158-4)] [Medline: [32420712](https://pubmed.ncbi.nlm.nih.gov/32420712/)]
75. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010 Nov;63(11):1179-1194. [doi: [10.1016/j.jclinepi.2010.04.011](https://doi.org/10.1016/j.jclinepi.2010.04.011)] [Medline: [20685078](https://pubmed.ncbi.nlm.nih.gov/20685078/)]
76. Hays RD, Spritzer KL, Schalet BD, Cella D. PROMIS®-29 v2.0 profile physical and mental health summary scores. *Qual Life Res* 2018 Jul;27(7):1885-1891. [doi: [10.1007/s11136-018-1842-3](https://doi.org/10.1007/s11136-018-1842-3)]
77. Bartlett SJ, Gutierrez AK, Andersen KM, et al. Identifying minimal and meaningful change in a Patient-Reported Outcomes Measurement Information System for rheumatoid arthritis: use of multiple methods and perspectives. *Arthritis Care Res (Hoboken)* 2022 Apr;74(4):588-597. [doi: [10.1002/acr.24501](https://doi.org/10.1002/acr.24501)] [Medline: [33166066](https://pubmed.ncbi.nlm.nih.gov/33166066/)]
78. Meyer K, Sprott H, Mannion AF. Cross-cultural adaptation, reliability, and validity of the German version of the Pain Catastrophizing Scale. *J Psychosom Res* 2008 May;64(5):469-478. [doi: [10.1016/j.jpsychores.2007.12.004](https://doi.org/10.1016/j.jpsychores.2007.12.004)] [Medline: [18440399](https://pubmed.ncbi.nlm.nih.gov/18440399/)]
79. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. *Psychol Assess* 1995;7(4):524-532. [doi: [10.1037/1040-3590.7.4.524](https://doi.org/10.1037/1040-3590.7.4.524)]
80. Monticone M, Portoghesi I, Rocca B, Giordano A, Campagna M, Franchignoni F. Responsiveness and minimal important change of the Pain Catastrophizing Scale in people with chronic low back pain undergoing multidisciplinary rehabilitation. *Eur J Phys Rehabil Med* 2022 Feb;58(1):68-75. [doi: [10.23736/S1973-9087.21.06729-0](https://doi.org/10.23736/S1973-9087.21.06729-0)] [Medline: [34042409](https://pubmed.ncbi.nlm.nih.gov/34042409/)]
81. Saragiotto BT, Maher CG, Yamato TP, et al. Motor control exercise for chronic non-specific low-back pain. *Cochrane Database Syst Rev* 2016 Jan 8;2016(1):CD012004. [doi: [10.1002/14651858.CD012004](https://doi.org/10.1002/14651858.CD012004)] [Medline: [26742533](https://pubmed.ncbi.nlm.nih.gov/26742533/)]
82. Surkitt LD, Ford JJ, Hahne AJ, Pizzari T, McMeeken JM. Efficacy of directional preference management for low back pain: a systematic review. *Phys Ther* 2012 May;92(5):652-665. [doi: [10.2522/ptj.20100251](https://doi.org/10.2522/ptj.20100251)] [Medline: [22247407](https://pubmed.ncbi.nlm.nih.gov/22247407/)]
83. Sluka KA, Frey-Law L, Hoeger Bement M. Exercise-induced pain and analgesia? Underlying mechanisms and clinical translation. *Pain* 2018 Sep;159 Suppl 1(Suppl 1):S91-S97. [doi: [10.1097/j.pain.0000000000001235](https://doi.org/10.1097/j.pain.0000000000001235)] [Medline: [30113953](https://pubmed.ncbi.nlm.nih.gov/30113953/)]
84. Apeldoorn AT, Swart NM, Conijn D, Meerhoff GA, Ostelo RW. Management of low back pain and lumbosacral radicular syndrome: the guideline of the Royal Dutch Society for Physical Therapy (KNGF). *Eur J Phys Rehabil Med* 2024 Apr;60(2):292-318. [doi: [10.23736/S1973-9087.24.08352-7](https://doi.org/10.23736/S1973-9087.24.08352-7)] [Medline: [38407016](https://pubmed.ncbi.nlm.nih.gov/38407016/)]
85. Bourdin P, Barberia I, Oliva R, Slater M. A virtual out-of-body experience reduces fear of death. *PLoS ONE* 2017;12(1):e0169343. [doi: [10.1371/journal.pone.0169343](https://doi.org/10.1371/journal.pone.0169343)] [Medline: [28068368](https://pubmed.ncbi.nlm.nih.gov/28068368/)]
86. Maselli A, Slater M. The building blocks of the full body ownership illusion. *Front Hum Neurosci* 2013;7:83. [doi: [10.3389/fnhum.2013.00083](https://doi.org/10.3389/fnhum.2013.00083)] [Medline: [23519597](https://pubmed.ncbi.nlm.nih.gov/23519597/)]
87. R: a language and environment for statistical computing. R Project. URL: <https://www.R-project.org/> [accessed 2026-03-17]
88. Kassambara A. rstatix: pipe-friendly framework for basic statistical tests. R Project. URL: <https://CRAN.R-project.org/package=rstatix> [accessed 2026-03-17]
89. Navarro D. lsr: companion to “learning statistics with R”. R Project. URL: <https://cran.r-project.org/web/packages/lsr/index.html> [accessed 2026-03-17]
90. Revelle W. psych: procedures for psychological, psychometric, and personality research. R Project. URL: <https://CRAN.R-project.org/package=psych> [accessed 2026-03-17]
91. ggplot2. URL: <https://ggplot2.tidyverse.org> [accessed 2026-03-17]
92. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001 Nov;94(2):149-158. [doi: [10.1016/S0304-3959\(01\)00349-9](https://doi.org/10.1016/S0304-3959(01)00349-9)] [Medline: [11690728](https://pubmed.ncbi.nlm.nih.gov/11690728/)]
93. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain* 2008 Feb;9(2):105-121. [doi: [10.1016/j.jpain.2007.09.005](https://doi.org/10.1016/j.jpain.2007.09.005)] [Medline: [18055266](https://pubmed.ncbi.nlm.nih.gov/18055266/)]
94. Maughan EF, Lewis JS. Outcome measures in chronic low back pain. *Eur Spine J* 2010 Sep;19(9):1484-1494. [doi: [10.1007/s00586-010-1353-6](https://doi.org/10.1007/s00586-010-1353-6)] [Medline: [20397032](https://pubmed.ncbi.nlm.nih.gov/20397032/)]
95. Li R, Li Y, Kong Y, et al. Virtual reality-based training in chronic low back pain: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2024 Feb 26;26:e45406. [doi: [10.2196/45406](https://doi.org/10.2196/45406)] [Medline: [38407948](https://pubmed.ncbi.nlm.nih.gov/38407948/)]

96. Kumar V, Vatkar AJ, Kataria M, Dhatt SS, Baburaj V. Virtual reality is effective in the management of chronic low back ache in adults: a systematic review and meta-analysis of randomized controlled trials. *Eur Spine J* 2024 Feb;33(2):474-480. [doi: [10.1007/s00586-023-08040-5](https://doi.org/10.1007/s00586-023-08040-5)] [Medline: [37979068](https://pubmed.ncbi.nlm.nih.gov/37979068/)]
97. Igna R, Ștefan S, Onac I, Ungur RA, Szentagotai Tatar A. Mindfulness-based cognitive-behavior therapy (MCBT) versus virtual reality (VR) enhanced CBT, versus treatment as usual for chronic back pain. A clinical trial. *J Evid Based Psychother* 2014;14(2):229-247 [[FREE Full text](#)]
98. Zhou T, Salman D, McGregor AH. Recent clinical practice guidelines for the management of low back pain: a global comparison. *BMC Musculoskelet Disord* 2024 May 1;25(1):344. [doi: [10.1186/s12891-024-07468-0](https://doi.org/10.1186/s12891-024-07468-0)] [Medline: [38693474](https://pubmed.ncbi.nlm.nih.gov/38693474/)]
99. Corp N, Mansell G, Stynes S, et al. Evidence-based treatment recommendations for neck and low back pain across Europe: a systematic review of guidelines. *Eur J Pain* 2021 Feb;25(2):275-295. [doi: [10.1002/ejp.1679](https://doi.org/10.1002/ejp.1679)] [Medline: [33064878](https://pubmed.ncbi.nlm.nih.gov/33064878/)]
100. Chhabra HS, Sharma S, Verma S. Smartphone app in self-management of chronic low back pain: a randomized controlled trial. *Eur Spine J* 2018 Nov;27(11):2862-2874. [doi: [10.1007/s00586-018-5788-5](https://doi.org/10.1007/s00586-018-5788-5)] [Medline: [30324496](https://pubmed.ncbi.nlm.nih.gov/30324496/)]
101. Afzal MW, Ahmad A, Hanif HMB, Chaudhary N, Gilani SA. Effects of virtual reality exercises on chronic low back pain: quasi-experimental study. *JMIR Rehabil Assist Technol* 2023 Sep 15;10:e43985. [doi: [10.2196/43985](https://doi.org/10.2196/43985)] [Medline: [37713252](https://pubmed.ncbi.nlm.nih.gov/37713252/)]
102. Spiegel B, Fuller G, Lopez M, et al. Virtual reality for management of pain in hospitalized patients: a randomized comparative effectiveness trial. *PLoS One* 2019;14(8):e0219115. [doi: [10.1371/journal.pone.0219115](https://doi.org/10.1371/journal.pone.0219115)] [Medline: [31412029](https://pubmed.ncbi.nlm.nih.gov/31412029/)]
103. Vlaeyen JWS, Kole-Snijders AMJ, Boeren RGB, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 1995 Sep;62(3):363-372. [doi: [10.1016/0304-3959\(94\)00279-N](https://doi.org/10.1016/0304-3959(94)00279-N)] [Medline: [8657437](https://pubmed.ncbi.nlm.nih.gov/8657437/)]
104. Sullivan MJL, Bishop SR, Pivik J. Pain Catastrophizing Scale (PCS). APA PsycNet. URL: <https://psycnet.apa.org/doiLanding?doi=10.1037%2F01304-000> [accessed 2026-03-17]
105. Fritz JM, George SZ. Identifying psychosocial variables in patients with acute work-related low back pain: the importance of fear-avoidance beliefs. *Phys Ther* 2002 Oct;82(10):973-983. [doi: [10.1093/ptj/82.10.973](https://doi.org/10.1093/ptj/82.10.973)] [Medline: [12350212](https://pubmed.ncbi.nlm.nih.gov/12350212/)]
106. Williamson E. Fear Avoidance Beliefs Questionnaire (FABQ). *Aust J Physiother* 2006;52(2):149. [doi: [10.1016/s0004-9514\(06\)70052-6](https://doi.org/10.1016/s0004-9514(06)70052-6)] [Medline: [16805041](https://pubmed.ncbi.nlm.nih.gov/16805041/)]
107. Thomas JS, France CR. The relationship between pain-related fear and lumbar flexion during natural recovery from low back pain. *Eur Spine J* 2008 Jan;17(1):97-103. [doi: [10.1007/s00586-007-0532-6](https://doi.org/10.1007/s00586-007-0532-6)] [Medline: [17972113](https://pubmed.ncbi.nlm.nih.gov/17972113/)]
108. Thomas JS, France CR. Pain-related fear is associated with avoidance of spinal motion during recovery from low back pain. *Spine (Phila Pa 1976)* 2007 Jul 15;32(16):E460-E466. [doi: [10.1097/BRS.0b013e3180bc1f7b](https://doi.org/10.1097/BRS.0b013e3180bc1f7b)] [Medline: [17632385](https://pubmed.ncbi.nlm.nih.gov/17632385/)]
109. Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev* 2021 Sep 28;9(9):CD009790. [doi: [10.1002/14651858.CD009790.pub2](https://doi.org/10.1002/14651858.CD009790.pub2)] [Medline: [34580864](https://pubmed.ncbi.nlm.nih.gov/34580864/)]
110. Harvie DS, Broecker M, Smith RT, Meulders A, Madden VJ, Moseley GL. Bogus visual feedback alters onset of movement-evoked pain in people with neck pain. *Psychol Sci* 2015 Apr;26(4):385-392. [doi: [10.1177/0956797614563339](https://doi.org/10.1177/0956797614563339)] [Medline: [25691362](https://pubmed.ncbi.nlm.nih.gov/25691362/)]
111. Runswick OR, Siegel L, Rafferty GF, et al. The effects of congruent and incongruent immersive virtual reality modulated exercise environments in healthy individuals: a pilot study. *Int J Human-comput Interact* 2024 Dec;40(23):7864-7874. [doi: [10.1080/10447318.2023.2276524](https://doi.org/10.1080/10447318.2023.2276524)]
112. Bruce RM, Rafferty GF, Finnegan SL, Sergeant M, Pattinson KTS, Runswick OR. Incongruent virtual reality cycling exercise demonstrates a role of perceived effort in cardiovascular control. *J Physiol* 2025 Sep;603(18):5149-5161. [doi: [10.1113/JP287421](https://doi.org/10.1113/JP287421)] [Medline: [39754534](https://pubmed.ncbi.nlm.nih.gov/39754534/)]
113. Kowatsch T, Lohse KM, Erb V, et al. Hybrid ubiquitous coaching with a novel combination of mobile and holographic conversational agents targeting adherence to home exercises: four design and evaluation studies. *J Med Internet Res* 2021 Feb 22;23(2):e23612. [doi: [10.2196/23612](https://doi.org/10.2196/23612)] [Medline: [33461957](https://pubmed.ncbi.nlm.nih.gov/33461957/)]
114. Data and analysis scripts. OSF. URL: <https://osf.io/kar9u/overview> [accessed 2026-03-17]

Abbreviations

- BPS:** Back Performance Scale
- CLBP:** chronic low back pain
- FABQ:** Fear Avoidance Beliefs Questionnaire
- HMD:** head-mounted display
- MCAR:** missing completely at random
- MCID:** minimal clinically important difference
- MDC:** minimal detectable change
- MDT:** McKenzie Method of Mechanical Diagnosis and Therapy
- NRS:** numerical rating scale
- PCS:** Pain Catastrophizing Scale
- PGIC:** Patient Global Impression of Change
- PROMIS 29:** Patient-Reported Outcomes Measurement Information System 29

PSFS: Patient-Specific Functional Scale
rmANOVA: repeated measures ANOVA
RMDQ: Roland-Morris Disability Questionnaire
ROM: range of motion
SSQ: Simulator Sickness Questionnaire
TSK: Tampa Scale for Kinesiophobia
VR: virtual reality

Edited by S Brini; submitted 22.Jul.2025; peer-reviewed by OR Runswick, RL Hsieh; revised version received 08.Feb.2026; accepted 18.Feb.2026; published 30.Mar.2026.

Please cite as:

*Neumann I, Lindner S, Nedilko Y, Zhivova R, Gödde M, Tischer-Zeitz T, Rittner HL, Käthner I
A Novel Virtual Reality Intervention Combining Movement Exercises and Body Illusions for the Treatment of Chronic Back Pain:
Prospective Feasibility Study
JMIR Serious Games 2026;14:e81051
URL: <https://games.jmir.org/2026/1/e81051>
doi: [10.2196/81051](https://doi.org/10.2196/81051)*

© Isabel Neumann, Stefan Lindner, Yevgeniya Nedilko, Ralitsa Zhivova, Michael Gödde, Tobias Tischer-Zeitz, Heike L Rittner, Ivo Käthner. Originally published in JMIR Serious Games (<https://games.jmir.org>), 30.Mar.2026. 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.

Virtual Reality–Based Pain Modulation in Subacute Musculoskeletal Injury: Functional Near-Infrared Spectroscopy Study of Neural and Behavioral Correlates

Ryan Andrew Mace^{1,2,3}, PhD; Ziyang Wu^{1,2,3}, PhD; Christine B Sieberg^{1,2,3,4}, PhD; Keerthana Deepti Karunakaran⁵, PhD; Ke Peng⁶, PhD, MSc; David Borsook^{2,3,7}, MD, PhD

¹Center for Health Outcomes and Interdisciplinary Research, Massachusetts General Hospital, 1 Bowdoin Square Suite 100, Boston, MA, United States

²Department of Psychiatry, Harvard Medical School, Boston, MA, United States

³Department of Psychiatry, Massachusetts General Hospital, Boston, MA, United States

⁴Division of Adolescent and Young Adult Medicine, Boston Children's Hospital, Boston, MA, United States

⁵Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, United States

⁶Department of Electrical and Computer Engineering, Price Faculty of Engineering, University of Manitoba, Winnipeg, MB, Canada

⁷Department of Radiology, Massachusetts General Hospital, Boston, MA, United States

Corresponding Author:

Ryan Andrew Mace, PhD

Center for Health Outcomes and Interdisciplinary Research, Massachusetts General Hospital, 1 Bowdoin Square Suite 100, Boston, MA, United States

Abstract

Background: Acute traumatic musculoskeletal injuries often result in persistent pain and disability despite physical recovery. Virtual reality (VR) provides an innovative approach for overcoming treatment barriers and may help address psychological risk factors for persistent pain and disability. However, the neural mechanisms underlying VR, particularly in subacute orthopedic pain, are insufficiently understood.

Objective: This pilot study examined the feasibility, acceptability, and neural correlates of an 8-week home-based therapeutic VR intervention (RelieVRx) for subacute (<2 mo postinjury) musculoskeletal pain, using functional near-infrared spectroscopy (fNIRS) to assess changes in prefrontal cortex (PFC) activation and coactivation.

Methods: Ten orthopedic patients (mean age 46.8, SD 11.86 years; 7/10, 70% female) completed the RelieVRx intervention and underwent fNIRS and behavioral assessments before and post treatment. Behavioral measures included pain intensity, pain interference, pain catastrophizing, pain anxiety, mindfulness, coping, and pain self-efficacy. fNIRS recorded PFC hemodynamic responses during movement-evoked pain and VR-based relaxation/distraction tasks. Feasibility and acceptability were assessed qualitatively and quantitatively (valid fNIRS recordings, participant feedback). Analyses evaluated pre-post changes in PFC activation, functional coactivation, and correlations with behavioral measures.

Results: fNIRS procedures demonstrated high feasibility (74/80, 93% valid recordings), acceptability, and there were no safety concerns. Significant improvements were observed across all behavioral measures, including reduced pain intensity at rest (mean difference [MD]=−2.50, $P<.001$, $d=2.24$), and with activity (MD=−3.40, $P<.001$, $d=1.98$), decreased pain interference (MD range −3.90 to −4.90, $P<.001$, d range 1.32 to 2.30), reduced pain anxiety (MD=−32.70, $P=.001$, $d=2.14$) and pain catastrophizing (MD=−16.40, $P=.003$, $d=2.13$), and improved mindfulness (MD=+3.29, $P=.01$, $d=0.94$), coping (MD=+0.59, $P=.01$, $d=1.01$), and self-efficacy (MD=15.40, $P=.008$, $d=1.51$). fNIRS showed significant posttreatment increases in medial PFC activation (right medial channel: distraction task $t=-4.473$, $P=.004$; relaxation task $t=-3.036$, $P=0.02$) and enhanced coactivation between medial and lateral PFC regions (distraction task $t=-2.784$, $P=.03$). Increased functional coactivation between the right medial and left lateral PFC was negatively correlated with improved mindfulness ($r=-0.716$, $P=.046$) and coping scores ($r=-0.709$, $P=.049$).

Conclusions: This study provides initial evidence of the feasibility and acceptability of integrating fNIRS neuroimaging into clinical VR interventions. Results indicate that engagement with VR therapy was associated with alterations in neural activity in key PFC regions implicated in pain regulation, correlating with significant improvements in pain and behavioral measures. The neural and behavioral changes highlight the potential of VR as a mechanistically informed, scalable nonpharmacological approach to managing subacute pain following orthopedic injuries. These findings justify larger trials that incorporate multimodal neuroimaging to further understand potential mechanistic processes that may underlie VR-based pain therapies.

Trial Registration: ClinicalTrials.gov NCT05552430; <https://clinicaltrials.gov/study/NCT05552430>

KEYWORDS

immersive virtual reality; musculoskeletal pain; functional near-infrared spectroscopy; prefrontal cortex; pain management; mindfulness

Introduction

Gaps in Pain Management After Injury

Acute traumatic musculoskeletal injuries, such as fractures and dislocations, are both prevalent and costly [1]. Current standard of care for patients with musculoskeletal injuries primarily focuses on the physical aspects of recovery, including the immediate stabilization of injury (eg, surgery, casting, splinting), pharmacological pain management (eg, opioid and nonopioid analgesics), and physiotherapy. However, an estimated 20% to 50% of these cases result in persistent pain and functional limitations, even after the bones and soft tissues have healed [1-3]. These patients present a significant public health challenge and contribute substantially to health care costs and utilization due to the need for frequent medical appointments and multiple surgeries. Further, these injuries and subsequent pain have contributed to the United States opioid epidemic [4,5]. Poor recovery outcomes may be explained by psychological factors after injury, including pain catastrophizing (pain misconceptions, rumination), pain anxiety (fear or worry about pain), and pain self-efficacy (perceived ability to carry out daily activities despite pain), which increase the risk for chronic pain and disability independent of injury type, location, or severity [6]. Behavioral interventions, such as relaxation and mindfulness-based approaches, have been associated with enhanced coping and reduced pain intensity and interference among chronic populations [7,8], including orthopedic injuries [9,10]. However, access to behavioral interventions is often limited in orthopedic settings because treatments prioritize the physical aspects of recovery and due to barriers including mental health stigma, time, cost, transportation, and clinician availability [11,12].

Virtual reality (VR) may help overcome these treatment barriers and demonstrates growing evidence as a pain management tool [13,14]. VR is “a computer-generated simulation of the real or imagined environment or world” [15] that involves immersive, multisensory feedback and user interaction, most commonly delivered through a headset [16]. Immersion and interaction with multisensory VR stimuli, such as nature landscapes or games, is hypothesized to engage the user’s limited attentional resources, potentially diverting focus away from pain perceptions and related stress [14,17]. Most VR studies have focused on distraction to improve pain tolerance limits [17] and pain intensity [14] in chronic pain [18] or other acute pain populations (eg, burns and painful medical procedures) [14]. Additionally, VR provides a novel opportunity to target maladaptive pain responses by teaching relaxation and mindfulness skills. VR relaxation mindfulness studies report demonstrated high adherence and reductions in pain and pain catastrophizing in chronic pain populations [19-21] and feasibility in subacute orthopedic pain [22]. The mechanisms underlying VR distraction and relaxation remain poorly

understood [17], particularly in subacute musculoskeletal pain, limiting our ability to identify the neural and psychological processes through which VR reduces pain.

Role of the Prefrontal Cortex in Pain Modulation

The prefrontal cortex (PFC) plays a pivotal role in the modulation of pain through cognitive and emotional processes and includes various subregions such as the anterior PFC or frontopolar cortex (Brodmann area 10) [23], dorsolateral PFC (dlPFC) [24], ventromedial PFC [25], and orbitofrontal cortex [26]. These areas are integral to processes including attention, emotion regulation, and decision-making, all of which influence pain perception [27]. Neuroimaging studies have consistently shown that many of these PFC subregions are actively engaged in nociceptive processing, playing key roles in both pain perception and modulation [28]. Specifically, the anterior PFC or frontopolar cortex, encompassing medial and lateral Brodmann areas 10, extending into regions of Brodmann areas 12 and 32, is implicated in emotion regulation, decision-making, and cognitive appraisal of pain [29]. Neuroimaging studies have shown that medial PFC (mPFC) activity tracks expectations about pain and mediates expectancy effects on pain-related activity in other brain regions [30]. Functional near-infrared spectroscopy (fNIRS) studies, including our own, have reported robust activity in medial and lateral Brodmann area 10 during acute pain [31,32] that is suppressed by opioids [33,34]. Research investigating the neural correlates of stressor controllability in humans indicates that the mPFC mediates perceptions of control and subsequently regulates stress responses [35]. These findings suggest that mPFC, along with other subregions, could be investigated as potential neural targets for pain management interventions such as VR-based therapies [36].

Indeed, several studies have shown that brain regions engaged during VR overlap extensively with those implicated in chronic pain and behavioral interventions, including the dlPFC [37,38]. The dlPFC is a complex structure (spanning Brodmann areas 9, 8a, 8b, and 46) responsible for executive control, including the maintenance and regulation of top-down modulation and driving appropriate behavioral responses [39]. When compared with controls, increased dlPFC activation has been observed in response to painful stimuli in patients with chronic pain, such as fibromyalgia and knee osteoarthritis [40], likely due to central sensitization [41]. Activation of the dlPFC is associated with cognitive control of pain and decreased pain catastrophizing, suggesting its important role in pain coping [42]. These findings highlight the pivotal role of multiple PFC subregions in integrating cognitive, emotional, and evaluative processes to modulate pain, underscoring their importance as potential neural targets for pain management interventions such as VR-based therapies [36].

fNIRS for VR Neuroimaging

fNIRS has gained significant attention for characterizing neural mechanisms of VR [43-45]. fNIRS uses low-energy near-infrared light to quantify cortical hemodynamic variations by measuring changes in oxygenated hemoglobin (HbO) and deoxygenated hemoglobin concentrations. Unlike other neuroimaging methods, fNIRS is generally less affected by electrical interference, making it highly compatible with VR headsets [43]. Additional advantages of fNIRS for VR research include that it (1) is noninvasive; (2) is relatively low-cost; (3) offers higher motion tolerance, enabling brain measurements during movement; (4) operates silently, which preserves the sense of immersion in VR environments [43,46]; and (5) is commonly used as a reliable measure of PFC activation and connectivity in pain [47,48]. Despite these advantages, only 2 studies have used fNIRS to characterize pain and VR-induced changes in PFC activity. Deng et al [49] used VR to divert participants' attention while electrical pain stimuli were applied, resulting in significant activations in the dLPFC and premotor cortex. Hu et al [50] explored the brain mechanisms of mindful breathing using immersive VR and found that meditation increased pain thresholds, likely by enhancing PFC connectivity. These findings underscore the potential of fNIRS to capture dynamic changes in PFC activity related to pain and its modulation, providing a robust framework for investigating how different VR-based pain management strategies influence neural mechanisms in subacute pain after musculoskeletal injury.

Study Aims and Hypotheses

The principal objective of this study was to explore the role of VR on neural mechanisms and behavioral measures for subacute pain following a pilot sample of acute traumatic musculoskeletal injury. We analyzed fNIRS data collected during a feasibility pilot study of an 8-week home-based therapeutic VR (RelieVRx) in 10 orthopedic patients with subacute pain after injury [22]. RelieVRx is Food and Drug Administration–authorized VR therapy for chronic lower back pain [51-53] and has demonstrated high feasibility, satisfaction, and preliminary improvements in orthopedic pain [22]. RelieVRx is a multicomponent therapeutic VR that includes both relaxation (eg, guided deep breathing) and distraction (eg, focused attention games) conditions, which enabled us to examine differential PFC activation patterns. We hypothesized that VR engagement would be associated with altered PFC hemodynamic activity and functional coactivation (FC), which refers to the temporal correlation between spatially distinct brain regions reflecting synchronized neural activity. Specifically, we expected that engagement in VR-based relaxation and distraction would lead to distinct patterns of PFC activation, with potential differences in medial versus lateral PFC engagement. Furthermore, we hypothesized that these neural adaptations would be associated with pre- and post-VR changes in behavioral measures, including pain outcomes (pain interference and intensity) and psychological mechanisms of action (pain catastrophizing, pain anxiety, pain self-efficacy, mindfulness, and coping).

We proposed 3 aims to understand the neural mechanisms and behavioral correlates of VR-based pain modulation in patients with subacute musculoskeletal injuries. First, we aimed to

evaluate the feasibility of our fNIRS data collection procedures and the acceptability of fNIRS from patients' perspectives using 30-minute qualitative exit interviews. This initial step to confirm the integrity of our data collection procedures and potential burden on patients is essential prior to subsequent mechanistic investigation because no study has used fNIRS to understand VR treatment effects after acute traumatic musculoskeletal injuries. Second, we investigated alterations in PFC activation and FC during 2 movement-evoked pain (injured and noninjured) and 2 VR (distraction and relaxation) conditions before and after participants completed the VR. Finally, we explored possible brain–behavior associations between pre- and post-VR changes in PFC activity and behavioral measures (pain outcomes, psychological mechanisms of action). The study aimed to provide a proof of concept for integrating VR and fNIRS neuroimaging in orthopedic populations. Consistent with the National Institute of Health Stage Model and National Center for Complementary and Integrative Health Research Framework, these aims establish a stepwise approach—from feasibility, to characterizing neural response, to brain-behavior relationships—designed to inform a subsequent mechanistic clinical trial of RelieVRx targeting reductions in pain interference after orthopedic injury.

Methods

Ethical Considerations

The study was approved by the Mass General Brigham Institutional Review Board (Protocol #2022P001500) and preregistered at ClinicalTrials.gov (NCT05552430). This study adheres to the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist [54]. The ancillary reviews by the Partners Research Information Security Office approved the VR (AppliedVR RelieVRx), and the Laser Safety Biomedical Engineering Review approved the fNIRS device (OctaMon, Artinis Medical Systems, The Netherlands). All participants provided written informed consent. The study protected participant privacy through secure data storage, restricted team access, encrypted transfers, coded identifiers, MGB-compliant communication, and use of a VR headset (Pico G2 4K) specifically selected to avoid the collection of personal information. Participants were compensated US \$50 total (US \$25 per assessment visit). Full methodological details, derived from a pilot feasibility study of RelieVRx for acute orthopedic injuries, are described in Mace et al [22]. Analyses of the fNIRS data have not been reported elsewhere.

Participants

Participants with acute orthopedic traumatic musculoskeletal injuries were recruited through surgeon referrals at 2 Mass General Brigham Level 1 Trauma Clinics, as well as via flyers and the Partners Rally online research platform. Our target sample size of 10 participants was determined based on guidelines for pilots [55,56], National Institute of Health recommendations for testing feasibility [57,58], and similar pilot studies involving novel applications of fNIRS [38] and VR [18] for pain. While the sample size was appropriate for testing the feasibility of integrating fNIRS with VR and

detecting neural responses [59-64] by design, it was not statistically powered for efficacy analysis or reproducibility.

Inclusion criteria were as follows: (1) age ≥ 18 years; (2) fluency and literacy in English; (3) upper or lower extremity orthopedic injury (eg, fracture, dislocation, or rupture) or surgical repair within the previous 2 months; (4) psychological risk for persistent pain and disability, indicated by a Pain Anxiety Symptom Scale-20 score ≥ 40 or a Pain Catastrophizing Scale-13 (PCS-13) score ≥ 20 [65,66]; (5) internet access; (6) willingness to participate and comply with the study protocol; (7) stable psychotropic medication regimen for >6 weeks; and (8) clearance by the referring orthopedic surgeon.

Exclusion criteria were as follows: (1) history of persistent pain (≥ 3 mo) prior to the injury; (2) complex polytrauma (≥ 1 injury location); (3) epilepsy, seizure disorder, dementia, migraines, or other neurological conditions; (4) nausea or dizziness; (5) light hypersensitivity; (6) significant vision or hearing impairment; (7) injuries to the eyes, face, or neck that hinder

VR use; (8) medical conditions expected to worsen within 3 months; (9) untreated severe mental illness (eg, bipolar disorder, schizophrenia, active substance use); (10) surgical complications (eg, infection and need for repeat surgery); (11) ongoing litigation or Worker's Comp claims; (12) pregnancy; (13) regular practice of cognitive-behavioral therapy or other mind-body techniques (>1 /wk, ≥ 45 min).

As depicted in [Multimedia Appendix 1](#), a total of 109 patients were referred for the study. Of these, 49 underwent screening, and 12 were deemed eligible. The primary reasons for exclusion were lack of clinically significant pain catastrophizing or pain anxiety ($n=21$), injury occurred more than 2 months ago ($n=7$), and history of chronic pain ($n=6$). Two eligible participants withdrew before the baseline visit. [Table 1](#) presents the characteristics of the 10 participants who completed the VR and assessments. [Multimedia Appendix 2](#) presents injury and treatment details for all participants at baseline. All participants were right-handed.

Table . Demographics and clinical characteristics (n=10).

Characteristics	Values
Age (years), mean (SD)	46.8 (11.86)
Time since injury (days), mean (SD)	38.40 (36.33)
Time since surgery (days), mean (SD)	23.29 (17.32)
Engagement in physiotherapy (d/w), mean (SD)	2.83 (2.14)
Gender, n (%)	
Woman	7 (70.0)
Man	3 (30.0)
Ethnicity, n (%)	
Not Hispanic or Latino/Latina	10 (100.0)
Race, n (%)	
White	10 (100.0)
Marital status, n (%)	
Married	3 (30.0)
Single, never married	5 (50.0)
Separated or Divorced	1 (10.0)
Widowed	1 (10.0)
Education, n (%)	
Some college/Associates degree (<16 y)	1 (10.0)
Completed college (16 y)	3 (30.0)
Graduate/professional degree (>16 y)	6 (60.0)
Employment, n (%)	
Employed full-time	8 (80.0)
Employed part-time	1 (10.0)
Unemployed	1 (10.0)
Income (US \$), n (%)	
15,000 to less than 20,000	1 (10.0)
50,000 to less than 75,000	3 (30.0)
75,000 or more	6 (60.0)
Injury location, n (%)	
Tibia	1 (10.0)
Ankle	2 (20.0)
Foot	1 (10.0)
Humerus	1 (10.0)
Wrist	2 (20.0)
Finger	3 (30.0)
Injury type, n (%)	
Fracture	9 (90.0)
Rupture	1 (10.0)
Surgery, n (%)	
Yes	7 (70.0)
No	3 (30.0)
Physiotherapy, n (%)	

Characteristics	Values
Yes	6 (60.0)
No	4 (40.0)
Past year prior pain, n (%)	
Yes	3 (30.0)
No	7 (70.0)
Depression diagnosis, n (%)	
Previously	1 (10.0)
Currently	3 (30.0)
Never diagnosed	6 (60.0)
Anxiety, n (%)	
Currently	3 (30.0)
Never diagnosed	7 (70.0)
Posttraumatic stress disorder (PTSD), n (%)	
Currently	1 (10.0)
Never diagnosed	9 (90.0)
Substance use, n (%)	
Never diagnosed	10 (100.0)
Psychotropic medications, n (%)	
Yes	3 (30.0)
No	7 (70.0)
Nonnarcotic medications, n (%)	
Yes	6 (60.0)
No	4 (40.0)

Enrollment Procedures

Procedures were based on prior orthopedic behavioral intervention trials [61,67] and VR pilot studies [51,52,68]. A research assistant screened participants for eligibility and scheduled baseline, pre-VR, and post-VR study visits at our research center. At the baseline visit, participants provided informed consent, completed assessments including fNIRS, and received a VR demonstration. The VR came with a travel case, charger, cleaning cloth, and user instructions. After the 8-week VR intervention, participants repeated the assessments and returned the VR headset.

VR Intervention

Participants received a rental Pico G2 4K headset preloaded with the therapeutic VR (RelieVRx). The Pico G2 4K headset is lightweight (278 g), high-resolution (4K VR 3840 × 2160 display, 75 Hz refresh rate), and affordable (\$245). The headset has several user-friendly features, including minimal setup, compatibility with glasses, and “hands-free” controls via the user’s head movements. The RelieVRx intervention, developed by AppliedVR [51], is a self-guided therapeutic program based on the biopsychosocial model of pain, incorporating cognitive behavioral therapy and mindfulness principles. The primary treatment components are VR relaxation and distraction. Relaxation conditions guide the participant through progressive

relaxation exercises that are enhanced by biofeedback and immersive environments (eg, deep breathing while watching a sunset and observing the breath). Distraction conditions involve interactive focused attention games (eg, earning points by tracking butterflies amid distractor visual stimuli). Participants also receive pain neuroscience education to explain the biopsychosocial aspects of pain and the role of relaxation and distraction techniques (eg, visualizing pain reduction through nervous system calming). Over 8 weeks, participants completed one daily module (total=56 conditions, average=6 min, range 2 to 16 min) in a fixed sequence. The conditions are designed to be accessible and minimize potential risks (eg, emotional distress or motion sickness).

Measures

Behavioral Measures

To explore brain–behavior relationships, we included several measures of pain outcomes (pain intensity and interference) and psychological mechanisms of action targeted by VR (pain catastrophizing, pain anxiety, pain self-efficacy, mindfulness, and coping). Pain intensity at rest and with activity in the last week was measured using the 2-item (0=“no pain”; 10=“worst ever”) Numerical Rating Scale. The Numerical Rating Scale is reliable and valid in orthopedic populations [69]. Pain interference with activity, sleep, mood, and stress was measured

with the 4-item (0=“does not interfere”; 10=“completely interferes”) Defense and Veterans Pain Rating Scale [70]. The Defense and Veterans Pain Rating Scale was the primary outcome of RelieVRx trials [51,53,68] and is reliable and valid in acute and persistent pain populations [71]. Pain catastrophizing was measured using the 13-item (0=“not at all”; 4=“all the time”) PCS [72]. The PCS-13 (range 0 - 52) is reliable and valid in populations with musculoskeletal pain [73]. Fear and anxiety related to pain were measured using the 20-item (0=“never”; 5=“always”) Pain Anxiety Symptoms Scale [74]. The Pain Anxiety Symptoms Scale-20 (range 0 - 100) is reliable and valid in pain populations [75]. Pain self-efficacy was measured using the 10-item (0=“not at all confident”; 6=“completely confident”) Pain Self-Efficacy Questionnaire [76]. The Pain Self-Efficacy Questionnaire (range 0 - 60) is reliable and valid in chronic pain populations [76]. Mindfulness was measured using the 12-item (1=“rarely/not at all”; 4=“almost always”) Cognitive and Affective Mindfulness Scale – Revised [77]. The Cognitive and Affective Mindfulness Scale – Revised (range 12 - 48) has been used to measure mindfulness in pain [60,78], including orthopedic populations [79]. Coping

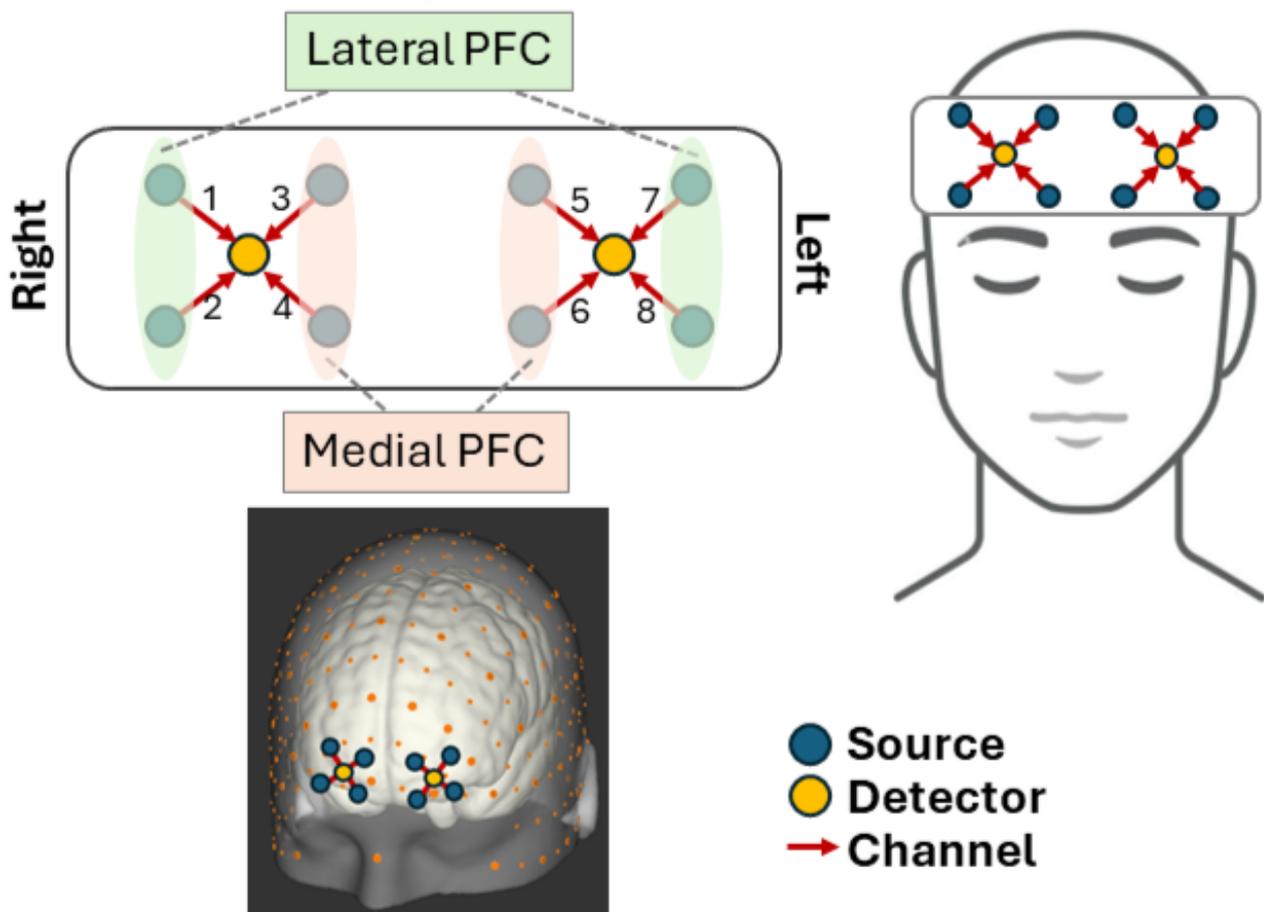
was measured (0=“I cannot do this at all”; 4=“I can do this extremely well”) using the 13-item Measure of Current Status [80]. The Measure of Current Status (range 0 - 52) is a reliable and valid measure of healthy coping skills in pain [60,78] and orthopedic populations [79].

fNIRS

Overview

We measured changes in PFC activation in response to acute pain and VR with a wireless 8-channel fNIRS optical topography system (OctaMon, Artinis Medical Systems, The Netherlands). The PFC is divided into 4 regions in each hemisphere: 2 medial and 2 lateral (Figure 1 displays sensor layout details). fNIRS measures changes in HbO and deoxygenated hemoglobin concentrations through the propagation of near-infrared light between LED emitters and receivers [17]. The OctaMon system used in this study uses 2 wavelengths of light at 760 nm and 850 nm [81]. This approach shares a similar physiological basis as the blood-oxygenation-level-dependent signal activation in traditional functional magnetic resonance imaging.

Figure 1. Artinis OctaMon 8-channel optical topography functional near-infrared spectroscopy (fNIRS) system. PFC: prefrontal cortex.



fNIRS Data Acquisition

Participants underwent OctaMon fNIRS recordings after completing the self-reports during the baseline and postintervention visits. Participants sat in a comfortable stationary chair. The trained research assistant and principal investigator placed the standard OctaMon head cap on the

participant’s forehead and firmly attached it with the side straps following standard operating procedures. The head cap had 8 mounted emitters (4 on each hemisphere) and 2 detectors (1 on each hemisphere) arranged in the OctaMon standard octode template over the PFC. The signal sampling rate was 10 Hz using a 2×4 emitter-detector channel layout. The differential pathlength factor was calculated based on participant age (range

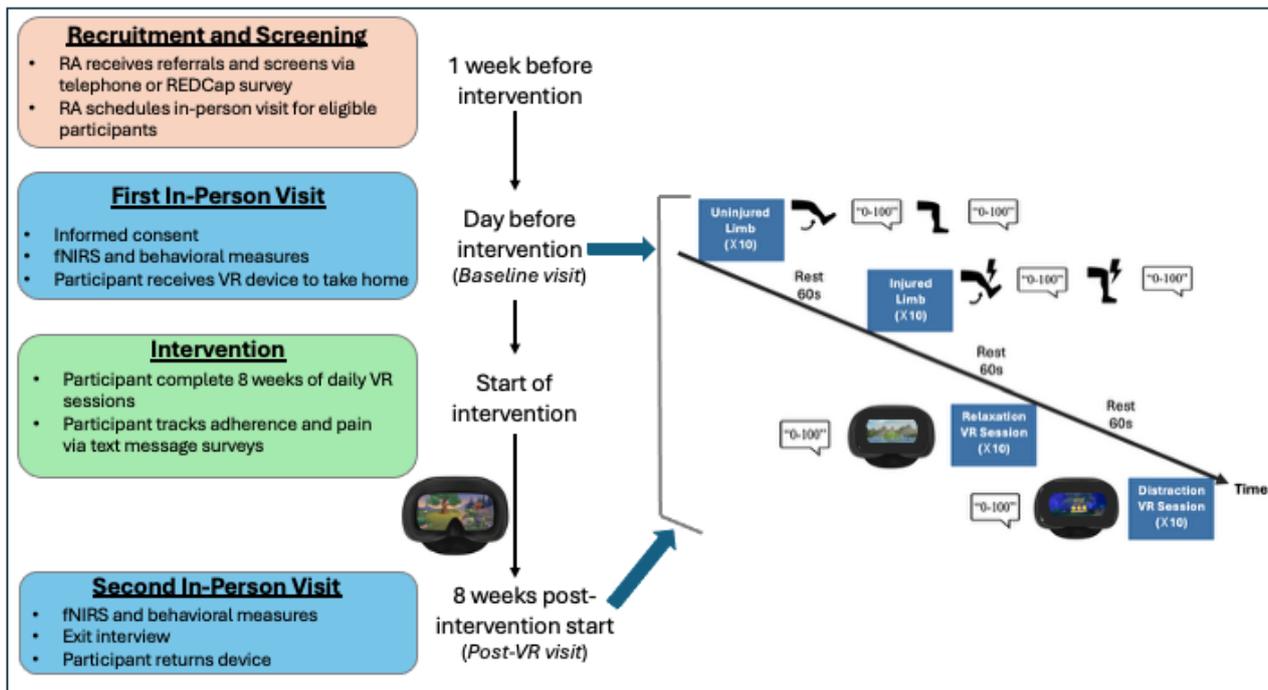
6.00 - 6.61 years). Participants were asked about their comfort level wearing the fNIRS, and none reported any problems. Additionally, the principal investigator (RAM) and research assistant monitored the continuous fNIRS data during the experiments to reduce potential confounds (eg, movement or ambient light) to ensure data quality. The headset was adjusted to improve signal quality as needed (eg, moving hair underneath optodes, ensuring skin contact). Raw fNIRS data were collected in Oxysoft (Artinis Medical Systems, The Netherlands).

fNIRS Experimental Design

Overview

An overview of the experimental design is presented in Figure 2. We describe the movement-evoked pain (injured and noninjured) and VR (distraction and relaxation) conditions in detail below.

Figure 2. An overview of the experimental design. fNIRS: functional near-infrared spectroscopy; RA: research assistant; REDCap: Research Electronic Data Capture; VR: virtual reality.



Movement-Evoked Pain Conditions

We instituted a standardized movement-evoked pain paradigm across participants and sessions to maximize internal consistency while preserving ecological validity. First, a scripted instruction set was delivered verbatim at every visit to avoid variation in verbal cues. The tasks began from a fixed starting posture and joint-specific neutral alignment, determined with orthopedic surgeon collaborators, and participants maintained consistent body posture throughout. A predefined plane of movement was specified for each injured joint (eg, sagittal flexion/extension for wrist or knee), and participants performed active movement only (no external assistance) at a slow, uniform tempo ($\approx 2 - 3$ s to end range) into the first pain threshold, held that position for 5 seconds, then relaxed for 3 seconds. After the initial hold, participants were instructed to move approximately 5 degrees further (visually estimated from anatomical landmarks) and hold for another 5 seconds, followed by return to rest. This hold-move-hold cycle was repeated 10 times per side with immediate pain ratings at the end of each hold (0 - 100 for granular repeated pain ratings), both to confirm pain evocation and to improve fNIRS block reliability (ie, signal averaging) [82].

Participants always performed the noninjured (control) side first, which provided a within-participant reference of minimal pain movement, followed by the injured side for movement-provoked pain. This ordering is consistent with orthopedic examination workflows and supports external validity for real-world limb testing. To standardize the “five degrees more” increment, examiners used trained visual estimation anchored to anatomical landmarks (eg, joint lines and bony prominences) and practiced on exemplar trials prior to data collection. Although visual estimation is known to introduce measurement error relative to devices, it is common in clinical range of motion exams (eg, in gait or outpatient orthopedic settings) and has acceptable clinical utility [83]. Across all visits, movement speed, plane, posture, and tempo remained constant. The same scripted instructions, starting alignment, and tempo were preserved from session to session. We marked the start and end of each block on the time series fNIRS data using OxySoft event labels. We reset the graph traces before and after we marked each block to reorient the device for each condition. Participants were given 2 minutes to rest before the VR conditions.

Relaxation and Distraction VR Conditions

Participants completed two 6-minute VR conditions while wearing the fNIRS. We designed 2 conditions to measure PFC

activation associated with 2 mechanisms of VR pain modulation: relaxation and distraction. Relaxation is intended to downregulate the sympathetic nervous system response to pain and related stress [17,84], while distraction aims to divert attention away from pain and stress signaling [85,86]. The order of the tasks was fixed to preserve the standardized sequence of the RelieVRx program and to maintain data quality during fNIRS acquisition. We considered counterbalancing, but it was not feasible without removing the fNIRS equipment, which would have disrupted signal quality and intervention fidelity. The relaxation condition was a 6-minute guided relaxation in a calming outdoor environment (module 6 of RelieVRx). Participants were instructed to notice internal physical sensations while showing and deepening their breath. Using the same range as the movement-evoked pain condition, participants rated their overall relaxation on a scale of 0 (“Not relaxed at all”) to 100 (“Very relaxed”) as an experimental manipulation check. The distraction condition was a 6-minute concentration game (module 8 of RelieVRx). Participants followed objects with their gaze to earn points across 3 levels of increasing difficulty: (1) tracking a single butterfly, (2) tracking a butterfly with a target color, and (3) tracking a butterfly with a target color among foil butterflies. Participants rated their overall focus on the visual target stimuli on a scale of 0 (“Not focused at all”) to 100 (“Very focused”) as a proxy for distraction from pain. To increase immersion, overhead lights were dimmed, participants wore headphones connected to the headset, and the study staff left the room during both VR conditions.

Individual-Level fNIRS Data Analyses

fNIRS Data Preprocessing

fNIRS data preprocessing was performed with a combination of OxySoft, Homer2 [87], and customized scripts written in MATLAB (Mathworks). Preprocessing included the following steps. First, initial quality checks were performed using OxySoft [88], a proprietary Artinis Medical Systems software platform that provides a live signal display feature for monitoring head motion artifacts. Large and abrupt changes in signal intensity, indicative of excessive head motion, were manually marked and trimmed in real-time to mitigate motion-induced noise and maintain high fidelity in fNIRS recordings. For fNIRS data that passed the quality check, raw light intensity data were then converted into optical density change using `hmrIntensity2OD` function in Homer2 [87]. We automatically flagged brief motion-related spikes on each channel using a standard algorithm from the Homer2 toolbox that marks samples when the instantaneous jump in the light signal exceeds an amplitude threshold or when the short-term variability exceeds an SD threshold. We used a 0.5 s window with a 1 s mask, a 5-SD criterion for variability, and an amplitude criterion of 0.1 in optical-density units. Motion-correlated optical density data were further filtered with a third-order Butterworth bandpass filter, a commonly used filtering approach in fNIRS signal processing to remove high-frequency noise, such as cardiac and respiratory artifacts, and low-frequency drifts [89], with high and low cutoff frequencies of 0.2 Hz and 0.01 Hz, respectively. After artifact screening and band-pass filtering (0.01 - 0.20 Hz), we converted optical-density data to oxy- and deoxy-hemoglobin concentration changes using the modified Beer–Lambert law

with age-adjusted differential pathlength factors. In this study, we focused on the analysis of the HbO time courses as HbO changes are reported to have a much higher signal-to-noise ratio and may have greater sensitivity to task-evoked changes [90].

Computation of Brain Imaging and Effect Size Metrics

Three types of brain imaging measures were then calculated based on the preprocessed fNIRS signals.

Task-Induced Activation Using General Linear Modeling

The neural activation magnitude for each region of interest was measured using a general linear model approach, which modeled the relationship between the hemodynamic response and task conditions by convolving the timing of the tasks with a canonical hemodynamic response function to estimate beta coefficients representing activation strength [91].

Task-Induced Change in HbO Using Effect Size

Additionally, for each condition (injured, noninjured, VR distraction, VR relaxation), we calculated an effect size metric (Cohen d) by taking the difference between the mean signal during the condition and the mean signal during the prestimulus baseline interval [38].

State-Related Functional Activity During VR

Pair-wise FC was computed using the correlation of HbO concentration time series among PFC subregions. FC measures can help understand how different regions of the PFC are engaged as a network during a given brain state (eg, pain or VR). Pearson correlation coefficients were calculated for all pair-wise combinations of channels within and across hemispheres to assess FC [92]. We further derived cross-regional metrics by averaging correlation values between specific channel pairs (eg, medial-left PFC and lateral-right PFC).

Statistical Analyses

For Aim 1, the feasibility and acceptability of fNIRS data collection were assessed using a mixed methods approach. Descriptive statistics (eg, completion rates and signal quality metrics) were used to evaluate the feasibility of fNIRS data collection. Reasons for missing or invalid fNIRS recordings were documented and categorized to provide further insight into procedural challenges. We conducted a 30-minute in-person or remote (participant preference) individual exit interview during the postintervention visit to further understand participants’ perceptions of the fNIRS data collection. We used rapid assessment procedures to analyze the transcripts consistent with established qualitative frameworks [93-95]. Rapid assessment is a valid alternative to in-depth qualitative methods for generating timely, actionable insights to guide [96,97]. All authors met to collaboratively discuss each exit interview, discuss key content, and summarize emergent themes in a shared matrix to identify patterns across participants [98]. All authors collaboratively reviewed each transcript, discussed key content, and summarized emergent themes in a shared matrix to identify patterns across participants. Two team members independently completed the rapid assessments, and discrepancies were adjudicated by the principal investigator (RAM).

For Aim 2, paired sample t tests were conducted to compare baseline and posttreatment fNIRS measures for each participant,

focusing on regional activation magnitudes and pair-wise FC across 4 conditions: movement-evoked pain in the injured and noninjured extremities and VR distraction and relaxation. These analyses aimed to identify significant neural changes following VR therapy. To further localize statistically significant hemodynamic activations at the group level, paired sample t tests were applied to the Cohen d parameter for each fNIRS channel separately. Cohen d was computed by taking the difference between the mean signal in the time range of the condition and the mean signal in the prestimulus baseline interval. This approach was adapted from previous fNIRS studies using Cohen d as a standardized effect size metric for assessing neural activation changes [38]. For each participant, Cohen d values were calculated for movement-evoked pain and VR conditions at both baseline and posttreatment. Multiple comparison corrections were applied using the false discovery rate [99].

For Aim 3, Pearson's correlation analyses were performed to examine associations between pre-to-post VR changes in PFC activity and FC and changes in behavioral outcomes. Specifically, correlations were conducted between statistically significant neural response changes and self-reported pain outcomes (pain interference, pain intensity) and psychological mechanisms of action (pain catastrophizing, pain anxiety, pain self-efficacy, mindfulness, and coping).

Finally, to assess potential effects of sample heterogeneity, we examined associations between time since injury and (1) baseline measures and (2) change scores (postbaseline) using Spearman correlations with Benjamini–Hochberg correction. We performed sensitivity analyses excluding the single 138-day case and using leave-one-out and robust regression. We further modeled postintervention outcomes via analysis of covariance, including baseline value, time since injury, and injury type (collapsed categories) as covariates. Injury type and time since injury were not significant predictors, and model inferences for the intervention remain unchanged.

Results

Aim 1: Feasibility and Acceptability

Of the 80 possible fNIRS recordings (ie, 10 participants, 2 time points, 4 conditions [pain, no pain, distraction, and relaxation]), 74 (93%) of them resulted in valid data that were included in the analysis, suggesting high feasibility of fNIRS data collection. Five recordings were missing, and one recording was invalid due to optode signal interference. Manipulation checks provided internal validity that injured movements evoked more pain (mean 44.11, SD 17.17) than the noninjured movements (mean 2.74, SD 7.60). Similarly, manipulation checks indicated that the VR conditions elicited high relaxation ratings (mean 80.25, SD 11.19) and distraction ratings (mean 82.50, SD 11.40) as intended (0=min, 100=max). No major safety concerns related to VR or fNIRS were noted during the study. Participants reported no discomfort or VR sickness.

Exit interviews revealed that fNIRS was acceptable to participants. Participants described the device as noninvasive, comfortable, and easy to use. Prior to the study, none of the

participants had prior knowledge of fNIRS. They found the informational pamphlet helpful before the initial visit. They valued the study team's emphasis on safety and appreciated opportunities to ask questions. Several participants viewed fNIRS as more accessible than MRI. Advance notice regarding hair adjustments prevented discomfort. They requested more information about the purpose and interpretation of fNIRS data, and they affirmed the importance of institutional approvals (eg, their orthopedic surgeon's involvement). Some indicated that it felt cumbersome to wear fNIRS and VR equipment. Despite these inconveniences, participants expressed willingness to use fNIRS in future orthopedic research and showed interest in the results.

VR Therapy and Behavioral Measures

We observed significant and large pre-post VR reductions in all outcomes, including decreases in pain intensity at rest (mean difference [MD]=−2.50, $P<.001$, $d=2.24$), pain intensity with activity (MD=−3.40, $P<.001$, $d=1.98$), and pain interference (MD range −3.90 to −4.90, $P<.001$, d range 1.32 to 2.30). Similarly, we also found significant and large pre-post VR improvements in the mechanisms of action. Pain anxiety (MD=−32.70, $P=.001$, $d=2.14$) and pain catastrophizing (MD=−16.40, $P=.003$, $d=2.13$) significantly decreased, while pain self-efficacy (MD=+15.40, $P=.008$, $d=1.51$), mindfulness (MD=+3.29, $P=.01$, $d=0.94$), and coping (MD=+0.59, $P=.01$, $d=1.01$) significantly increased.

Aim 2: Group-Level Hemodynamic Activation

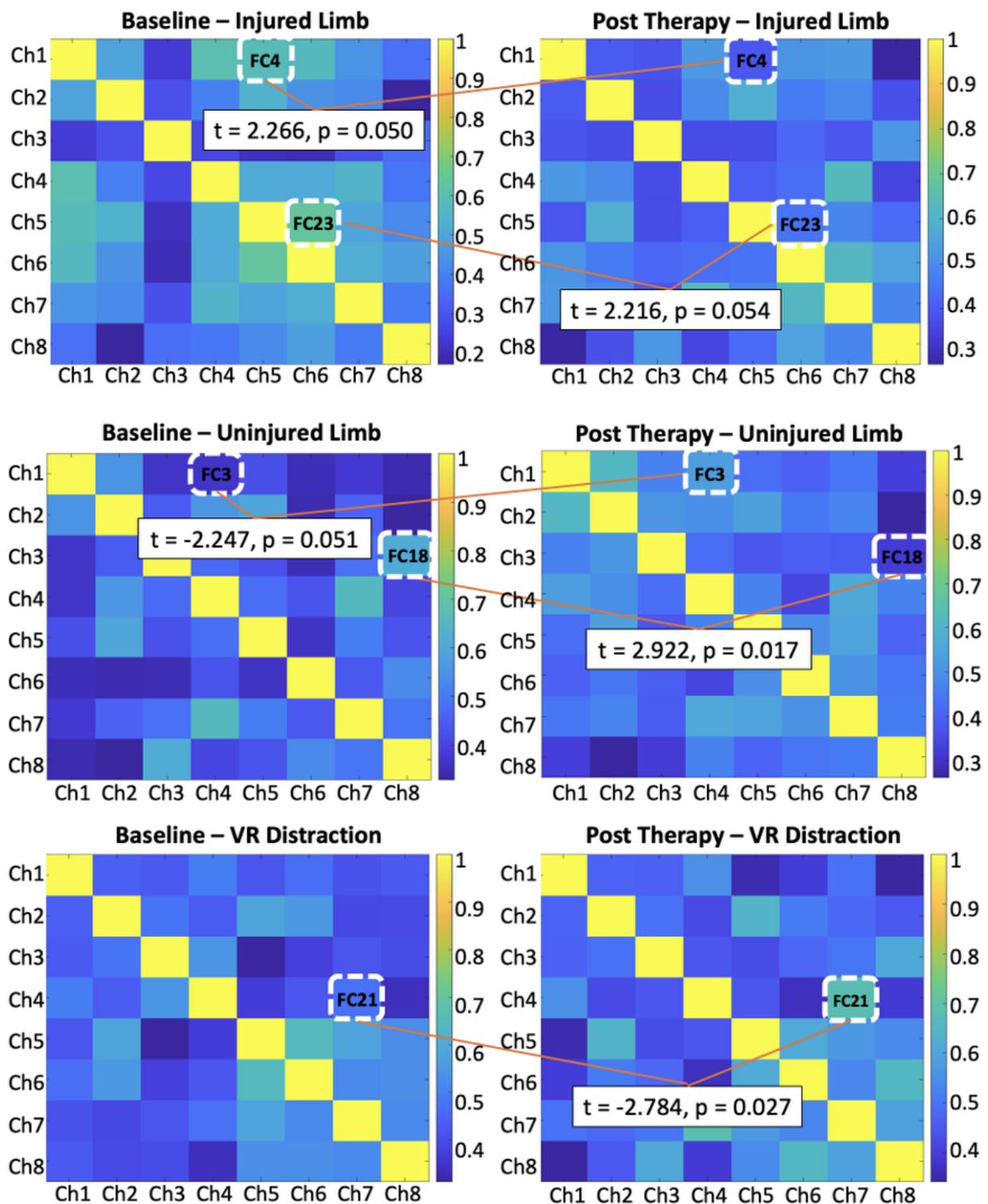
Regarding hemodynamic activation magnitudes (ie, beta value measures) for the VR distraction and VR relaxation conditions at baseline and posttreatment, significant differences in activation were observed between pre and posttreatment fNIRS recordings. Prior to the initiation of the noninjured condition, a significant reduction in HbO concentration was observed at channel 8, corresponding to the left lateral PFC ($t=-2.585$, $P=.03$).

Analysis of changes in Cohen d parameter between baseline and posttreatment fNIRS measurements revealed a significant increase in HbO at channel 4, which represents the right mPFC. Specifically, this increase was significant for both conditions (VR distraction condition: $t=-4.473$, $P=.004$; VR relaxation: $t=-3.036$, $P=.03$).

Aim 2: Functional Coactivation Analysis

FC analyses identified significant differences between baseline and posttreatment conditions across multiple comparisons (Figure 3). In the injured condition, changes were observed in FC23 ($t=2.216$, $P=.05$), representing coactivation within the left mPFC, and FC4 ($t=2.266$, $P=.05$), representing coactivation between the right lateral PFC and left mPFC. In the noninjured condition, alterations were detected in FC18 ($t=2.922$, $P=.02$), which represents coactivation between the right mPFC and left lateral PFC, and FC3 ($t=-2.247$, $P=.05$), representing coactivation between the right lateral PFC and right mPFC. In the VR distraction condition, significant differences were identified in FC21 ($t=-2.784$, $P=.03$), which represents coactivation between the right mPFC and left lateral PFC.

Figure 3. Distinct patterns of functional connectivity between baseline and posttreatment conditions. Ch: channel; FC: functional coactivation; *P*: statistical significance; *t*: average difference; VR: virtual reality.



Aim 3: Brain-Behavior Associations

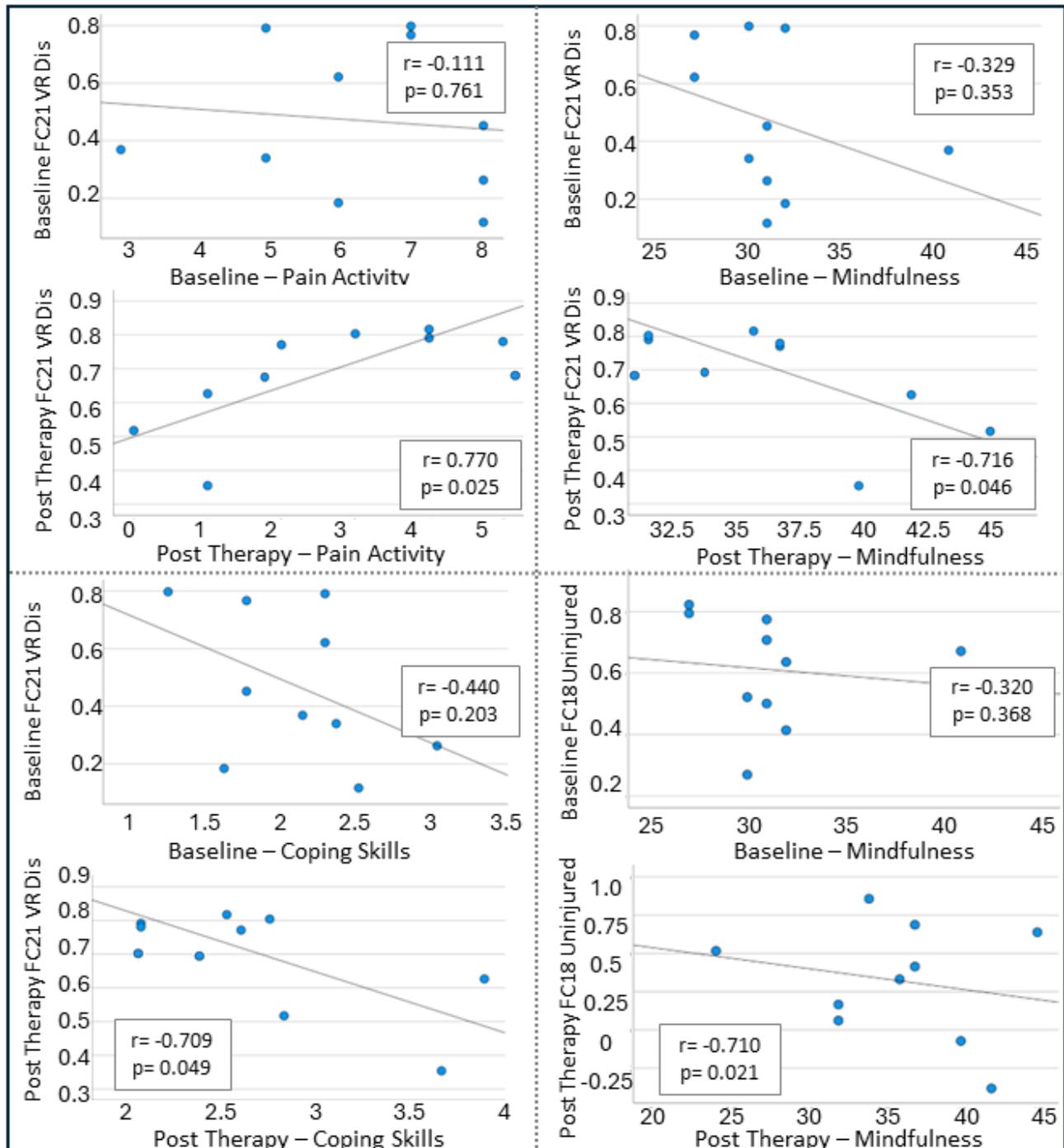
To investigate the relationship between HbO response changes and behavioral outcomes, correlations between alterations in Cohen *d* parameters and self-reported measures were examined. A significant correlation was observed between the hemodynamic response to painful stimuli at channel 4 (right mPFC) and pain-related activity, with greater increases in pain

levels associated with increased activation in this region ($r=0.760, P=.047$). FC analyses further revealed key findings (Figure 4). In the VR distraction condition, baseline FC21 (right mPFC – left lateral PFC activity) was not significantly correlated with pain activity ($r=-0.111, P=.76$). However, posttreatment FC21 demonstrated a strong positive correlation with pain activity ($r=0.770, P=.02$). Additionally, posttreatment FC21 (right mPFC – left lateral PFC activity) exhibited significant

negative correlations with mindfulness ($r=-0.716$, $P=.046$) and coping skills ($r=-0.709$, $P=.049$); however, such correlation patterns were not observed at baseline (mindfulness: $r=-0.329$, $P=.353$; coping skills: $r=-0.440$, $P=.203$). Additionally, in the

noninjured condition, posttreatment FC18 (right mPFC – left lateral PFC activity) showed a significant negative correlation with mindfulness ($r=-0.710$, $P=.021$), whereas no significant correlation was observed at baseline ($r=-0.320$, $P=.368$).

Figure 4. Brain-behavior correlation results. The x-axis refers to behavior measures of pain activity (Numerical Rating Scale [NRS]), mindfulness (Cognitive and Affective Mindfulness Scale – Revised, CAMS-R), and coping skills (measure of current status, MOCS) completed at the baseline and postintervention visits. FC: functional coactivation; VRDi: virtual reality distraction condition.



Discussion

Principal Findings

This study systematically investigated the neural mechanisms and behavioral measures underlying pain modulation in patients with recent acute musculoskeletal injuries undergoing an 8-week

home-based therapeutic VR intervention. fNIRS characterized PFC activity and FC in response to VR distraction and relaxation conditions, as well as movement-evoked pain. Our findings indicate significant neurophysiological adaptations in the PFC alongside significant and large improvements in patient-reported pain outcomes (interference and intensity) and psychological mechanisms (pain catastrophizing, pain anxiety, pain

self-efficacy, mindfulness, and coping), reinforcing the potential of VR therapy as a clinically viable nonpharmacological treatment for subacute pain after musculoskeletal injury. Notably, this interventional study complements cross-sectional and provides preliminary evidence of pre-post differences in brain-behavior associations during a VR program. Our study contributes to the literature by suggesting (1) the feasibility of integrating fNIRS into clinical trial settings to evaluate nonpharmacological pain treatments and (2) the utility of combining neuroimaging and behavioral measures to improve understanding of treatment mechanisms and therapeutic effects. Given the small sample size and single-arm design, the observed associations between VR and neurobehavioral adaptations are preliminary and should be interpreted with caution, as described in the sections below.

Differentiated Neural Response Patterns in the PFC

A primary objective of this study was to elucidate the role of the PFC in VR-mediated pain modulation. Our findings show alterations in hemodynamic activity and FC during painful movement across multiple PFC subregions after VR therapy, reinforcing the central role of PFC in top-down pain regulation [100] and its potential as a neurobiological target for VR-based pain therapy. The PFC plays a crucial role in pain perception, integrating cognitive, affective, and sensory components of pain processing, with the medial and dorsolateral areas of PFC particularly implicated in cognitive control mechanisms, including pain suppression and emotion regulation [27,39].

In chronic pain conditions, hyperactivity in the dIPFC is often associated with increased pain perception, reflecting its role in pain chronification rather than actively suppressing pain [37,101]. Conversely, effective behavioral pain management approaches, including cognitive behavioral therapy and mindfulness-based interventions, have been shown to normalize or reduce excessive PFC activity, suggesting an adaptive regulatory role [102,103]. Our findings indicate that VR therapy may similarly engage PFC mechanisms involved in cognitive pain modulation. Increased activation in the right mPFC (channel 4) following both VR distraction and relaxation may reflect increased engagement of neural processes implicated in pain regulation. Prior neuroimaging studies have linked the mPFC to cognitive reappraisal of pain, self-referential processing, and emotion regulation [104,105], indicating that VR therapy may strengthen these adaptive neural pathways. Increased activation in the mPFC during both conditions aligns with prior findings that distraction-based VR predominantly engages attentional networks, while relaxation-based VR enhances interoceptive awareness and downregulates stress responses [106,107]. These findings underscore the need to further delineate these overlapping but functionally distinct neural networks involved in VR-based pain management strategies.

Beyond regional activation, modulations in PFC coactivation helped characterize state-dependent network engagement across conditions. The presence of comparable or stronger connectivity changes during uninjured limb movements indicates that these effects are not specific to nociceptive input, but instead likely reflect domain-general regulatory and attentional processes that

are engaged during movement and modulated by VR. Notably, FC21 (right mPFC-left lateral PFC coactivation during VR distraction) exhibited a significant increase following VR therapy. While postintervention associations between medial-lateral PFC coactivation and behavioral outcomes are consistent with models of pain regulation, the absence of corresponding baseline relationships precludes conclusions about stable individual differences in regulatory capacity. This pathway has been associated with executive control and the integration of affective and sensory aspects of pain processing [39]. However, rather than reflecting a monotonic marker of regulatory efficiency, medial-lateral PFC coactivation likely indexes regulatory demand or engagement, with greater coupling occurring when pain-related cognitive control requirements are higher. Under this interpretation, increased FC21 may reflect compensatory recruitment of control networks in the presence of greater pain or task demands, rather than direct evidence of improved regulation.

Although potentially counterintuitive, the inverse associations between medial-lateral PFC coactivation and mindfulness/coping are consistent with models in which effective self-regulation reflects reduced reliance on concurrent default-mode network (DMN) and executive-control network engagement, indicating greater functional segregation and more efficient network switching. In this framework, individuals with stronger psychological resources may require less medial-lateral PFC coactivation to perform the same task, whereas individuals experiencing greater pain or lower coping capacity may exhibit increased coactivation as a compensatory response. Thus, individuals with stronger psychological resources (such as mindfulness) may require less concurrent DMN-executive control networks engagement during the VR distraction task, yielding lower medial-lateral coactivation after treatment. Conversely, people with lower coping skills may require greater PFC engagement to perform the same task. This view is supported by DMN-control anti-correlations observed during adaptive regulation [36,100,108], by prior fNIRS findings of opposing mPFC versus lateral PFC responses under analgesic/pain-modulatory states [23], and by chronic-pain literature linking disrupted DMN connectivity (including DMN-mPFC coupling) to pain rumination [40,105]. Collectively, the present inverse correlations may reflect recovery/normalization of DMN-executive control networks interactions following VR training, with potential trait-dependent differential effects on attentional control and coping; however, these findings are exploratory and merit replication in larger mechanistic cohorts. These findings contribute to the growing body of evidence demonstrating that therapeutic interventions and their relationships with behavioral outcomes can dynamically modulate regional and network activity in the PFC.

Therapeutic Implications

Analyses of brain-behavior correlates provided further evidence of the therapeutic potential of VR interventions for subacute pain management. A significant correlation was observed between the hemodynamic response to painful stimuli at channel 4 (right mPFC) and self-reported pain, such that greater pain intensity was associated with increased activation in this region during VR distraction. This finding implicates the mPFC in the

ffective processing of pain and suggests that interventions capable of dampening mPFC hyperactivity, such as VR distraction, may attenuate perceived pain intensity [104,105]. Additionally, mindfulness and coping were negatively associated with FC between the right medial and left lateral PFC (FC18 during uninjured limb movement, and FC21 during VR distraction) following treatment. These associations suggest that medial–lateral PFC coactivation reflects context-dependent regulatory engagement, such that individuals with stronger psychological resources may achieve regulation with reduced concurrent PFC coupling, whereas greater coactivation may reflect increased regulatory effort in the presence of higher pain or affective demand. Specifically, individuals with higher mindfulness may require less effortful top-down control to manage pain, consistent with evidence that mindfulness training reduces prefrontal over-engagement and promotes more automatic emotion regulation [108–110]. Although significant associations were observed between neural and psychological outcomes, the direction of causality remains uncertain. It is unclear whether changes in brain coactivation facilitated improvements in mindfulness and coping or if psychological gains induced neurophysiological adaptation. Nevertheless, these findings align with prior research emphasizing the critical role of cognitive-affective PFC networks in adaptive pain modulation [111] and underscore the potential of VR-based interventions to foster beneficial neurobehavioral adaptations.

Last, several nonexclusive explanations may account for why brain–behavior relationships were observed only postintervention. First, limited measurement reliability or restricted variance at baseline—particularly within a subacute pain cohort—may have attenuated detectable associations prior to intervention. Second, postintervention associations may be confounded by unmeasured factors such as engagement or expectancy that covary with both neural responses and self-reported outcomes. Third, given the modest sample size and multiple exploratory tests, posttreatment correlations have the risk of reflecting type I error.

Taken together, the emergence of postintervention brain–behavior correlations, though preliminary, suggests potentially meaningful VR-related neuroadaptations. For instance, the positive association between medial–lateral PFC coactivation and pain activity serves as a dynamic rather than a unidimensional marker of treatment success. It likely reflects a state-dependent coupling metric that varies as a function of pain intensity, cognitive demand, and regulatory capacity, potentially indexing compensatory engagement or regulatory effort. Though the single-arm design and modest sample size preclude a definitive causal mechanism, these findings provide an empirical foundation for evaluating VR-based pain interventions. By identifying specific neurobiological targets, these findings are intended to inform the design of future randomized, adequately powered mechanistic studies incorporating appropriate control conditions, repeated longitudinal neural assessments, and direct measurement of engagement-related factors to rigorously test how brain–behavior relationships evolve with intervention.

Limitations

Five limitations warrant consideration: (1) fNIRS spatial resolution, (2) movement-evoked pain variability, (3) sample size, (4) single-arm design, and (5) fixed VR protocol.

First, the restricted spatial coverage of fNIRS limits insights into broader pain-processing networks, including the insula, anterior cingulate cortex, and sensorimotor regions—key areas in pain perception and modulation [24]. fNIRS lacks the spatial resolution to precisely localize activity within PFC, highlighting the need for complementary neuroimaging methods like functional magnetic resonance imaging or high-density fNIRS to refine our understanding of prefrontal involvement in VR-mediated pain processing.

Second, individual variability in pain and range of motion introduces heterogeneity in our movement-evoked pain condition, despite our protocol to standardize posture, plan, speed, and scripted instructions. These constraints reflect the ecological intent of the protocol; yet, future trials should use device-based angle capture (eg, digital inclinometers) to reduce variability and consider including additional standardized pain assessments (eg, quantitative sensory testing).

Third, the small sample size ($N=10$) constrains generalizability. However, this study was explicitly designed as a feasibility pilot to evaluate the acceptability and technical integration of fNIRS with a therapeutic VR intervention in subacute pain. In line with established guidance for pilot studies [18,38,55–58], findings from this sample are intended to be descriptive and hypothesis-generating rather than confirmatory. Small samples can produce unstable effect estimates and elevate both type I and type II error risk; accordingly, observed brain–behavior associations—particularly those with P values near the $\alpha=.05$ threshold (eg, $P=.046$ to $.049$)—should be interpreted as exploratory. These preliminary results are meant to inform the design, targeting, and analytic strategy of future adequately powered mechanistic trials, not to establish efficacy or definitive effect sizes.

Fourth, the single-arm design precludes causal inference and limits our ability to distinguish natural recovery from VR-specific effects. In this pre-post feasibility pilot, all participants received the VR intervention, making it unclear whether observed neural and behavioral changes reflect VR-related processes, spontaneous recovery following subacute musculoskeletal injury, or their interaction. Given the bidirectional relationship between cognitive-affective processes and prefrontal coactivation, it is also possible that neural and psychological changes emerged in parallel rather than through a unidirectional, treatment-driven mechanism. In the absence of a control condition, improvements in pain outcomes and putative mechanisms cannot be attributed specifically to VR and may reflect confounds commonly identified in pain research, including regression to the mean (eg, elevated pain at baseline that abates over time), expectancy and placebo effects, and nonspecific time and attention associated with study visits and staff interactions [112]. Importantly, isolated acute orthopedic injuries are also typically associated with substantial pain reduction within the first several months following injury or surgery, independent of adjunctive behavioral interventions

[113]. Accordingly, all findings are interpreted as pre-post associations within a single cohort rather than evidence of efficacy or VR-specific neuroplasticity. These limitations underscore the necessity of future randomized controlled mechanistic trials incorporating appropriate control conditions to disentangle natural recovery from intervention-specific neural and behavioral effects.

Finally, the fixed VR task order may have introduced order effects [114]. We were constrained by the sequential RelieVRx program and the practical need to maintain optode placement and data quality. We did not observe participant fatigue, as evidenced by the VR task manipulation checks. While similar fixed-order designs appear across fNIRS feasibility and real-world studies [115-117], future studies should incorporate counterbalancing to isolate task-specific neural adaptations. Although tasks and acquisition were standardized and correlations used within-subject change scores, individualized movement amplitudes and residual measurement noise may still contribute variance; results should be viewed as exploratory pending confirmation in larger samples.

Conclusions

The findings of this study provide proof of concept that VR engagement was associated with measurable pre-post differences in PFC activation and FC, which are associated with improvements in behavioral pain outcomes. The involvement of the PFC in modulating pain through VR also supports investigating whether neuromodulatory targeting of this region could enhance therapeutic effects. Notably, both VR and fNIRS

have emerged as complementary technologies with accelerating momentum in neuroscience and rehabilitation research. Their combined application, especially within immersive, ecologically valid environments, enables simultaneous intervention and measurement of cortical dynamics during real-time engagement [43]. This convergence supports the development of closed-loop, precision medicine approaches, such as adaptive VR-based neurofeedback systems that can tailor therapeutic content based on individualized brain-behavior profiles. Future research should include larger, more diverse cohorts, integrate multimodal neuroimaging to validate and extend these findings, and use longitudinal designs to evaluate the durability of VR-induced effects. Investigating interindividual variability in neural response will also be critical for optimizing and personalizing VR-based care.

To address these gaps, we are planning a mechanistic clinical trial of RelieVRx that will examine brain-behavior relationships using whole-head fNIRS and investigate the biomechanical and neural mechanisms underlying rehabilitation following subacute traumatic musculoskeletal injuries. Additional planned studies include a multisite randomized clinical trial to confirm the feasibility of the VR and fNIRS protocol across different orthopedic settings with more diverse patient samples and a subsequent efficacy trial of RelieVRx versus a sham VR control. This program of research will advance the refinement of VR-based therapies toward precision subacute pain management. Such trials, if successful, may advance this approach to individualized medicine.

Acknowledgments

The authors would like to thank Julie Brewer, Joshua Cohen, Thuan Ly, Mike Weaver, Bassem Elhassam, Abhiram Bhashyam, Rohit Garg, Chaitanya Mudgal, Neal Chen, Rachel Cross, Zane Madi, Makenna Law, AppliedVR, and Artinis Medical Systems for supporting this study. Special thanks to Androu Adalmalak and Angelo Herrera from NIRx for their functional near-infrared spectroscopic analysis support.

Funding

This study was funded by The Cathedral Fund.

Data Availability

The datasets generated or analyzed during this study are not publicly available but are available to researchers via a secure data transfer agreement upon email request to the principal investigator (RAM).

Conflicts of Interest

AppliedVR is the developer and owner of RelieVRx. Ten RelieVRx virtual reality headsets (one-year rentals) were purchased from AppliedVR for this study. The Cathedral Fund or AppliedVR had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Multimedia Appendix 1

CONSORT diagram of study participants.

[[PPTX File, 45 KB - games_v14i1e77713_app1.pptx](#)]

Multimedia Appendix 2

Baseline injury and treatment details.

[[DOCX File, 16 KB - games_v14i1e77713_app2.docx](#)]

References

1. Morris S, Lenihan B, Duddy L, O'Sullivan M. Outcome after musculoskeletal trauma treated in a regional hospital. *J Trauma* 2000 Sep;49(3):461-469. [doi: [10.1097/00005373-200009000-00013](https://doi.org/10.1097/00005373-200009000-00013)] [Medline: [11003324](https://pubmed.ncbi.nlm.nih.gov/11003324/)]
2. Busse JW, Heels-Ansdell D, Makosso-Kallyth S, et al. Patient coping and expectations predict recovery after major orthopaedic trauma. *Br J Anaesth* 2019 Jan;122(1):51-59. [doi: [10.1016/j.bja.2018.06.021](https://doi.org/10.1016/j.bja.2018.06.021)] [Medline: [30579406](https://pubmed.ncbi.nlm.nih.gov/30579406/)]
3. Clay FJ, Newstead SV, Watson WL, Ozanne-Smith J, Guy J, McClure RJ. Bio-psychosocial determinants of persistent pain 6 months after non-life-threatening acute orthopaedic trauma. *J Pain* 2010 May;11(5):420-430. [doi: [10.1016/j.jpain.2009.12.002](https://doi.org/10.1016/j.jpain.2009.12.002)] [Medline: [20439055](https://pubmed.ncbi.nlm.nih.gov/20439055/)]
4. Proctor TJ, Mayer TG, Gatchel RJ, McGeary DD. Unremitting health-care-utilization outcomes of tertiary rehabilitation of patients with chronic musculoskeletal disorders. *J Bone Joint Surg Am* 2004 Jan;86(1):62-69. [doi: [10.2106/00004623-200401000-00011](https://doi.org/10.2106/00004623-200401000-00011)] [Medline: [14711947](https://pubmed.ncbi.nlm.nih.gov/14711947/)]
5. Mojtabai R. National trends in long-term use of prescription opioids. *Pharmacoepidemiol Drug Saf* 2018 May;27(5):526-534. [doi: [10.1002/pds.4278](https://doi.org/10.1002/pds.4278)] [Medline: [28879660](https://pubmed.ncbi.nlm.nih.gov/28879660/)]
6. Vranceanu AM, Bachoura A, Weening A, Vrahas M, Smith RM, Ring D. Psychological factors predict disability and pain intensity after skeletal trauma. *J Bone Joint Surg Am* 2014 Feb 5;96(3):e20. [doi: [10.2106/JBJS.L.00479](https://doi.org/10.2106/JBJS.L.00479)] [Medline: [24500592](https://pubmed.ncbi.nlm.nih.gov/24500592/)]
7. Turner JA, Anderson ML, Balderson BH, Cook AJ, Sherman KJ, Cherkin DC. Mindfulness-based stress reduction and cognitive behavioral therapy for chronic low back pain: similar effects on mindfulness, catastrophizing, self-efficacy, and acceptance in a randomized controlled trial. *Pain* 2016 Nov;157(11):2434-2444. [doi: [10.1097/j.pain.0000000000000635](https://doi.org/10.1097/j.pain.0000000000000635)] [Medline: [27257859](https://pubmed.ncbi.nlm.nih.gov/27257859/)]
8. Cherkin DC, Sherman KJ, Balderson BH, et al. Effect of mindfulness-based stress reduction vs cognitive behavioral therapy or usual care on back pain and functional limitations in adults with chronic low back pain: a randomized clinical trial. *JAMA* 2016;315(12):1240-1249. [doi: [10.1001/jama.2016.2323](https://doi.org/10.1001/jama.2016.2323)] [Medline: [27002445](https://pubmed.ncbi.nlm.nih.gov/27002445/)]
9. Westenberg RF, Zale EL, Heinhuis TJ, et al. Does a brief mindfulness exercise improve outcomes in upper extremity patients? a randomized controlled trial. *Clin Orthop Relat Res* 2018 Apr;476(4):790-798. [doi: [10.1007/s11999-0000000000000086](https://doi.org/10.1007/s11999-0000000000000086)] [Medline: [29480886](https://pubmed.ncbi.nlm.nih.gov/29480886/)]
10. Chad-Friedman E, Talaei-Khoei M, Ring D, Vranceanu AM. First use of a brief 60-second mindfulness exercise in an orthopedic surgical practice; results from a pilot study. *Arch Bone Jt Surg* 2017 Nov;5(6):400-405. [Medline: [29299495](https://pubmed.ncbi.nlm.nih.gov/29299495/)]
11. Darnall BD, Scheman J, Davin S, et al. Pain psychology: a global needs assessment and national call to action. *Pain Med* 2016 Feb;17(2):250-263. [doi: [10.1093/pm/pnv095](https://doi.org/10.1093/pm/pnv095)] [Medline: [26803844](https://pubmed.ncbi.nlm.nih.gov/26803844/)]
12. Vranceanu AM, Beks RB, Guitton TG, Janssen SJ, Ring D. How do orthopaedic surgeons address psychological aspects of illness? *Arch Bone Jt Surg* 2017 Jan;5(1):2-9. [Medline: [28271080](https://pubmed.ncbi.nlm.nih.gov/28271080/)]
13. Mallari B, Spaeth EK, Goh H, Boyd BS. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis. *J Pain Res* 2019;12:2053-2085. [doi: [10.2147/JPR.S200498](https://doi.org/10.2147/JPR.S200498)] [Medline: [31308733](https://pubmed.ncbi.nlm.nih.gov/31308733/)]
14. Keefe FJ, Huling DA, Coggins MJ, et al. Virtual reality for persistent pain: a new direction for behavioral pain management. *Pain* 2012 Nov;153(11):2163-2166. [doi: [10.1016/j.pain.2012.05.030](https://doi.org/10.1016/j.pain.2012.05.030)] [Medline: [22770840](https://pubmed.ncbi.nlm.nih.gov/22770840/)]
15. Mantovani F, Castelnuovo G, Gaggioli A, Riva G. Virtual reality training for health-care professionals. *Cyberpsychol Behav* 2003 Aug;6(4):389-395. [doi: [10.1089/109493103322278772](https://doi.org/10.1089/109493103322278772)] [Medline: [14511451](https://pubmed.ncbi.nlm.nih.gov/14511451/)]
16. Sherman WR, Craig AB. *Understanding Virtual Reality: Interface, Application, and Design*; Morgan Kaufmann; 2018.
17. Colloca L, Raghuraman N, Wang Y, et al. Virtual reality: physiological and behavioral mechanisms to increase individual pain tolerance limits. *Pain* 2020 Sep 1;161(9):2010-2021. [doi: [10.1097/j.pain.0000000000001900](https://doi.org/10.1097/j.pain.0000000000001900)] [Medline: [32345915](https://pubmed.ncbi.nlm.nih.gov/32345915/)]
18. Garrett B, Taverner T, McDade P. Virtual reality as an adjunct home therapy in chronic pain management: an exploratory study. *JMIR Med Inform* 2017 May 11;5(2):e11. [doi: [10.2196/medinform.7271](https://doi.org/10.2196/medinform.7271)] [Medline: [28495661](https://pubmed.ncbi.nlm.nih.gov/28495661/)]
19. Eccleston C, Fisher E, Liikkanen S, et al. A prospective, double-blind, pilot, randomized, controlled trial of an "embodied" virtual reality intervention for adults with low back pain. *Pain* 2022 Sep 1;163(9):1700-1715. [doi: [10.1097/j.pain.0000000000002617](https://doi.org/10.1097/j.pain.0000000000002617)] [Medline: [35324507](https://pubmed.ncbi.nlm.nih.gov/35324507/)]
20. MacIntyre E, Sigersteth M, Larsen TF, et al. Get your head in the game: a replicated single-case experimental design evaluating the effect of a novel virtual reality intervention in people with chronic low back pain. *J Pain* 2023 Aug;24(8):1449-1464. [doi: [10.1016/j.jpain.2023.03.013](https://doi.org/10.1016/j.jpain.2023.03.013)] [Medline: [37030584](https://pubmed.ncbi.nlm.nih.gov/37030584/)]
21. Groenveld TD, Smits MLM, Knoop J, et al. Effect of a behavioral therapy-based virtual reality application on quality of life in chronic low back pain. *Clin J Pain* 2023 Jun 1;39(6):278-285. [doi: [10.1097/AJP.0000000000001110](https://doi.org/10.1097/AJP.0000000000001110)] [Medline: [37002877](https://pubmed.ncbi.nlm.nih.gov/37002877/)]
22. Mace RA, Brewer JR, Cohen JE, Ly TV, Weaver MJ, Borsook D. Virtual reality for subacute pain after orthopedic traumatic musculoskeletal injuries: a mixed methods pilot study. *Clin J Pain* 2024 Sep 1;40(9):526-541. [doi: [10.1097/AJP.0000000000001231](https://doi.org/10.1097/AJP.0000000000001231)] [Medline: [39016312](https://pubmed.ncbi.nlm.nih.gov/39016312/)]
23. Peng K, Steele SC, Becerra L, Borsook D. Brodmann area 10: collating, integrating and high level processing of nociception and pain. *Prog Neurobiol* 2018 Feb;161:1-22. [doi: [10.1016/j.pneurobio.2017.11.004](https://doi.org/10.1016/j.pneurobio.2017.11.004)] [Medline: [29199137](https://pubmed.ncbi.nlm.nih.gov/29199137/)]
24. Apkarian AV, Bushnell MC, Treede RD, Zubieta JK. Human brain mechanisms of pain perception and regulation in health and disease. *Eur J Pain* 2005 Aug;9(4):463-484. [doi: [10.1016/j.ejpain.2004.11.001](https://doi.org/10.1016/j.ejpain.2004.11.001)] [Medline: [15979027](https://pubmed.ncbi.nlm.nih.gov/15979027/)]

25. Roy M, Shohamy D, Wager TD. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends Cogn Sci* 2012 Mar;16(3):147-156. [doi: [10.1016/j.tics.2012.01.005](https://doi.org/10.1016/j.tics.2012.01.005)] [Medline: [22310704](https://pubmed.ncbi.nlm.nih.gov/22310704/)]
26. Krummenacher P, Candia V, Folkers G, Schedlowski M, Schönbachler G. Prefrontal cortex modulates placebo analgesia. *Pain* 2010 Mar;148(3):368-374. [doi: [10.1016/j.pain.2009.09.033](https://doi.org/10.1016/j.pain.2009.09.033)] [Medline: [19875233](https://pubmed.ncbi.nlm.nih.gov/19875233/)]
27. Wiech K, Ploner M, Tracey I. Neurocognitive aspects of pain perception. *Trends Cogn Sci* 2008 Aug;12(8):306-313. [doi: [10.1016/j.tics.2008.05.005](https://doi.org/10.1016/j.tics.2008.05.005)] [Medline: [18606561](https://pubmed.ncbi.nlm.nih.gov/18606561/)]
28. Seminowicz DA, Davis KD. A re-examination of pain-cognition interactions: implications for neuroimaging. *Pain* 2007 Jul;130(1-2):8-13. [doi: [10.1016/j.pain.2007.03.036](https://doi.org/10.1016/j.pain.2007.03.036)] [Medline: [17482364](https://pubmed.ncbi.nlm.nih.gov/17482364/)]
29. Etkin A, Egner T, Kalisch R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn Sci* 2011 Feb;15(2):85-93. [doi: [10.1016/j.tics.2010.11.004](https://doi.org/10.1016/j.tics.2010.11.004)] [Medline: [21167765](https://pubmed.ncbi.nlm.nih.gov/21167765/)]
30. Atlas LY, Wager TD. How expectations shape pain. *Neurosci Lett* 2012 Jun 29;520(2):140-148. [doi: [10.1016/j.neulet.2012.03.039](https://doi.org/10.1016/j.neulet.2012.03.039)] [Medline: [22465136](https://pubmed.ncbi.nlm.nih.gov/22465136/)]
31. Yücel MA, Aasted CM, Petkov MP, Borsook D, Boas DA, Becerra L. Specificity of hemodynamic brain responses to painful stimuli: a functional near-infrared spectroscopy study. *Sci Rep* 2015 Mar 30;5(1):9469. [doi: [10.1038/srep09469](https://doi.org/10.1038/srep09469)]
32. Peng K, Deepti Karunakaran K, Lee A, et al. Rhythmic change of cortical hemodynamic signals associated with ongoing nociception in awake and anesthetized individuals: an exploratory functional near infrared spectroscopy study. *Anesthesiology* 2021 Nov 1;135(5):877-892. [doi: [10.1097/ALN.0000000000003986](https://doi.org/10.1097/ALN.0000000000003986)] [Medline: [34610092](https://pubmed.ncbi.nlm.nih.gov/34610092/)]
33. Peng K, Yücel MA, Steele SC, et al. Morphine attenuates fNIRS signal associated with painful stimuli in the medial frontopolar cortex (medial BA 10). *Front Hum Neurosci* 2018;12:394. [doi: [10.3389/fnhum.2018.00394](https://doi.org/10.3389/fnhum.2018.00394)] [Medline: [30349466](https://pubmed.ncbi.nlm.nih.gov/30349466/)]
34. Karunakaran KD, Kussman BD, Peng K, et al. Brain-based measures of nociception during general anesthesia with remifentanyl: a randomized controlled trial. *PLOS Med* 2022 Apr;19(4):e1003965. [doi: [10.1371/journal.pmed.1003965](https://doi.org/10.1371/journal.pmed.1003965)] [Medline: [35452458](https://pubmed.ncbi.nlm.nih.gov/35452458/)]
35. Maier SF, Watkins LR. Role of the medial prefrontal cortex in coping and resilience. *Brain Res* 2010 Oct 8;1355(52-60):52-60. [doi: [10.1016/j.brainres.2010.08.039](https://doi.org/10.1016/j.brainres.2010.08.039)] [Medline: [20727864](https://pubmed.ncbi.nlm.nih.gov/20727864/)]
36. Kucyi A, Davis KD. The dynamic pain connectome. *Trends Neurosci* 2015 Feb;38(2):86-95. [doi: [10.1016/j.tins.2014.11.006](https://doi.org/10.1016/j.tins.2014.11.006)] [Medline: [25541287](https://pubmed.ncbi.nlm.nih.gov/25541287/)]
37. Seminowicz DA, Moayed M. The dorsolateral prefrontal cortex in acute and chronic pain. *J Pain* 2017 Sep;18(9):1027-1035. [doi: [10.1016/j.jpain.2017.03.008](https://doi.org/10.1016/j.jpain.2017.03.008)] [Medline: [28400293](https://pubmed.ncbi.nlm.nih.gov/28400293/)]
38. Öztürk Ö, Alğun ZC, Bombacı H, Erdoğan SB. Changes in prefrontal cortex activation with exercise in knee osteoarthritis patients with chronic pain: an fNIRS study. *J Clin Neurosci* 2021 Aug;90:144-151. [doi: [10.1016/j.jocn.2021.05.055](https://doi.org/10.1016/j.jocn.2021.05.055)] [Medline: [34275540](https://pubmed.ncbi.nlm.nih.gov/34275540/)]
39. Lorenz J, Minoshima S, Casey KL. Keeping pain out of mind: the role of the dorsolateral prefrontal cortex in pain modulation. *Brain (Bacau)* 2003 May;126(Pt 5):1079-1091. [doi: [10.1093/brain/awg102](https://doi.org/10.1093/brain/awg102)] [Medline: [12690048](https://pubmed.ncbi.nlm.nih.gov/12690048/)]
40. Loggia ML, Berna C, Kim J, et al. Disrupted brain circuitry for pain-related reward/punishment in fibromyalgia. *Arthritis Rheumatol* 2014 Jan;66(1):203-212. [doi: [10.1002/art.38191](https://doi.org/10.1002/art.38191)] [Medline: [24449585](https://pubmed.ncbi.nlm.nih.gov/24449585/)]
41. Gracely RH, Petzke F, Wolf JM, Clauw DJ. Functional magnetic resonance imaging evidence of augmented pain processing in fibromyalgia. *Arthritis Rheum* 2002 May;46(5):1333-1343. [doi: [10.1002/art.10225](https://doi.org/10.1002/art.10225)] [Medline: [12115241](https://pubmed.ncbi.nlm.nih.gov/12115241/)]
42. Crombez G, Eccleston C, Van Damme S, Vlaeyen JWS, Karoly P. Fear-avoidance model of chronic pain: the next generation. *Clin J Pain* 2012 Jul;28(6):475-483. [doi: [10.1097/AJP.0b013e3182385392](https://doi.org/10.1097/AJP.0b013e3182385392)] [Medline: [22673479](https://pubmed.ncbi.nlm.nih.gov/22673479/)]
43. Peng K, Moussavi Z, Karunakaran KD, Borsook D, Lesage F, Nguyen DK. iVR-fNIRS: studying brain functions in a fully immersive virtual environment. *Neurophotonics* 2024 Apr;11(2):020601. [doi: [10.1117/1.NPh.11.2.020601](https://doi.org/10.1117/1.NPh.11.2.020601)] [Medline: [38577629](https://pubmed.ncbi.nlm.nih.gov/38577629/)]
44. Ansado J, Chasen C, Bouchard S, Northoff G. How brain imaging provides predictive biomarkers for therapeutic success in the context of virtual reality cognitive training. *Neurosci Biobehav Rev* 2021 Jan;120:583-594. [doi: [10.1016/j.neubiorev.2020.05.018](https://doi.org/10.1016/j.neubiorev.2020.05.018)] [Medline: [32533997](https://pubmed.ncbi.nlm.nih.gov/32533997/)]
45. Phillips V Z, Canoy RJ, Paik SH, Lee SH, Kim BM. Functional near-infrared spectroscopy as a personalized digital healthcare tool for brain monitoring. *J Clin Neurol* 2023 Mar;19(2):115-124. [doi: [10.3988/jcn.2022.0406](https://doi.org/10.3988/jcn.2022.0406)] [Medline: [36854332](https://pubmed.ncbi.nlm.nih.gov/36854332/)]
46. Hoshi Y. Functional near-infrared spectroscopy: current status and future prospects. *J Biomed Opt* 2007;12(6):062106. [doi: [10.1117/1.2804911](https://doi.org/10.1117/1.2804911)] [Medline: [18163809](https://pubmed.ncbi.nlm.nih.gov/18163809/)]
47. Luo Y, Du J, Yu H, Fang F, Shi P. Resting-state fNIRS reveals changes in prefrontal cortex functional connectivity during TENS in patients with chronic pain. *Sci Rep* 2024 Nov 25;14(1):29187. [doi: [10.1038/s41598-024-79820-2](https://doi.org/10.1038/s41598-024-79820-2)] [Medline: [39587185](https://pubmed.ncbi.nlm.nih.gov/39587185/)]
48. Hall M, Kidgell D, Perraton L, Morrissey J, Jaberzadeh S. Pain induced changes in brain oxyhemoglobin: a systematic review and meta-analysis of functional NIRS studies. *Pain Med* 2021 Jun 4;22(6):1399-1410. [doi: [10.1093/pm/pnaa453](https://doi.org/10.1093/pm/pnaa453)] [Medline: [33659994](https://pubmed.ncbi.nlm.nih.gov/33659994/)]
49. Deng X, Jian C, Yang Q, Jiang N, Huang Z, Zhao S. The analgesic effect of different interactive modes of virtual reality: a prospective functional near-infrared spectroscopy (fNIRS) study. *Front Neurosci* 2022;16:1033155. [doi: [10.3389/fnins.2022.1033155](https://doi.org/10.3389/fnins.2022.1033155)] [Medline: [36458040](https://pubmed.ncbi.nlm.nih.gov/36458040/)]

50. Hu XS, Beard K, Sherbel MC, et al. Brain mechanisms of virtual reality breathing versus traditional mindful breathing in pain modulation: observational functional near-infrared spectroscopy study. *J Med Internet Res* 2021 Oct 12;23(10):e27298. [doi: [10.2196/27298](https://doi.org/10.2196/27298)] [Medline: [34636731](https://pubmed.ncbi.nlm.nih.gov/34636731/)]
51. Garcia LM, Birkhead BJ, Krishnamurthy P, et al. An 8-week self-administered at-home behavioral skills-based virtual reality program for chronic low back pain: double-blind, randomized, placebo-controlled trial conducted during COVID-19. *J Med Internet Res* 2021 Feb 22;23(2):e26292. [doi: [10.2196/26292](https://doi.org/10.2196/26292)] [Medline: [33484240](https://pubmed.ncbi.nlm.nih.gov/33484240/)]
52. Darnall BD, Krishnamurthy P, Tsuei J, Minor JD. Self-administered skills-based virtual reality intervention for chronic pain: randomized controlled pilot study. *JMIR Form Res* 2020 Jul 7;4(7):e17293. [doi: [10.2196/17293](https://doi.org/10.2196/17293)] [Medline: [32374272](https://pubmed.ncbi.nlm.nih.gov/32374272/)]
53. Maddox T, Sparks C, Oldstone L, et al. Durable chronic low back pain reductions up to 24 months after treatment for an accessible, 8-week, in-home behavioral skills-based virtual reality program: a randomized controlled trial. *Pain Med* 2023 Oct 3;24(10):1200-1203. [doi: [10.1093/pm/pnad070](https://doi.org/10.1093/pm/pnad070)] [Medline: [37220894](https://pubmed.ncbi.nlm.nih.gov/37220894/)]
54. 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. [doi: [10.2196/jmir.1923](https://doi.org/10.2196/jmir.1923)] [Medline: [22209829](https://pubmed.ncbi.nlm.nih.gov/22209829/)]
55. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. *J Eval Clin Pract* 2004 May;10(2):307-312. [doi: [10.1111/j.2002.384.doc.x](https://doi.org/10.1111/j.2002.384.doc.x)] [Medline: [15189396](https://pubmed.ncbi.nlm.nih.gov/15189396/)]
56. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *J Psychiatr Res* 2011 May;45(5):626-629. [doi: [10.1016/j.jpsychires.2010.10.008](https://doi.org/10.1016/j.jpsychires.2010.10.008)] [Medline: [21035130](https://pubmed.ncbi.nlm.nih.gov/21035130/)]
57. NCCIH research framework. NCCIH. URL: <https://www.nccih.nih.gov/grants/nccih-clinical-trials-funding-opportunities/research-framework> [accessed 2025-10-16]
58. Nielsen L, Riddle M, King JW, et al. The NIH science of behavior change program: transforming the science through a focus on mechanisms of change. *Behav Res Ther* 2018 Feb;101:3-11. [doi: [10.1016/j.brat.2017.07.002](https://doi.org/10.1016/j.brat.2017.07.002)] [Medline: [29110885](https://pubmed.ncbi.nlm.nih.gov/29110885/)]
59. Lester EG, Hopkins SW, Popok PJ, Vranceanu AM. Adaptation of a live video mind-body program to a web-based platform for english-speaking adults with neurofibromatosis: protocol for the NF-web study. *JMIR Res Protoc* 2021 Jun 10;10(6):e27526. [doi: [10.2196/27526](https://doi.org/10.2196/27526)] [Medline: [34110294](https://pubmed.ncbi.nlm.nih.gov/34110294/)]
60. Greenberg J, Lin A, Zale EL, et al. Development and early feasibility testing of a mind-body physical activity program for patients with heterogeneous chronic pain; the GetActive study. *J Pain Res* 2019;12:3279-3297. [doi: [10.2147/JPR.S222448](https://doi.org/10.2147/JPR.S222448)] [Medline: [31849515](https://pubmed.ncbi.nlm.nih.gov/31849515/)]
61. Jacobs CA, Mace RA, Greenberg J, et al. Development of a mind body program for obese knee osteoarthritis patients with comorbid depression. *Contemp Clin Trials Commun* 2021 Mar;21(January):100720. [doi: [10.1016/j.conctc.2021.100720](https://doi.org/10.1016/j.conctc.2021.100720)] [Medline: [33553798](https://pubmed.ncbi.nlm.nih.gov/33553798/)]
62. Lester E, DiStefano S, Mace R, Macklin E, Plotkin S, Vranceanu AM. Virtual mind-body treatment for geographically diverse youth with neurofibromatosis: a pilot randomized controlled trial. *Gen Hosp Psychiatry* 2020;62:72-78. [doi: [10.1016/j.genhosppsy.2019.12.001](https://doi.org/10.1016/j.genhosppsy.2019.12.001)] [Medline: [31841875](https://pubmed.ncbi.nlm.nih.gov/31841875/)]
63. Mace RA, Greenberg J, Stauder M, Reynolds G, Vranceanu AM. My healthy brain: a multimodal lifestyle program to promote brain health. *Aging Ment Health* 2022 May;26(5):980-991. [doi: [10.1080/13607863.2021.1904828](https://doi.org/10.1080/13607863.2021.1904828)] [Medline: [33784902](https://pubmed.ncbi.nlm.nih.gov/33784902/)]
64. Mace RA, Doorley JD, Popok PJ, Vranceanu AM. Live video adaptations to a mind-body activity program for chronic pain and cognitive decline: protocol for the virtual active brains study. *JMIR Res Protoc* 2021 Jan 4;10(1):e25351. [doi: [10.2196/25351](https://doi.org/10.2196/25351)] [Medline: [33208301](https://pubmed.ncbi.nlm.nih.gov/33208301/)]
65. Marin TJ, Van Eerd D, Irvin E, et al. Multidisciplinary biopsychosocial rehabilitation for subacute low back pain. *Cochrane Database Syst Rev* 2017 Jun 28;6(6):CD002193. [doi: [10.1002/14651858.CD002193.pub2](https://doi.org/10.1002/14651858.CD002193.pub2)] [Medline: [28656659](https://pubmed.ncbi.nlm.nih.gov/28656659/)]
66. Bussi eres AE, Stewart G, Al-Zoubi F, et al. The treatment of neck pain-associated disorders and whiplash-associated disorders: a clinical practice guideline. *J Manipulative Physiol Ther* 2016 Oct;39(8):523-564. [doi: [10.1016/j.jmpt.2016.08.007](https://doi.org/10.1016/j.jmpt.2016.08.007)] [Medline: [27836071](https://pubmed.ncbi.nlm.nih.gov/27836071/)]
67. Vranceanu AM, Jacobs C, Lin A, et al. Results of a feasibility randomized controlled trial (RCT) of the toolkit for optimal recovery (TOR): a live video program to prevent chronic pain in at-risk adults with orthopedic injuries. *Pilot Feasibility Stud* 2019;5(30):30. [doi: [10.1186/s40814-019-0416-7](https://doi.org/10.1186/s40814-019-0416-7)] [Medline: [30820341](https://pubmed.ncbi.nlm.nih.gov/30820341/)]
68. Garcia LM, Darnall BD, Krishnamurthy P, et al. Self-administered behavioral skills-based at-home virtual reality therapy for chronic low back pain: protocol for a randomized controlled trial. *JMIR Res Protoc* 2021 Jan 19;10(1):e25291. [doi: [10.2196/25291](https://doi.org/10.2196/25291)] [Medline: [33464215](https://pubmed.ncbi.nlm.nih.gov/33464215/)]
69. Alghadir AH, Anwer S, Iqbal A, Iqbal ZA. Test-retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. *J Pain Res* 2018;11:851-856. [doi: [10.2147/JPR.S158847](https://doi.org/10.2147/JPR.S158847)] [Medline: [29731662](https://pubmed.ncbi.nlm.nih.gov/29731662/)]
70. Buckenmaier CC, Galloway KT, Polomano RC, McDuffie M, Kwon N, Gallagher RM. Preliminary validation of the defense and veterans pain rating scale (DVPRS) in a military population. *Pain Med* 2013 Jan;14(1):110-123. [doi: [10.1111/j.1526-4637.2012.01516.x](https://doi.org/10.1111/j.1526-4637.2012.01516.x)] [Medline: [23137169](https://pubmed.ncbi.nlm.nih.gov/23137169/)]

71. Polomano RC, Galloway KT, Kent ML, et al. Psychometric testing of the defense and veterans pain rating scale (DVPRS): a new pain scale for military population. *Pain Med* 2016 Aug;17(8):1505-1519. [doi: [10.1093/pm/pnw105](https://doi.org/10.1093/pm/pnw105)] [Medline: [27272528](https://pubmed.ncbi.nlm.nih.gov/27272528/)]
72. Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess* 1995 Dec;7(4):524-532. [doi: [10.1037/1040-3590.7.4.524](https://doi.org/10.1037/1040-3590.7.4.524)]
73. Bascour-Sandoval C, Albayay J, Martínez-Molina A, et al. Psychometric properties of the PCS and the PCS-4 in individuals with musculoskeletal pain. *Psicothema* 2022 May;34(2):323-331. [doi: [10.7334/psicothema2021.21](https://doi.org/10.7334/psicothema2021.21)] [Medline: [35485547](https://pubmed.ncbi.nlm.nih.gov/35485547/)]
74. McCracken LM, Dhingra L. A short version of the pain anxiety symptoms scale (PASS-20): preliminary development and validity. *Pain Res Manag* 2002;7(1):45-50. [doi: [10.1155/2002/517163](https://doi.org/10.1155/2002/517163)] [Medline: [16231066](https://pubmed.ncbi.nlm.nih.gov/16231066/)]
75. Roelofs J, McCracken L, Peters ML, Crombez G, van Breukelen G, Vlaeyen JW. Psychometric evaluation of the pain anxiety symptoms scale (PASS) in chronic pain patients. *J Behav Med* 2004 Apr;27(2):167-183. [doi: [10.1023/b:jobm.0000019850.51400.a6](https://doi.org/10.1023/b:jobm.0000019850.51400.a6)] [Medline: [15171105](https://pubmed.ncbi.nlm.nih.gov/15171105/)]
76. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. *Eur J Pain* 2007 Feb;11(2):153-163. [doi: [10.1016/j.ejpain.2005.12.008](https://doi.org/10.1016/j.ejpain.2005.12.008)] [Medline: [16446108](https://pubmed.ncbi.nlm.nih.gov/16446108/)]
77. Feldman G, Hayes A, Kumar S, Greeson J, Laurenceau JP. Mindfulness and emotion regulation: the development and initial validation of the cognitive and affective mindfulness scale-revised (CAMS-R). *J Psychopathol Behav Assess* 2007 Jul 13;29(3):177-190. [doi: [10.1007/s10862-006-9035-8](https://doi.org/10.1007/s10862-006-9035-8)]
78. Doorley JD, Mace RA, Popok PJ, Grunberg VA, Ragnhildstveit A, Vranceanu AM. Feasibility randomized controlled trial of a mind-body activity program for older adults with chronic pain and cognitive decline: the virtual “Active Brains” study. *Gerontologist* 2022 Aug 12;62(7):1082-1094. [doi: [10.1093/geront/gnab135](https://doi.org/10.1093/geront/gnab135)] [Medline: [34487167](https://pubmed.ncbi.nlm.nih.gov/34487167/)]
79. Vranceanu AM, Bakhshaiie J, Reichman M, et al. A live video program to prevent chronic pain and disability in at-risk adults with acute orthopedic injuries (toolkit for optimal recovery): protocol for a multisite feasibility study. *JMIR Res Protoc* 2021 Apr 28;10(4):e28155. [doi: [10.2196/28155](https://doi.org/10.2196/28155)] [Medline: [33908886](https://pubmed.ncbi.nlm.nih.gov/33908886/)]
80. Antoni MH, Lechner SC, Kazi A, et al. How stress management improves quality of life after treatment for breast cancer. *J Consult Clin Psychol* 2006;74(6):1143-1152. [doi: [10.1037/0022-006X.74.6.1143](https://doi.org/10.1037/0022-006X.74.6.1143)] [Medline: [17154743](https://pubmed.ncbi.nlm.nih.gov/17154743/)]
81. OctaMon user manual. : Artinis Medical Systems; 2007.
82. Berry MP, Lutz J, Schuman-Olivier Z, et al. Brief self-compassion training alters neural responses to evoked pain for chronic low back pain: a pilot study. *Pain Med* 2020 Oct 1;21(10):2172-2185. [doi: [10.1093/pm/pnaa178](https://doi.org/10.1093/pm/pnaa178)] [Medline: [32783054](https://pubmed.ncbi.nlm.nih.gov/32783054/)]
83. Hanks J, Myers B. Validity, reliability, and efficiency of a standard goniometer, medical inclinometer, and builder’s inclinometer. *Int J Sports Phys Ther* 2023;18(4):989-996. [doi: [10.26603/001c.83944](https://doi.org/10.26603/001c.83944)] [Medline: [37547826](https://pubmed.ncbi.nlm.nih.gov/37547826/)]
84. Vambheim SM, Kyllö TM, Hegland S, Bystad M. Relaxation techniques as an intervention for chronic pain: a systematic review of randomized controlled trials. *Heliyon* 2021 Aug;7(8):e07837. [doi: [10.1016/j.heliyon.2021.e07837](https://doi.org/10.1016/j.heliyon.2021.e07837)] [Medline: [34485731](https://pubmed.ncbi.nlm.nih.gov/34485731/)]
85. Gold JL, Belmont KA, Thomas DA. The neurobiology of virtual reality pain attenuation. *Cyberpsychol Behav* 2007 Aug;10(4):536-544. [doi: [10.1089/cpb.2007.9993](https://doi.org/10.1089/cpb.2007.9993)] [Medline: [17711362](https://pubmed.ncbi.nlm.nih.gov/17711362/)]
86. Loreto-Quijada D, Gutiérrez-Maldonado J, Nieto R, et al. Differential effects of two virtual reality interventions: distraction versus pain control. *Cyberpsychol Behav Soc Netw* 2014 Jun;17(6):353-358. [doi: [10.1089/cyber.2014.0057](https://doi.org/10.1089/cyber.2014.0057)] [Medline: [24892197](https://pubmed.ncbi.nlm.nih.gov/24892197/)]
87. Huppert TJ, Diamond SG, Franceschini MA, Boas DA. HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Appl Opt* 2009 Apr 1;48(10):D280-D298. [doi: [10.1364/ao.48.00d280](https://doi.org/10.1364/ao.48.00d280)] [Medline: [19340120](https://pubmed.ncbi.nlm.nih.gov/19340120/)]
88. OxySoft. Artinis Medical Systems. URL: <https://www.artinis.com/oxysoft> [accessed 2025-12-04]
89. Hocke LM, Oni IK, Duszynski CC, Corrigan AV, Frederick BD, Dunn JF. Automated processing of fNIRS data—a visual guide to the pitfalls and consequences. *Algorithms* 2018 May;11(5):67. [doi: [10.3390/a11050067](https://doi.org/10.3390/a11050067)] [Medline: [30906511](https://pubmed.ncbi.nlm.nih.gov/30906511/)]
90. Kinder KT, Heim HLR, Parker J, et al. Systematic review of fNIRS studies reveals inconsistent chromophore data reporting practices. *Neurophotonics* 2022 Oct;9(4):040601. [doi: [10.1117/1.NPh.9.4.040601](https://doi.org/10.1117/1.NPh.9.4.040601)] [Medline: [36578778](https://pubmed.ncbi.nlm.nih.gov/36578778/)]
91. von Lüthmann A, Ortega-Martinez A, Boas DA, Yücel MA. Using the general linear model to improve performance in fNIRS single trial analysis and classification: a perspective. *Front Hum Neurosci* 2020;14(30):30. [doi: [10.3389/fnhum.2020.00030](https://doi.org/10.3389/fnhum.2020.00030)] [Medline: [32132909](https://pubmed.ncbi.nlm.nih.gov/32132909/)]
92. Pearson K. VII. Note on regression and inheritance in the case of two parents. *Proc R Soc Lond* 1895 Dec 31;58(347-352):240-242. [doi: [10.1098/rsp.1895.0041](https://doi.org/10.1098/rsp.1895.0041)]
93. Vindrola-Padros C, Johnson GA. Rapid techniques in qualitative research: a critical review of the literature. *Qual Health Res* 2020 Aug;30(10):1596-1604. [doi: [10.1177/1049732320921835](https://doi.org/10.1177/1049732320921835)] [Medline: [32667277](https://pubmed.ncbi.nlm.nih.gov/32667277/)]
94. Gale RC, Wu J, Erhardt T, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the veterans health administration. *Implement Sci* 2019 Feb 1;14(1):11. [doi: [10.1186/s13012-019-0853-y](https://doi.org/10.1186/s13012-019-0853-y)] [Medline: [30709368](https://pubmed.ncbi.nlm.nih.gov/30709368/)]
95. Renfro CP, Rome Z, Gatwood J, Hohmeier KC. Use of rapid assessment procedures when analyzing qualitative data in pharmacy research. *Res Social Adm Pharm* 2022 Jan;18(1):2249-2253. [doi: [10.1016/j.sapharm.2021.05.013](https://doi.org/10.1016/j.sapharm.2021.05.013)] [Medline: [34116965](https://pubmed.ncbi.nlm.nih.gov/34116965/)]

96. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. *Psychiatry Res* 2019 Oct;280:112516. [doi: [10.1016/j.psychres.2019.112516](https://doi.org/10.1016/j.psychres.2019.112516)] [Medline: [31437661](https://pubmed.ncbi.nlm.nih.gov/31437661/)]
97. Taylor B, Henshall C, Kenyon S, Litchfield I, Greenfield S. Can rapid approaches to qualitative analysis deliver timely, valid findings to clinical leaders? a mixed methods study comparing rapid and thematic analysis. *BMJ Open* 2018 Oct 8;8(10):e019993. [doi: [10.1136/bmjopen-2017-019993](https://doi.org/10.1136/bmjopen-2017-019993)] [Medline: [30297341](https://pubmed.ncbi.nlm.nih.gov/30297341/)]
98. Averill JB. Matrix analysis as a complementary analytic strategy in qualitative inquiry. *Qual Health Res* 2002 Jul;12(6):855-866. [doi: [10.1177/104973230201200611](https://doi.org/10.1177/104973230201200611)] [Medline: [12109729](https://pubmed.ncbi.nlm.nih.gov/12109729/)]
99. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995 Jan 1;57(1):289-300. [doi: [10.1111/j.2517-6161.1995.tb02031.x](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x)]
100. Ochsner KN, Silvers JA, Buhle JT. Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Ann N Y Acad Sci* 2012 Mar;1251(1):E1-24. [doi: [10.1111/j.1749-6632.2012.06751.x](https://doi.org/10.1111/j.1749-6632.2012.06751.x)] [Medline: [23025352](https://pubmed.ncbi.nlm.nih.gov/23025352/)]
101. Baliki MN, Geha PY, Apkarian AV, Chialvo DR. Beyond feeling: chronic pain hurts the brain, disrupting the default-mode network dynamics. *J Neurosci* 2008 Feb 6;28(6):1398-1403. [doi: [10.1523/JNEUROSCI.4123-07.2008](https://doi.org/10.1523/JNEUROSCI.4123-07.2008)] [Medline: [18256259](https://pubmed.ncbi.nlm.nih.gov/18256259/)]
102. Zeidan F, Martucci KT, Kraft RA, McHaffie JG, Coghill RC. Neural correlates of mindfulness meditation-related anxiety relief. *Soc Cogn Affect Neurosci* 2014 Jun;9(6):751-759. [doi: [10.1093/scan/nst041](https://doi.org/10.1093/scan/nst041)] [Medline: [23615765](https://pubmed.ncbi.nlm.nih.gov/23615765/)]
103. Jensen KB, Kosek E, Wicksell R, et al. Cognitive behavioral therapy increases pain-evoked activation of the prefrontal cortex in patients with fibromyalgia. *Pain* 2012 Jul;153(7):1495-1503. [doi: [10.1016/j.pain.2012.04.010](https://doi.org/10.1016/j.pain.2012.04.010)] [Medline: [22617632](https://pubmed.ncbi.nlm.nih.gov/22617632/)]
104. Schweinhardt P, Seminowicz DA, Jaeger E, Duncan GH, Bushnell MC. The anatomy of the mesolimbic reward system: a link between personality and the placebo analgesic response. *J Neurosci* 2009 Apr 15;29(15):4882-4887. [doi: [10.1523/JNEUROSCI.5634-08.2009](https://doi.org/10.1523/JNEUROSCI.5634-08.2009)] [Medline: [19369556](https://pubmed.ncbi.nlm.nih.gov/19369556/)]
105. Kucyi A, Salomons TV, Davis KD. Mind wandering away from pain dynamically engages antinociceptive and default mode brain networks. *Proc Natl Acad Sci U S A* 2013 Nov 12;110(46):18692-18697. [doi: [10.1073/pnas.1312902110](https://doi.org/10.1073/pnas.1312902110)] [Medline: [24167282](https://pubmed.ncbi.nlm.nih.gov/24167282/)]
106. Sharar SR, Carrougier GJ, Nakamura D, Hoffman HG, Blough DK, Patterson DR. Factors influencing the efficacy of virtual reality distraction analgesia during postburn physical therapy: preliminary results from 3 ongoing studies. *Arch Phys Med Rehabil* 2007 Dec;88(12 Suppl 2):S43-S49. [doi: [10.1016/j.apmr.2007.09.004](https://doi.org/10.1016/j.apmr.2007.09.004)] [Medline: [18036981](https://pubmed.ncbi.nlm.nih.gov/18036981/)]
107. Li A, Montañó Z, Chen VJ, Gold JI. Virtual reality and pain management: current trends and future directions. *Pain Manag* 2011 Mar;1(2):147-157. [doi: [10.2217/pmt.10.15](https://doi.org/10.2217/pmt.10.15)] [Medline: [21779307](https://pubmed.ncbi.nlm.nih.gov/21779307/)]
108. Brewer JA, Worhunsky PD, Gray JR, Tang YY, Weber J, Kober H. Meditation experience is associated with differences in default mode network activity and connectivity. *Proc Natl Acad Sci U S A* 2011 Dec 13;108(50):20254-20259. [doi: [10.1073/pnas.1112029108](https://doi.org/10.1073/pnas.1112029108)] [Medline: [22114193](https://pubmed.ncbi.nlm.nih.gov/22114193/)]
109. Taren AA, Gianaros PJ, Greco CM, et al. Mindfulness meditation training alters stress-related amygdala resting state functional connectivity: a randomized controlled trial. *Soc Cogn Affect Neurosci* 2015 Dec;10(12):1758-1768. [doi: [10.1093/scan/nsv066](https://doi.org/10.1093/scan/nsv066)] [Medline: [26048176](https://pubmed.ncbi.nlm.nih.gov/26048176/)]
110. Tang YY, Hölzel BK, Posner MI. The neuroscience of mindfulness meditation. *Nat Rev Neurosci* 2015 Apr;16(4):213-225. [doi: [10.1038/nrn3916](https://doi.org/10.1038/nrn3916)] [Medline: [25783612](https://pubmed.ncbi.nlm.nih.gov/25783612/)]
111. Tracey I, Mantyh PW. The cerebral signature for pain perception and its modulation. *Neuron* 2007 Aug 2;55(3):377-391. [doi: [10.1016/j.neuron.2007.07.012](https://doi.org/10.1016/j.neuron.2007.07.012)] [Medline: [17678852](https://pubmed.ncbi.nlm.nih.gov/17678852/)]
112. Zywił MG, Mahomed A, Gandhi R, Perruccio AV, Mahomed NN. Measuring expectations in orthopaedic surgery: a systematic review. *Clin Orthop Relat Res* 2013 Nov;471(11):3446-3456. [doi: [10.1007/s11999-013-3013-8](https://doi.org/10.1007/s11999-013-3013-8)] [Medline: [23633186](https://pubmed.ncbi.nlm.nih.gov/23633186/)]
113. Castillo RC, Huang Y, Scharfstein D, et al. Association between 6-week postdischarge risk classification and 12-month outcomes after orthopedic trauma. *JAMA Surg* 2019 Feb 1;154(2):e184824. [doi: [10.1001/jamasurg.2018.4824](https://doi.org/10.1001/jamasurg.2018.4824)] [Medline: [30566192](https://pubmed.ncbi.nlm.nih.gov/30566192/)]
114. Brooks JL. Counterbalancing for serial order carryover effects in experimental condition orders. *Psychol Methods* 2012 Dec;17(4):600-614. [doi: [10.1037/a0029310](https://doi.org/10.1037/a0029310)] [Medline: [22799624](https://pubmed.ncbi.nlm.nih.gov/22799624/)]
115. Kaimal G, Carroll-Haskins K, Topoglu Y, Ramakrishnan A, Arslanbek A, Ayaz H. Exploratory fNIRS assessment of differences in activation in virtual reality visual self-expression including with a fragrance stimulus. *Art Ther (Alex)* 2022 Jul 3;39(3):128-137. [doi: [10.1080/07421656.2021.1957341](https://doi.org/10.1080/07421656.2021.1957341)]
116. Mirelman A, Maidan I, Bernad-Elazari H, et al. Increased frontal brain activation during walking while dual tasking: an fNIRS study in healthy young adults. *J Neuroeng Rehabil* 2014 May 12;11(1):85. [doi: [10.1186/1743-0003-11-85](https://doi.org/10.1186/1743-0003-11-85)] [Medline: [24886198](https://pubmed.ncbi.nlm.nih.gov/24886198/)]
117. Rahimpour Jounghani A, Kumar A, Moreno Carbonell L, et al. Wearable fNIRS platform for dense sampling and precision functional neuroimaging. *NPJ Digit Med* 2025 May 13;8(1):271. [doi: [10.1038/s41746-025-01690-3](https://doi.org/10.1038/s41746-025-01690-3)] [Medline: [40355575](https://pubmed.ncbi.nlm.nih.gov/40355575/)]

Abbreviations

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth

dIPFC: dorsolateral prefrontal cortex

DMN: default-mode network

FC: functional coactivation

fNIRS: functional near-infrared spectroscopy

HbO: oxygenated hemoglobin

MD: mean difference

mPFC: medial PFC

PCS: Pain Catastrophizing Scale

PFC: prefrontal cortex

VR: virtual reality

Edited by F Spyridonis; submitted 22.May.2025; peer-reviewed by E Güleç, K Krishnamurthy, VG Prabhu; revised version received 18.Feb.2026; accepted 24.Feb.2026; published 30.Mar.2026.

Please cite as:

Mace RA, Wu Z, Sieberg CB, Karunakaran KD, Peng K, Borsook D

Virtual Reality–Based Pain Modulation in Subacute Musculoskeletal Injury: Functional Near-Infrared Spectroscopy Study of Neural and Behavioral Correlates

JMIR Serious Games 2026;14:e77713

URL: <https://games.jmir.org/2026/1/e77713>

doi: [10.2196/77713](https://doi.org/10.2196/77713)

© Ryan Andrew Mace, Ziyang Wu, Christine B Sieberg, Keerthana Deepti Karunakaran, Ke Peng, David Borsook. Originally published in JMIR Serious Games (<https://games.jmir.org>), 30.Mar.2026. 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.

Design Requirements for Gamified Exercise Apps for Adults With Prehypertension Based on the Octalysis Framework and Self-Determination Theory: Qualitative Interview Study

Sijia Li^{1*}, MMed; Genying Liu^{2*}, MBA; Li Qiao¹, PhD; Hong-Yu Yu³, PhD; Yu-Yang Zhang¹, BSN; Runyuan Pei¹, PhD; Yue-Tao Lin⁴, BSN; Hao-Ming Ma¹, PhD; Mei-Hua Piao¹, PhD

¹School of Nursing, Chinese Academy of Medical Sciences & Peking Union Medical College, 33 Badachu Road, Shijingshan District, Beijing, China

²Outpatient Department, West China Hospital, Sichuan University, Chengdu, China

³School of Nursing, University of Washington, Seattle, WA, United States

⁴Comprehensive Disease Center, West China Xiamen Hospital of Sichuan University, Fujian, China

*these authors contributed equally

Corresponding Author:

Mei-Hua Piao, PhD

School of Nursing, Chinese Academy of Medical Sciences & Peking Union Medical College, 33 Badachu Road, Shijingshan District, Beijing, China

Abstract

Background: Managing blood pressure (BP) in prehypertensive individuals is crucial to prevent the incidence of hypertension. While physical activity has proven effective in BP management, physical inactivity remains prevalent. Gamification has shown promise in addressing physical inactivity; however, its effectiveness is limited due to the suboptimal intervention design.

Objective: This study aimed to develop a comprehensive understanding of prehypertensive individuals' needs, design preferences, and motivational drives for gamified exercise apps by integrating the Octalysis Gamification Framework with Self-Determination Theory (SDT) through a participatory qualitative methodology.

Methods: The study adopted a cross-sectional qualitative interview study design and used semistructured interviews conducted from June to July 2025 across mainland China. Participants were recruited through online platforms using purposive sampling to select adults with prehypertension (systolic BP 120 - 139 mm Hg and/or diastolic BP 80 - 89 mm Hg). Interview guides were systematically structured around SDT constructs and the Octalysis Framework's 8 core gamification drives to explore participants' needs, design preferences, and motivational drives for gamified exercise apps. An inductive or deductive hybrid thematic analysis was used to identify key themes.

Results: A total of 14 eligible prehypertensive individuals participated in the study. Their needs and preferences for exercise apps, including exercise guidance, data monitoring and feedback, and wearable device portability, were summarized. Eight core gamification motivational drives, such as epic meaning and social influence, according to the Octalysis Framework, were explored with interviewees, including their attitudes and creative design considerations. Moreover, the study examined how the satisfaction of basic psychological needs (eg, sense of volition, technical challenges, and social connections) influences the transformation of motivation.

Conclusions: This theory-informed qualitative study is among the first to explore the needs and preferences of individuals with prehypertension for gamified exercise apps by integrating the Octalysis Gamification Framework with SDT. The findings suggest that successful gamified exercise apps for prehypertensive individuals may benefit from going beyond traditional game mechanics to address deeper psychological needs. Participants emphasized the importance of personalized exercise programming, robust health monitoring capabilities, and human-centered gamification designs that support autonomy, competence, and relatedness. Satisfaction of these basic psychological needs is critical for transforming extrinsic motivations, such as weight management and health values, into intrinsic enjoyment of exercise. By integrating a theoretical basis with user-centered perspectives, this study provides context-specific design implications for future development of gamified exercise apps tailored to prehypertensive individuals in similar digital health contexts.

(*JMIR Serious Games* 2026;14:e86793) doi:[10.2196/86793](https://doi.org/10.2196/86793)

KEYWORDS

prehypertension; gamification; mHealth; qualitative study; self-determination theory; mobile health

Introduction

Hypertension is the leading risk factor for global disease burden [1]. The World Health Organization (WHO) estimated that 1.28 billion adults aged 30 - 79 years are affected by hypertension worldwide. Additionally, the WHO aims to reduce the prevalence of hypertension by 33% from 2010 to 2030, a goal that necessitates effective hypertension prevention strategies [2]. Data from a randomized controlled trial indicate that two-thirds of untreated prehypertension cases can progress to hypertension within 4 years [3]. Managing blood pressure (BP) in prehypertensive individuals is crucial to prevent the incidence of hypertension. Lifestyle interventions, such as adopting physical activity (PA) programs, are strongly recommended in clinical guidelines for managing BP among individuals with prehypertension [4]. In addition, a previous clinical trial involving prehypertensive individuals found that Tai Chi reduced systolic BP by 7.01 mm Hg and diastolic BP by 3.73 mm Hg, demonstrating notable effectiveness [5]. Although PA has shown effectiveness in BP management in both clinical guidelines and randomized controlled trials, physical inactivity remains prevalent, with one-third of adults being insufficiently physically active in 2022 [6]. This disconnect between evidence and practice presents a significant challenge in cardiovascular disease prevention.

Digital health technologies offer scalable solutions to this challenge. Gamification—the strategic integration of game elements such as points, levels, badges, and avatars into nongame contexts such as health care, education, and training—has shown promise in addressing physical inactivity [7]. The meta-analysis demonstrated a small to medium effect (Hedges $g=0.42$) of gamification on PA, though notably, long-term effects diminish significantly, and current interventions have not translated into measurable health benefits [8,9]. This limited effectiveness stems from 3 critical gaps: interventions often lack robust theoretical foundations, fail to incorporate user-centered design principles, and neglect the psychological mechanisms underlying sustained behavior change [10,11]. Therefore, given the current limitations of gamification intervention design and its limited effectiveness, our team is developing a gamified exercise app that is theory-driven, evidence-based, and incorporates user perspectives. This qualitative study, based on user interviews, is part of the broader research topic. This interview aimed to understand the needs, design preferences, and motivations of prehypertensive individuals for gamified exercise apps.

In contrast to prior approaches that rely on either gamification frameworks or psychological theories in isolation, we integrated both to provide a deeper understanding of how gamification elements relate to users' psychological need satisfaction and support health behavior change [12,13].

First, we selected the Octalysis Gamification Framework as our primary theoretical basis for analyzing gamification mechanics. Created by Chou in 2019 [14], this human-focused framework is well known for emphasizing motivational drives rather than functional features, identifying 8 core drives that underlie user motivation in gamified systems. However, while Octalysis

explains what motivates users, it provides limited insight into how these motivations translate into sustained behavioral engagement.

To address this, we integrated Self-Determination Theory (SDT), a well-established motivational theory, to elucidate the psychological pathways through which external gamification elements may foster intrinsic motivation [15]. SDT offers a theoretical lens for linking gamification drives to the satisfaction of 3 fundamental psychological needs—autonomy, competence, and relatedness—with specific Octalysis core drives (eg, achievement and social influence) mapping to distinct needs (eg, achievement to competence), thereby facilitating the transition from extrinsic to intrinsic motivation.

By combining Octalysis and SDT, this dual-theory approach enables us to bridge the gap between surface-level gamification features and deep psychological engagement [16]. Guided by this, we conducted a user-centered qualitative study to understand the needs of prehypertensive individuals in relation to exercise apps, as well as to explore gamification design considerations.

Specifically, our objectives were to (1) explore the needs and design preferences for exercise apps among individuals with prehypertension, (2) identify optimal gamification design strategies based on the Octalysis Framework from users' perspectives, and (3) elucidate how exercise apps can support motivational transformation in the target population according to SDT. The insights derived from this interview will inform the development of next-generation gamified health interventions and contribute to the broader discourse on theory-informed digital health design.

Methods

Overview

This study used a cross-sectional qualitative interview study design grounded in a pragmatic and interpretive orientation, using semistructured interviews to generate user-centered insights. This approach was chosen to generate practical design implications for gamified exercise apps while facilitating an interpretive understanding of participants' exercise-related experiences. To ensure adherence to qualitative research reporting standards, we followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist (Checklist 1) and the American Psychological Association Journal Article Reporting Standards for Qualitative Research (JARS-Qual) [17,18].

Study Setting and Recruitment

The study was conducted from June to July 2025 through online poster recruitment in mainland China. To ensure the richness and relevance of the findings, purposive sampling was used to select individuals with prehypertension [19]. We aimed to achieve gender balance and include individuals from diverse socioeconomic backgrounds and occupations in our recruitment. The sample size was determined in accordance with the principle of data saturation. Interviews were discontinued after a minimum of 10 sessions, when no new codes or themes emerged after 3 consecutive interviews [20].

Inclusion criteria included the following: (1) aged between 18 and 60 years old; (2) met the 2024 Chinese diagnostic criteria for prehypertension, defined as systolic BP between 120 and 139 mm Hg and/or diastolic BP between 80 and 89 mm Hg [21]; (3) able to communicate verbally in Mandarin and willing to participate voluntarily; (4) expressed an interest in the association between exercise and BP; and (5) had prior experience using smartphone apps. Exclusion criteria were as follows: (1) individuals with any known cognitive impairments or mental illnesses that could affect their ability to participate in an interview and (2) individuals with physical conditions that may impact their PA behavior.

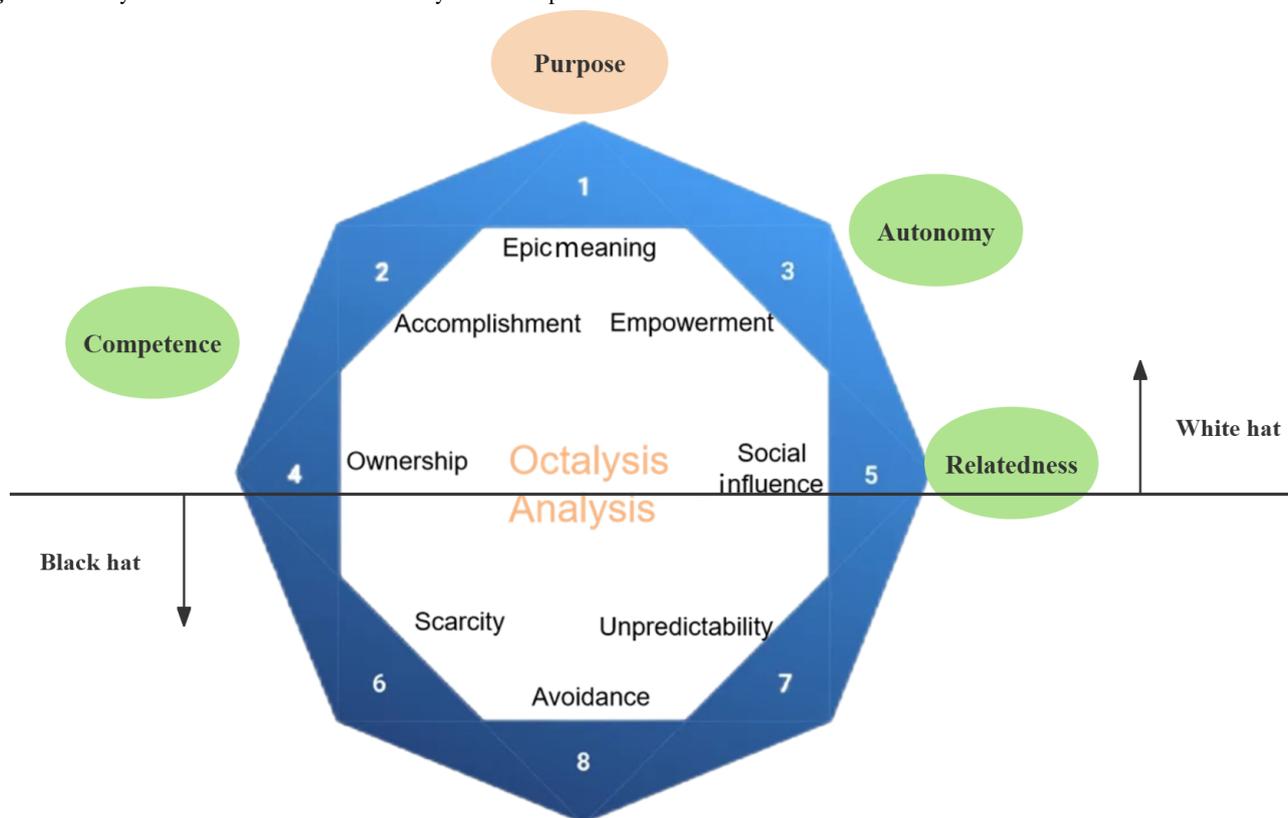
Theoretical Framework

The interview guidelines were systematically developed through integration of SDT and the Octalysis Gamification Framework, encompassing questions regarding 3 aspects: exercise behavior, needs, and preferences for exercise apps, and gamified incentive mechanisms (Multimedia Appendix 1) [14,16]. According to SDT, human motivation undergoes a continuous transition from amotivation to extrinsic motivation and then to intrinsic motivation. When the external environment provides appropriate support and satisfies an individual’s basic psychological needs (autonomy, competency, and relatedness), it can stimulate intrinsic motivation or facilitate the transformation of extrinsic motivation [16,22]. The Octalysis Gamification Framework,

developed by Chou in 2019, has been widely adopted across diverse fields and is recognized for emphasizing human motivation rather than mere functionality to improve the impact of gamification [14,23]. It categorizes game elements into 8 fundamental psychological drives (ie, epic meaning, accomplishment, empowerment, ownership, social influence, scarcity, unpredictability, and avoidance) that shape human behavior and decision-making. These core drives serve to motivate, empower, and influence behavioral patterns, ultimately prompting individuals to take action.

The relationship between the Octalysis Framework and SDT is illustrated in Figure 1. The figure conceptually depicts how selected white hat Octalysis motivational drives may be interpreted through an SDT-informed lens in relation to basic psychological needs of autonomy (empowerment), competence (accomplishment and ownership), and relatedness (social influence), as well as a sense of purpose (epic meaning). Ownership is depicted as having a weaker association with competence, as it often relies on responsibility rather than intrinsic mastery. The horizontal axis distinguishes between predominantly psychological need support (white hat) and more externally driven (black hat) motivation. Black hat drives (scarcity, avoidance, and unpredictability) are shown as operating outside psychological need satisfaction and are therefore not theoretically aligned with SDT.

Figure 1. Octalysis and Self-Determination Theory relationship.



This dual-theory approach enabled us to systematically explore both the functional features of gamification (Octalysis) and its psychological aspects (SDT) in health behavior change.

Data Collection

Prior to data collection, the interview guide was rigorously reviewed by 3 qualitative methodology experts, pilot-tested with 2 prehypertensive individuals, and iteratively finalized. Data were gathered through one-on-one semistructured

interviews, which were conducted by a trained qualitative researcher (SJL), a female master's student with 2 years of experience in health behavior research and a research interest in gamified exercise apps, to ensure consistency. The researcher's prior understanding of behavior change theories and gamification informed the development of the interview guide. Interviews were conducted either through online Tencent Meetings (Tencent) or face-to-face in a quiet meeting room with no other person present, depending on the specific circumstances and participant preferences. Before the interviews, the researcher provided participants with an explanation of the topic, reasons, and purposes for doing the research, and privacy protections to foster trust with the participants. Participants were also invited to complete demographic questionnaires prior to the interview. Those participating in face-to-face interviews completed paper-based questionnaires, whereas those interviewed through Tencent Meetings completed questionnaires through Wenjuanxing (Changsha Ranxing Information Technology Co, Ltd), a widely used online survey platform in China.

During the interview, before discussing gamification, the interviewer introduced the concept to the participants using examples from well-known apps such as Starbucks (Starbucks Corporation) and Duolingo (Duolingo, Inc). Interviews took place in private, comfortable environments, with each session lasting approximately 30 to 60 minutes. After the interviews, the interviewer took notes on key information. With the participant's consent, all sessions were audio-recorded and transcribed verbatim within 24 hours. The researchers then returned the transcriptions to the participants to verify the accuracy of the content.

Data Analysis

The interviewer (SJL) transcribed the interview recordings verbatim using the transcription software iFlytek Hearing. Subsequently, the transcriptions were reviewed and verified for accuracy by listening to the audio recordings again in Microsoft Word, with any necessary corrections made based on the recordings. The finalized transcriptions were then uploaded to NVivo 20 (QSR International) for systematic qualitative data analysis, where coding and thematic analysis were conducted. Thematic analysis is a flexible method that can be applied to both inductive and deductive approaches to data analysis [24]. According to our research aim, we used an inductive or deductive hybrid thematic analysis, a methodologically rigorous approach that combines both theory-driven and data-driven approaches to identify key themes [25]. Specifically, inductive thematic analysis was used to explore the needs and preferences of prehypertensive individuals regarding exercise apps, while the deductive analysis was applied using predetermined codes derived from SDT and the Octalysis Gamification Framework. For example, SDT-informed deductive codes included autonomy (eg, sense of volition and preferences), while Octalysis-informed codes included motivational drives such as accomplishment (eg, points and progress bars).

Two researchers (SJL and HMM) with expertise in qualitative methods and health gamification independently coded the first 6 transcripts, then compared and discussed their results to establish preliminary themes and insights. In cases of

disagreement, a third researcher (MHP) with extensive experience in behavioral health research was consulted to reach a consensus. The interviewer then proceeded to code the remaining transcripts, collaborating with the second researcher to incorporate any newly emerging themes. Finally, the entire research team reviewed and finalized the themes through iterative discussion until consensus was achieved.

Methodological Rigor

Rigor was determined using Lincoln and Guba (1985) 4 criteria of credibility, dependability, confirmability, and transferability [26]. Credibility was established through the use of probing questions during interviews, such as "Can you elaborate?" These questions encouraged participants to provide more detailed responses, allowing researchers to gain a more comprehensive understanding of their experiences. Member checking, where participants were asked to review and confirm the accuracy of the interview data, was also used to ensure that the research findings reflected their true perspectives. An audit trail, which included a detailed record of how data were collected, kept, analyzed, and interpreted, was used to ensure dependability by providing a transparent research process and confirmability by ensuring that the qualitative findings were not influenced by researchers' biases. Additionally, a detailed description of the research aim, participants, and context was provided to facilitate transferability to similar populations and settings.

Methodological Integrity

Fidelity was ensured through consistent alignment among the research objectives, the guiding theoretical frameworks (SDT and the Octalysis Gamification Framework), the semistructured interview design, and the hybrid inductive-deductive thematic analysis. Utility was demonstrated by addressing a meaningful research problem and generating context-specific design implications to inform the future development of gamified exercise apps for individuals with prehypertension. These findings offer actionable insights for researchers and software designers working in digital behavior change contexts.

Ethical Considerations

The study was approved by the Ethics Committee of Peking Union Medical College School of Nursing (approval number: PUMCSON-2024 - 36) before data collection began and was conducted in compliance with the ethical principles outlined in the Declaration of Helsinki. Written informed consent was obtained from all participants after providing comprehensive information about study objectives, procedures, and data protection measures. All data were anonymized prior to analysis, and no information could be used to identify individual participants. Each interviewee received a compensation of 200 RMB (approximately US \$28) as appreciation for their time. No identifiable images of participants were included in the manuscript or supplementary materials.

Results

Overview

The following results are based on interviews with 14 adults with prehypertension recruited online in mainland China.

Characteristics of Participants

A total of 14 individuals with prehypertension participated in the interviews. Seventeen individuals were invited in total, but 3 declined due to work commitments. We initially interviewed 11 participants, and after identifying the key topics, we conducted 3 additional interviews. Thematic saturation was reached, as no new themes or codes emerged in the final interviews. Data collection was therefore concluded. The

characteristics of the 14 interviewed prehypertensive participants from mainland China are presented in [Table 1](#). The mean age of participants was 33.93 (SD 12.86) years old, with the majority (64.29%) being female and single. The average BMI was 25.10 kg/m² and the mean BP was 129 (SD 8)/81 (SD 6) mm Hg. All participants had at least a college degree and reported using the internet for approximately 8 hours daily. Nearly half of the participants engaged in regular exercise or frequently used their mobile phones.

Table 1. Characteristics of the prehypertensive participants from mainland China (N=14).

Variables	Value
Age (years), mean (SD)	33.93 (12.86)
Sex, n (%)	
Male	9 (64.29)
Female	5 (35.71)
BMI (kg/m ²), mean (SD)	25.10 (4.28)
Marital status, n (%)	
Married	4 (28.57)
Single	9 (64.29)
Divorced	1 (7.14)
Educational background, n (%)	
College	11 (78.57)
Postgraduate	3 (21.43)
Occupation, n (%)	
Employee of an enterprise or institution	6 (42.86)
Student	4 (28.57)
Civil servant or official	1 (7.14)
Other	3 (21.43)
Daily internet use (hours), mean (SD)	8.64 (2.13)
Regular exercise habits, n (%)	
Yes	7 (50.00)
No	7 (50.00)
Mobile phone usage habits, n (%)	
Yes	6 (42.86)
No	8 (57.14)
Systolic or diastolic blood pressure (mm Hg), mean (SD)	129 (8)/81 (6)

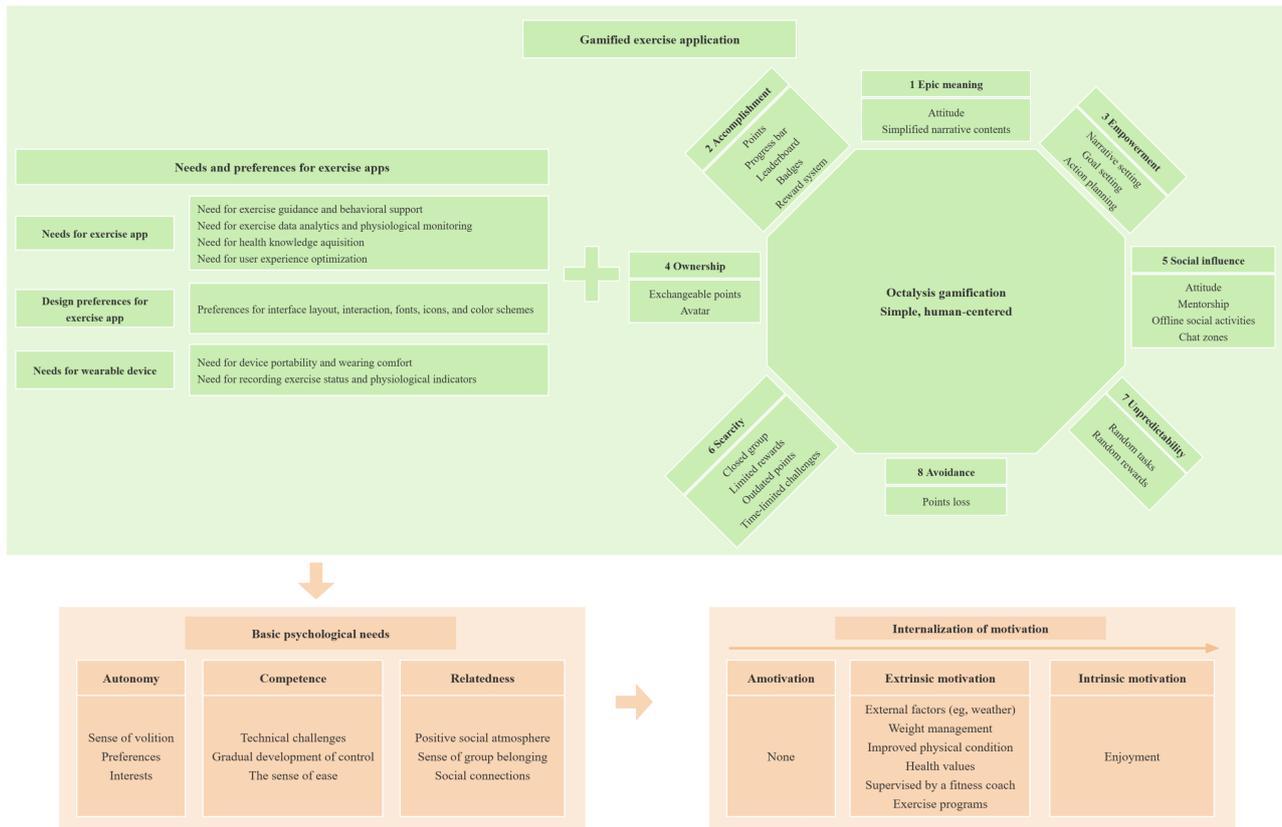
Key Themes

Thematic Analysis Overview

Three overarching domains emerged from our hybrid inductive-deductive analysis, as shown in [Figure 2](#). An inductive or deductive hybrid thematic analysis was used to analyze the data. Therefore, the theme of needs and preferences for exercise

apps was divided into 3 subthemes: needs for exercise apps, design preferences for exercise apps, and needs for wearable devices. Based on the Gamification Octalysis Framework, 8 core psychological themes were predefined, with specific topics categorized under each core theme. The entire conceptual model was grounded in the SDT. The 3 fundamental psychological needs and motivation types were carefully identified and outlined.

Figure 2. Conceptual framework in qualitative study.



Domain 1: Needs and Preferences for Exercise Apps

Overview

Themes, subthemes, and detailed quotations in domain 1 were summarized in Table S1 in Multimedia Appendix 2.

Needs for Exercise App

Participants identified several key needs for exercise apps, including real-time guidance, data feedback, and user-centered design. They expected the app to provide posture correction, integrate with calendars for automated reminders, and offer individualized exercise plans with clear progress schedules. Comprehensive tracking of exercise performance, calorie expenditure, and physiological indicators such as BP was also viewed as essential for maintaining motivation and monitoring health outcomes. In addition, participants emphasized intuitive onboarding, practical and content-oriented interfaces, and personalized interaction features that customized to different user preferences to enhance overall usability.

Design Preferences for Exercise App

Participants suggested that exercise apps should feature a more user-friendly and visually appealing interface, with simple fonts, intuitive icons (eg, a running figure), and color schemes such as green to evoke nature or orange to convey energy and vitality. They also emphasized the value of integrating conversational features powered by large language models to enhance personalization and sustain user engagement.

Needs for Wearable Device

Participants emphasized that wearable devices should be lightweight, comfortable, and focused on essential functions. They valued features such as exercise tracking, physiological monitoring, and alert notifications for indicators such as heart rate and BP, noting that these functions enhance safety and promote consistent exercise engagement.

Domain 2: Gamification Design Insights Based on the Octalysis Framework

Overview

Themes, subthemes, and detailed quotations in domain 2 are shown in Table 2. Participants generally expressed positive attitudes toward incorporating gamified elements into exercise apps but emphasized that such features should remain simple, time-efficient, and seamlessly integrated into daily life to maintain practicality.

According to the Octalysis Gamification Framework, the 8 core motivation drives are divided into white hat core drives (ie, epic meaning, accomplishment, and empowerment), which are designed to make users feel powerful, fulfilled, and satisfied. In contrast, black hat core drives (ie, scarcity, unpredictability, and avoidance) tend to make users feel obsessed, anxious, and addicted. The remaining 2 drives (ie, ownership and social influence) are neutral and do not fall under either category.

Table . Summary of the themes, subthemes, and detailed quotations in domain 2 (gamification design insights based on the Octalysis Framework).

Themes and subthemes	Supporting quotations
Overall attitudes	<ul style="list-style-type: none"> • “I feel pretty good about it. It might be better than just using something basic, like providing video instructions alone.” (P2, male) • “I think it's great because I really love playing games myself.” (P6, female) • “For me, it's because I value my time a lot. I wouldn't want to spend too much time on this.” (P4, female) • “I hope the game's process and content can be simplified, as it's primarily an exercise app.” (P12, male) • “I'd prefer it to be simpler, something that can seamlessly integrate into our daily life.” (P14, male)
White hat core drives	
Epic meaning	
Attitude	<ul style="list-style-type: none"> • “It would be nice if it worked like this, you know, like in a game where you're playing something like CS. For example, if I weigh 200 pounds and can only reach the first level, but after walking 10,000 steps today, I can move on to the second level. The second level could look more luxurious, with lots of rewards that I can claim.” (P3, male) • “I'm not really a fan of these kinds of games, so I probably wouldn't enjoy this one either. It doesn't motivate me much, even if you added a storyline.” (P6, female)
Simplicity	<ul style="list-style-type: none"> • “Exercise is already tiring enough, so why add all these flashy features? The simpler, the better. As you mentioned with the tree-growing idea, just seeing progress little by little can be enough. If it's too complicated, it just wastes time, and people won't want to engage with it.” (P5, male)
Accomplishment	
Points	<ul style="list-style-type: none"> • “From my perspective, I don't care about the extra points; what's important is that I want to exercise. Even now, I feel like exercising because I believe it benefits my body, so I just go ahead and do it. The number of points doesn't matter much.” (P9, female) • “If it were possible to exchange points for physical rewards or other things, it would be more motivating for users to earn points.” (P12, male)
Progress bar	<ul style="list-style-type: none"> • “I think that since it's my own goal, having a progress bar would make it clearer and more straightforward.” (P10, female) • “The progress bar allows me to see my progress, but when I see that there's still a long way to go, I feel a sense of difficulty and pressure. However, with points, I don't feel this pressure.” (P8, male)
Leaderboard	<ul style="list-style-type: none"> • “If there's a leaderboard with points, comparing myself to friends or even strangers can be motivating. When this kind of system is in place, I might feel a little pressure to rank higher, driven by competition or even a bit of vanity.” (P12, male) • “But honestly, I think the leaderboard is too competitive. It just feels overwhelming, and I don't like it.” (P6, female) • “A real-time leaderboard can be particularly motivating, as you might be very close to second place and could potentially surpass them. At that point, you would likely choose to check in more often in order to overtake them.” (P7, male)
Badges	<ul style="list-style-type: none"> • “A meaningful badge, especially one like the ones I got for running before, which are specific and limited, like a special badge for running during the Dragon Boat Festival. It's a bit like collecting.” (P7, male)
Reward system	<ul style="list-style-type: none"> • “Something visually appealing, perhaps with a higher rarity, would be great. The simpler the reward, the better. If the reward system is too complicated, it just becomes confusing, and I might end up not wanting it at all.” (P2, male)

Themes and subthemes	Supporting quotations
Empowerment	
Narrative setting	<ul style="list-style-type: none"> • “The scenario setup should certainly allow for choice, right? Because there are 50 scenarios, and through effort, I can only reach 10 of them. The rest are unattainable, so I would just choose the 10 that I want to focus on.” (P3, male)
Goal setting	<ul style="list-style-type: none"> • “Progressive goals are better, so that I don’t stop after just a day or two.” (P1, female)
Action planning	<ul style="list-style-type: none"> • “It would be ideal if you could create a basic, foundational health plan or goal based on some basic information I provide. I would then be able to adjust it according to my own needs and preferences. Essentially, you would provide a framework, and I could make flexible adjustments within that structure. This approach would work better for me.” (P4, female) • “What’s important in making a plan is how to connect it with daily habits.” (P7, male)
Black hat core drives	
Scarcity	
Closed group	<ul style="list-style-type: none"> • “I don’t accept it. It should be voluntary—you can join if you want. But if there are thresholds, it’s like separating the top students from the bottom ones.” (P7, male)
Limited rewards	<ul style="list-style-type: none"> • “There was a cycling challenge before that I tried. It started with a 10-kilometer challenge, then 20 kilometers, 50 kilometers, and different levels of difficulty. If you completed them, you could go to an offline store to claim your reward. A lot of people managed to stick with it for more than half, so the participation was pretty high. The rewards were limited, though, and many of them got snapped up quickly.” (P7, male)
Outdated points	<ul style="list-style-type: none"> • “If I had the choice to set it myself, I would accept the rules. It would indirectly encourage me to exercise. If I don’t exercise, I would lose the corresponding points, but if I do exercise, I would earn more points than I lost.” (P12, male)
Time-limited challenges	<ul style="list-style-type: none"> • “In my view, for a challenge like this, is the reward given to anyone who meets the requirements, or is it just for a select few? Are the spots limited? I think a lot of people will be drawn in by the rewards. If the rewards aren’t attractive, though, I doubt many people would be willing to participate.” (P13, male) • “But the time span needs to be long enough. For example, if you only give one or two days, and I happen to be on a business trip during that time, then I simply won’t be able to complete it.” (P8, male)
Unpredictability	
Random tasks	<ul style="list-style-type: none"> • “There’s no penalty for failure—in other words, if you don’t complete it, no points will be deducted. That’s acceptable.” (P12, male)
Random rewards	<ul style="list-style-type: none"> • “Rewards can be divided into two types, kind of like a blind box. There are hidden or special editions—if you want one of those, you can pay a higher price to get it. But if you’d rather pay less, then you rely purely on luck to draw one.” (P11, male)
Avoidance	

Themes and subthemes	Supporting quotations
Points loss	<ul style="list-style-type: none"> “I don't want there to be any points that can be deducted in this process. After all, the points I earn are the result of my own effort.” (P6, female) “If you don't complete it, there could be a penalty or a supervision mechanism, which encourages you to finish the task. Of course, the points deducted shouldn't be too many—just a symbolic loss. However, that sense of loss might motivate you to push yourself to complete it.” (P8, male)
Ownership and social influence	
Ownership	
Exchangeable points	<ul style="list-style-type: none"> “It mainly depends on what you can use the points for. If you're collecting points to exchange for a specific item, you'll definitely work harder to earn those points in order to get that item.” (P13, male)
Avatar	<ul style="list-style-type: none"> “I would prefer it to be more customizable, like being able to change outfits. For example, you could earn things like different hairstyles or clothes through exercise milestones.” (P10, female) “You just input your data in, and that's me, even though it's virtual.” (P3, male) “It might just be that rarity is more important. Some things might be ugly, like certain luxury items, but because they're rare, people still really want them.” (P2, male)
Social influence	
Attitude	<ul style="list-style-type: none"> “This doesn't really appeal to me. If it's someone I know, I wouldn't want them to know my step count or anything like that. And if it's a stranger, I have no desire to share or engage in any kind of exchange.” (P1, female) “I think mutual supervision is fine, and it's okay to check in with each other to complete tasks.” (P7, male) If it's something like posting on WeChat, I think that would be better. After all, exercise is also a chance to connect and interact. A lot of people end up becoming friends through things like running, and I think that's great.” (P8, male)
Mentorship	<ul style="list-style-type: none"> “Actually, it was a friend who encouraged me to start running. If I didn't have anyone pushing me, I definitely wouldn't have kept going. At the beginning, there are so many challenges, and it's easy to give up. But my friend kept encouraging me and guiding me.” (P4, female)
Offline social activities	<ul style="list-style-type: none"> “Meeting some strangers can be quite exciting, because life can be too mundane at times. Sometimes, having something new and fresh can really make a difference.” (P4, female) “It's not limited to the types of activities we organize, such as sports. It could also include hosting lectures, either online or offline. For example, as mentioned earlier, a lecture about the relationship between exercise and hypertension could be held, either online or in person—both options are possible.” (P14, male)
Chat zones	<ul style="list-style-type: none"> “In a group of 50 people, everyone can have a conversation. For example, if someone is 50 years old and hasn't seen much change after three months of exercise, and I'm 36, and after two months, I've seen good results, I could share my experience in just a few simple words.” (P3, male)

White Hat Core Drives (Epic Meaning, Accomplishment, and Empowerment)

Participants had mixed views on epic meaning gamification. While some found story-driven progress motivating, others preferred simpler designs, emphasizing that excessive

storytelling could complicate the exercise experience and reduce engagement. Participants expressed diverse views on accomplishment-related gamification elements, such as points, progress bars, leaderboards, badges, and rewards. While some found these features motivating, others felt they could create pressure or distraction. They preferred meaningful, visually

appealing, and simple rewards, such as limited-edition badges or tangible incentives, to enhance a sense of accomplishment. In addition, participants emphasized the need for empowerment through personalized, flexible goal setting and action planning. They preferred progressive goals, customizable plans that balance system guidance with personal choice, and seamless integration of exercise routines into daily life to cultivate exercise habit.

Black Hat Core Drives (Scarcity, Unpredictability, and Avoidance)

Participants showed mixed attitudes toward scarcity elements. They preferred voluntary participation over closed groups, viewed limited rewards as motivating when fairly attainable, and generally supported outdated points as a self-regulated incentive. For time-limited challenges, they emphasized the need for appealing rewards and flexible durations to fit diverse schedules. Participants favored random rewards without penalty, suggesting that unpredictability could enhance engagement if implemented positively. Creative ideas, such as blind box-style rewards, were viewed as appealing ways to maintain user interest and excitement. Moreover, participants expressed mixed views on avoidance mechanisms. While some opposed point deductions, others accepted small, symbolic penalties as a mild motivator to encourage task completion without causing frustration.

Ownership and Social Influence

Participants valued a sense of ownership through features such as exchangeable points and customizable avatars. They preferred avatars that reflect personal identity and include collectible or rare items, enhancing motivation and emotional connection to the app. Participants expressed diverse views on social influence in gamified exercise apps. While some preferred privacy and minimal sharing, others valued mutual support, peer interaction, and community features. Many favored social activities and opportunities for mentorship or experience sharing, viewing them as effective ways to enhance motivation and engagement.

Domain 3: Psychological Needs Satisfaction in SDT

Overview

According to SDT, individuals engage in behaviors because their basic psychological needs (ie, autonomy, competence, and relatedness) are fulfilled by the external environment. We interviewed participants on these relevant aspects. Themes, subthemes, and detailed quotations in domain 3 were summarized in Table S2 in [Multimedia Appendix 3](#).

Participants highlighted the importance of autonomy in exercise, preferring activities that align with their personal interests and preferences. They believed that such self-chosen exercises enhance enjoyment and support long-term engagement. Participants associated competence with mastering technical skills, gaining control over exercises, and experiencing a sense of ease. They also highlighted that perceived benefits, including improved sleep quality, overall health, emotional well-being, and weight management, further strengthened their motivation to continue exercising. Participants also emphasized that social connections, atmosphere, and group activities enhance

motivation for exercise. Engaging with friends or workout partners made PA more enjoyable and supportive.

Extrinsic Motivation

Participants were primarily motivated by extrinsic factors such as weight management, physical health, and external support. Environmental convenience, personal trainers, and structured exercise programs helped maintain engagement, though many relied on external drives to stay active.

Discussion

Principal Findings

In this qualitative study, we interviewed 14 prehypertensive individuals, providing valuable insights into their needs and preferences for exercise apps, gamified design considerations based on the Octalysis Gamification Framework, and basic psychological needs satisfaction and motivation according to SDT.

To the best of our knowledge, this study is among the first to systematically integrate these 2 theoretical frameworks to understand user requirements for digital health interventions in prehypertension management. Our research advances both theoretical understanding and practical application by revealing how the Octalysis Gamification Framework and SDT can synergistically inform exercise behavior change interventions, while providing evidence-based guidance for app design. The design implications discussed below should be interpreted according to the study context, which involved a small qualitative sample of individuals with prehypertension recruited online in mainland China. Notably, these findings are most applicable to prehypertension populations in similar Chinese digital health contexts. They are not directly generalizable to offline-recruited prehypertension groups, individuals from other cultural backgrounds, or nonexercise-focused gamified apps.

Interpretation of Findings

Objective 1: Exercise App Needs and Design Preferences Among Individuals With Prehypertension

This section addresses the first objective of the study, which aimed to explore the needs and design preferences for exercise apps among individuals with prehypertension.

Since the emergence and popularity of the concept of user-centered design, many qualitative studies have been conducted to explore individuals' needs and preferences for app design, especially in the health sector, such as apps targeting PA [27-30]. According to previous findings from interviews conducted with individuals who have chronic obstructive pulmonary disease, physical disabilities, and frailty, similar results were observed, such as the need to motivate users to engage in exercise and the importance of incorporating gamification and social elements into these apps. These functional requirements were also identified in our interview results. However, different populations have varied demands depending on their primary exercise goals.

In our study, we focused on prehypertensive individuals, whose primary motivation for exercising was to lower BP and prevent

the incidence of hypertension. As a result, they place high value on their exercise data and require feedback like weekly or monthly reports of their behavior, which differs from other populations. Moreover, physiological data such as BP, weight, and heart rate were particularly important for them to monitor their health status and the effectiveness of exercise. Therefore, they expressed a strong desire to integrate wearable devices into these apps. From the participants' perspectives, they want the device to be designed to be portable and comfortable, so that it does not hinder their use, which is consistent with previous findings [31,32]. With regard to the functions of these wearable devices, they should be connected to apps to present exercise and physiological data, thereby informing users about their physical conditions. However, there are many challenges when incorporating wearable devices into apps, such as data privacy, standardization, accuracy, and seamless integration with existing systems, all of which should be carefully considered by software developers [33,34].

Objective 2: User-Centered Gamification Design Strategies Based on the Octalysis Framework

In line with the second objective, this section discusses user-centered gamification design strategies based on the Octalysis Framework.

According to previously published studies, most qualitative studies have mentioned the need for integrating gamification nudges or elements into apps, but they lack specific suggestions on how to design gamification techniques, leaving software developers confused about what to do next [32,35,36].

In our study, participants have not only provided their attitudes and suggestions about existing gamification functions and techniques, but also offered innovative ideas for the design of gamification rules, reflecting a participant-incorporated design perspective. For example, participants expressed that they want the gamification techniques used in the apps to be simple, easy to understand, and easy to start, which aligns with a previous systematic review stating that gamification systems should be easy to use and simplify content [37].

Almost all participants expressed very positive stances toward gamification used in these apps, consistent with previous studies [38,39]. In addition, varied attitudes were observed toward different gamification motivational drives, based on participants' previous experiences, personal preferences, and personalities. With regard to specific motivational drives, participants expressed conservative attitudes toward narratives, as they think it is difficult to design an engaging storyline. However, a story without attraction would be ineffective, which aligns with previous studies that emphasize the effective use of narrative element [40].

Accomplishment is the most commonly used motivational drive in gamified apps, with traditional points, badges, and leaderboards elements underpinning this drive. During the interviews, most participants had encountered these elements in their previous experiences. They provided creative suggestions for enhancing the motivational functions of these techniques, such as exchangeable points, rare and meaningful

badges, and dynamic leaderboards. These insights are valuable for gamification designers and software developers.

Avatars are a common gamification element in the ownership drive, and participants mentioned that they can be dynamic, personalized, and rare to attract users, a technique emphasized in a previous study [41]. Social influence is a core motivational drive in the Octalysis Gamification Framework, and social factors should always be considered when triggering a behavior. Based on our interview observations, the acceptance of social elements is highly related to a person's personality, which is confirmed by previous studies [42]. Therefore, generating personality-matched social designs is worth exploring in the future.

In addition to other commonly used social features in apps, such as group tasks, friend circles, and chat zones, offline social activities and mentorship received very positive acceptance, which should be considered in future designs. Regarding black hat drives, participants did not show any notable preferences for these motivational drives, mainly because they are driven by the fear of losing. If these techniques are used appropriately, they may produce unexpected results.

Objective 3: Supporting Motivational Transformation Through Exercise Apps: The SDT Perspective

Addressing the third objective, this section elucidates how exercise apps may facilitate motivational transformation in individuals with prehypertension. Participants' engagement in exercise was closely linked to the satisfaction of autonomy, competence, and relatedness.

With regard to autonomy, participants mainly valued a sense of volition, preferences, and interests, which shows similar results to a previous study that used SDT to explain trial retention behavior [43]. The key difference between exercise and retention behavior is the complexity of exercise and its various types. Therefore, the future designs should emphasize exercise choices based on users' interests.

Participants also illustrated the importance of competence satisfaction. Many expressed that their progress or a sense of achievement in handling more difficult tasks would be a significant reason for them to continue exercising, which aligns with the results of a previous study [44]. Accordingly, progressively difficult tasks could be considered in future challenge designs. Relatedness, or the connection with others, should not be neglected when explaining behavior. Many participants shared that the involvement of others would be a notable drive to promote their exercise, a result also confirmed by quantitative data [45]. Incorporating social factors appropriately to satisfy the needs of different groups warrants consideration.

According to our interview results, most participants exercise due to extrinsic motivations, such as weight management and health values, while intrinsic enjoyment was less commonly observed. However, intrinsic motivation is predicted to support long-term, habitual physical activity [46].

Therefore, future research should focus on strategies that can transform extrinsic motivation to intrinsic motivation. With the

rapid popularity of digital technologies, integrating these technologies with behavior change strategies could play a crucial role in influencing behavior and motivating participants. Digital tools, such as apps, large language models, and wearable devices, combined with strategies such as gamification, nudge, and behavioral economics, could contribute to addressing physical inactivity by fostering participants' intrinsic motivation.

Strengths and Limitations

The qualitative study has several strengths. In the digital age and the era of technological explosion, integrating user-centered design principles into technology development is crucial. The study used a hybrid inductive or deductive thematic analysis to explore prehypertensive individuals' needs and preferences for gamified exercise apps through semistructured interviews. Based on SDT and the Octalysis Gamification Framework, the study enriches the theoretical basis for behavior change experts and provides practical design insights for software developers. It is timely and aligns well with global trends.

The study also has limitations. We recruited participants through an online poster, and individuals who responded were more likely to pay attention to their physical health compared with the general populations. Therefore, the insights derived from the included participants may not be generalizable to the broader population. Intervention designs based on these insights may not always be effective. Future quantitative studies, including participants from diverse cultural backgrounds and regions, should be conducted to validate our findings.

Moreover, since motivation and psychological need satisfaction are likely to change over time, longitudinal studies could provide a better understanding of exercise behavior according to SDT.

Additionally, ethical considerations warrant reflection when designing gamified exercise apps, as some participants expressed concerns about pressure, competition, and the possibility that excessive entertainment-oriented gamification may weaken the app's exercise-focused purpose. Future researchers remain attentive to the potential psychological burden of such apps and ensure that participation is voluntary and supportive. A further limitation is that interrater reliability (eg, Cohen κ) was not formally calculated. Although discrepancies in coding were resolved through consensus among researchers, future studies can incorporate statistical measures of interrater reliability to strengthen methodological rigor.

Conclusions

This theory-informed qualitative study is among the first to explore the needs and preferences of individuals with prehypertension for gamified exercise apps by integrating the Octalysis Gamification Framework with SDT. The findings suggest that successful gamified exercise apps for prehypertensive individuals may benefit from going beyond traditional game mechanics to address deeper psychological needs. Participants emphasized the importance of personalized exercise programming, robust health monitoring capabilities, and human-centered gamification designs that support autonomy, competence, and relatedness. Satisfaction of these basic psychological needs is critical for transforming extrinsic motivations, such as weight management and health values, into intrinsic enjoyment of exercise. By integrating a theoretical basis with user-centered perspectives, this study provides context-specific design implications for future development of gamified exercise apps tailored to prehypertensive individuals in similar digital health contexts.

Acknowledgments

The authors declare the use of generative artificial intelligence (AI) in the research and writing process. According to the GAIDeT taxonomy (2025), the generative AI tools were used under full human supervision for formulating research questions and hypotheses, proofreading and editing, and translation. The generative AI tool used was ChatGPT-5.2. Responsibility for the final manuscript lies entirely with the authors. Generative AI tools are not listed as authors and do not bear responsibility for the final outcomes. This declaration was submitted by SL.

Funding

This study was funded by the National Key R&D Program of China (No.2025YFE0204500) and Non-Profit Central Research Institute Fund of the Chinese Academy of Medical Sciences (Grant No. 2023-RC320-01).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

HMM and MHP contributed equally as co-corresponding authors. MHP can be contacted by email at parkmihua@snu.ac.kr.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guidelines.

[[DOCX File, 17 KB - games_v14i1e86793_app1.docx](#)]

Multimedia Appendix 2

Themes of domain 1.

[\[DOCX File, 16 KB - games_v14i1e86793_app2.docx \]](#)

Multimedia Appendix 3

Themes of domain 3.

[\[DOCX File, 15 KB - games_v14i1e86793_app3.docx \]](#)

Checklist 1

COREQ checklist.

[\[PDF File, 639 KB - games_v14i1e86793_app4.pdf \]](#)**References**

1. Chen R, Dharmarajan K, Kulkarni VT, et al. Most important outcomes research papers on hypertension. *Circ Cardiovasc Qual Outcomes* 2013 Jul;6(4):e26-e35. [doi: [10.1161/CIRCOUTCOMES.113.000424](https://doi.org/10.1161/CIRCOUTCOMES.113.000424)] [Medline: [23838106](https://pubmed.ncbi.nlm.nih.gov/23838106/)]
2. Hypertension. World Health Organization. 2023. URL: <https://www.who.int/news-room/fact-sheets/detail/hypertension> [accessed 2025-08-23]
3. Julius S, Nesbitt SD, Egan BM, et al. Feasibility of treating prehypertension with an angiotensin-receptor blocker. *N Engl J Med* 2006 Apr 20;354(16):1685-1697. [doi: [10.1056/NEJMoa060838](https://doi.org/10.1056/NEJMoa060838)] [Medline: [16537662](https://pubmed.ncbi.nlm.nih.gov/16537662/)]
4. Writing Committee Members*, Jones DW, Ferdinand KC, et al. 2025 AHA/ACC/AANP/AAPA/ABC/ACCP/ACPM/AGS/AMA/ASPC/NMA/PCNA/SGIM guideline for the prevention, detection, evaluation and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. *Hypertension* 2025 Oct;82(10):e212-e316. [doi: [10.1161/HYP.0000000000000249](https://doi.org/10.1161/HYP.0000000000000249)] [Medline: [40811516](https://pubmed.ncbi.nlm.nih.gov/40811516/)]
5. Li X, Chang P, Wu M, et al. Effect of Tai Chi vs aerobic exercise on blood pressure in patients with prehypertension: a randomized clinical trial. *JAMA Netw Open* 2024 Feb 5;7(2):e2354937. [doi: [10.1001/jamanetworkopen.2023.54937](https://doi.org/10.1001/jamanetworkopen.2023.54937)] [Medline: [38335001](https://pubmed.ncbi.nlm.nih.gov/38335001/)]
6. Strain T, Flaxman S, Guthold R, et al. National, regional, and global trends in insufficient physical activity among adults from 2000 to 2022: a pooled analysis of 507 population-based surveys with 5.7 million participants. *Lancet Glob Health* 2024 Aug;12(8):e1232-e1243. [doi: [10.1016/S2214-109X\(24\)00150-5](https://doi.org/10.1016/S2214-109X(24)00150-5)] [Medline: [38942042](https://pubmed.ncbi.nlm.nih.gov/38942042/)]
7. Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D. Gamification. using game-design elements in non-gaming contexts. 2011 May 7 Presented at: CHI '11; May 7-12, 2011; Vancouver BC Canada p. 2425-2428 URL: <https://dl.acm.org/doi/proceedings/10.1145/1979742> [doi: [10.1145/1979742.1979575](https://doi.org/10.1145/1979742.1979575)]
8. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the effectiveness of gamification on physical activity: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2022 Jan 4;24(1):e26779. [doi: [10.2196/26779](https://doi.org/10.2196/26779)] [Medline: [34982715](https://pubmed.ncbi.nlm.nih.gov/34982715/)]
9. Li SJ, Ma HM, Zhu LQ, et al. Exergame-based behavior change interventions for promoting physical activity: systematic review and meta-analysis of randomized controlled studies. *J Med Internet Res* 2025 Aug 8;27:e62906. [doi: [10.2196/62906](https://doi.org/10.2196/62906)] [Medline: [40779747](https://pubmed.ncbi.nlm.nih.gov/40779747/)]
10. van Velsen L, Ludden G, Grünloh C. The limitations of user-and human-centered design in an eHealth context and how to move beyond them. *J Med Internet Res* 2022 Oct 5;24(10):e37341. [doi: [10.2196/37341](https://doi.org/10.2196/37341)] [Medline: [36197718](https://pubmed.ncbi.nlm.nih.gov/36197718/)]
11. Xu L, Shi H, Shen M, et al. The effects of mHealth-based gamification interventions on participation in physical activity: systematic review. *JMIR Mhealth Uhealth* 2022 Feb 3;10(2):e27794. [doi: [10.2196/27794](https://doi.org/10.2196/27794)] [Medline: [35113034](https://pubmed.ncbi.nlm.nih.gov/35113034/)]
12. Wang M, Xu J, Zhou X, Li X, Zheng Y. Effectiveness of gamification interventions to improve physical activity and sedentary behavior in children and adolescents: systematic review and meta-analysis. *JMIR Serious Games* 2025 Sep 18;13:e68151. [doi: [10.2196/68151](https://doi.org/10.2196/68151)] [Medline: [40966596](https://pubmed.ncbi.nlm.nih.gov/40966596/)]
13. Gao Y, Zhang J, He Z, Zhou Z. Feasibility and usability of an artificial intelligence-powered gamification intervention for enhancing physical activity among college students: quasi-experimental study. *JMIR Serious Games* 2025 Mar 24;13:e65498. [doi: [10.2196/65498](https://doi.org/10.2196/65498)] [Medline: [40127464](https://pubmed.ncbi.nlm.nih.gov/40127464/)]
14. Chou YK. Actionable Gamification: Beyond Points, Badges, and Leaderboards: Packt Publishing Ltd; 2019.
15. Kian TW, Sunar MS, Su GE. The analysis of intrinsic game elements for undergraduates gamified platform based on learner type. *IEEE Access* 2022;10:120659-120679. [doi: [10.1109/ACCESS.2022.3218625](https://doi.org/10.1109/ACCESS.2022.3218625)]
16. Ryan RM, Deci EL. Intrinsic and extrinsic motivation from a self-determination theory perspective: definitions, theory, practices, and future directions. *Contemp Educ Psychol* 2020 Apr;61:101860. [doi: [10.1016/j.cedpsych.2020.101860](https://doi.org/10.1016/j.cedpsych.2020.101860)]
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](https://doi.org/10.1093/intqhc/mzm042)] [Medline: [17872937](https://pubmed.ncbi.nlm.nih.gov/17872937/)]

18. Levitt HM, Bamberg M, Creswell JW, Frost DM, Josselson R, Suárez-Orozco C. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: the APA Publications and Communications Board task force report. *Am Psychol* 2018 Jan;73(1):26-46. [doi: [10.1037/amp0000151](https://doi.org/10.1037/amp0000151)] [Medline: [29345485](https://pubmed.ncbi.nlm.nih.gov/29345485/)]
19. Campbell S, Greenwood M, Prior S, et al. Purposive sampling: complex or simple? Research case examples. *J Res Nurs* 2020 Dec;25(8):652-661. [doi: [10.1177/1744987120927206](https://doi.org/10.1177/1744987120927206)] [Medline: [34394687](https://pubmed.ncbi.nlm.nih.gov/34394687/)]
20. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? *Qual Health Res* 2017 Mar;27(4):591-608. [doi: [10.1177/1049732316665344](https://doi.org/10.1177/1049732316665344)] [Medline: [27670770](https://pubmed.ncbi.nlm.nih.gov/27670770/)]
21. Wang JG. Chinese guidelines for the prevention and treatment of hypertension (2024 revision). *J Geriatr Cardiol* 2025 Jan 28;22(1):1-149. [doi: [10.26599/1671-5411.2025.01.008](https://doi.org/10.26599/1671-5411.2025.01.008)] [Medline: [40151633](https://pubmed.ncbi.nlm.nih.gov/40151633/)]
22. 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](https://doi.org/10.1037//0003-066x.55.1.68)] [Medline: [11392867](https://pubmed.ncbi.nlm.nih.gov/11392867/)]
23. Mohanty S, Christopher B P. A bibliometric analysis of the use of the Gamification Octalysis Framework in training: evidence from Web of Science. *Humanit Soc Sci Commun* 2023;10(1):836. [doi: [10.1057/s41599-023-02243-3](https://doi.org/10.1057/s41599-023-02243-3)]
24. Majumdar A. Thematic Analysis in Qualitative Research Qualitative Techniques for Workplace Data Analysis: IGI Global Scientific Publishing; 2019:197-220. [doi: [10.4018/978-1-5225-5366-3.ch009](https://doi.org/10.4018/978-1-5225-5366-3.ch009)]
25. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. *Int J Qual Methods* 2006 Mar;5(1):80-92. [doi: [10.1177/160940690600500107](https://doi.org/10.1177/160940690600500107)]
26. Lincoln YS, Guba EG. *Naturalistic Inquiry*: Sage; 1985. [doi: [10.1016/0147-1767\(85\)90062-8](https://doi.org/10.1016/0147-1767(85)90062-8)]
27. Olsen SH, Saperstein SL, Gold RS. Content and feature preferences for a physical activity app for adults with physical disabilities: focus group study. *JMIR Mhealth Uhealth* 2019 Oct 11;7(10):e15019. [doi: [10.2196/15019](https://doi.org/10.2196/15019)] [Medline: [31605518](https://pubmed.ncbi.nlm.nih.gov/31605518/)]
28. Peng R, Cao Z, Hu S, et al. Frail older adults' needs and preferences for mobile health exercise interventions guided by nudge theory: a qualitative analysis. *J Clin Nurs* 2025 May;34(5):1830-1839. [doi: [10.1111/jocn.17419](https://doi.org/10.1111/jocn.17419)] [Medline: [39215431](https://pubmed.ncbi.nlm.nih.gov/39215431/)]
29. Sönnerrfors P, Skavberg Roaldsen K, Lundell S, Toots A, Wadell K, Halvarsson A. Preferences for an eHealth tool to support physical activity and exercise training in COPD: a qualitative study from the viewpoint of prospective users. *BMC Pulm Med* 2023 Feb 13;23(1):65. [doi: [10.1186/s12890-023-02353-3](https://doi.org/10.1186/s12890-023-02353-3)] [Medline: [36782155](https://pubmed.ncbi.nlm.nih.gov/36782155/)]
30. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Bainbridge W, editor. *Encyclopedia of Human-Computer Interaction* 2004:445-456 URL: <https://www.scirp.org/reference/referencespapers?referenceid=2501037> [accessed 2026-02-08]
31. Muñoz Esquivel K, Gillespie J, Kelly D, et al. Factors influencing continued wearable device use in older adult populations: quantitative study. *JMIR Aging* 2023 Jan 19;6:e36807. [doi: [10.2196/36807](https://doi.org/10.2196/36807)] [Medline: [36656636](https://pubmed.ncbi.nlm.nih.gov/36656636/)]
32. Moore K, O'Shea E, Kenny L, et al. Older adults' experiences with using wearable devices: qualitative systematic review and meta-synthesis. *JMIR Mhealth Uhealth* 2021 Jun 3;9(6):e23832. [doi: [10.2196/23832](https://doi.org/10.2196/23832)] [Medline: [34081020](https://pubmed.ncbi.nlm.nih.gov/34081020/)]
33. Liu Y, Li S, Yu J, et al. Demand analysis of transitional care for patients undergoing minimally invasive cardiac interventions with AI-driven solutions: a mixed-methods approach. *BMC Nurs* 2025 Apr 23;24(1):453. [doi: [10.1186/s12912-025-03037-5](https://doi.org/10.1186/s12912-025-03037-5)] [Medline: [40269856](https://pubmed.ncbi.nlm.nih.gov/40269856/)]
34. Canali S, Schiaffonati V, Aliverti A. Challenges and recommendations for wearable devices in digital health: data quality, interoperability, health equity, fairness. *PLOS Digit Health* 2022 Oct;1(10):e0000104. [doi: [10.1371/journal.pdig.0000104](https://doi.org/10.1371/journal.pdig.0000104)] [Medline: [36812619](https://pubmed.ncbi.nlm.nih.gov/36812619/)]
35. Ryu E, Farraday D, Fanaroff AC, et al. What motivates participants: a qualitative analysis of gamification and financial incentives to increase physical activity. *BMC Public Health* 2025 May 16;25(1):1804. [doi: [10.1186/s12889-025-22717-0](https://doi.org/10.1186/s12889-025-22717-0)] [Medline: [40380122](https://pubmed.ncbi.nlm.nih.gov/40380122/)]
36. Kappen DL, Mirza-Babaei P, Nacke LE. Technology facilitates physical activity through gamification: a thematic analysis of an 8-week study. *Front Comput Sci* 2020 Oct 22;2:2020. [doi: [10.3389/fcomp.2020.530309](https://doi.org/10.3389/fcomp.2020.530309)]
37. Krath J, Schürmann L, von Korfflesch HFO. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Comput Human Behav* 2021 Dec;125:106963. [doi: [10.1016/j.chb.2021.106963](https://doi.org/10.1016/j.chb.2021.106963)]
38. Bai S. Investigating the effects of gamification on student learning engagement and performance. : The University of Hong Kong (Pokfulam, Hong Kong); 2022 URL: <https://hub.hku.hk/handle/10722/322936> [accessed 2026-02-08]
39. Vrcelj A, Hoic-bozic N, Holenko Dlab M. Attitudes of secondary school teachers towards gamification. 2023 Oct Presented at: EPESS; Jul 6-9, 2023; Budapest, Hungary p. 54-66. [doi: [10.55549/epess.1381514](https://doi.org/10.55549/epess.1381514)]
40. Li M, Wang Y, Wu Y, Liu H. Gamification narrative design as a predictor for mobile fitness app user persistent usage intentions: a goal priming perspective. *Enterp Inf Syst* 2021 Nov 26;15(10):1501-1545. [doi: [10.1080/17517575.2021.1941272](https://doi.org/10.1080/17517575.2021.1941272)]
41. Zhu Y, Long Y, Wei L, et al. Developing cue-behavior association for habit formation: a qualitative study to explore the role of avatar in hypertension. *Digit Health* 2024;10:20552076241265217. [doi: [10.1177/20552076241265217](https://doi.org/10.1177/20552076241265217)] [Medline: [39099680](https://pubmed.ncbi.nlm.nih.gov/39099680/)]
42. Oyibo K, Vassileva J. The relationship between personality traits and susceptibility to social influence. *Comput Human Behav* 2019 Sep;98:174-188. [doi: [10.1016/j.chb.2019.01.032](https://doi.org/10.1016/j.chb.2019.01.032)]

43. Draper CE, Tshetu N, Nkosi N, Lye S, Norris SA. Retention in the Bukhali trial in Soweto, South Africa: a qualitative analysis using self-determination theory. *BMJ Glob Health* 2025 Feb 10;10(2):39929537. [doi: [10.1136/bmjgh-2024-017729](https://doi.org/10.1136/bmjgh-2024-017729)] [Medline: [39929537](https://pubmed.ncbi.nlm.nih.gov/39929537/)]
44. 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):24353482. [doi: [10.1177/1559827610392704](https://doi.org/10.1177/1559827610392704)] [Medline: [24353482](https://pubmed.ncbi.nlm.nih.gov/24353482/)]
45. Shelton RC, McNeill LH, Puleo E, Wolin KY, Emmons KM, Bennett GG. The association between social factors and physical activity among low-income adults living in public housing. *Am J Public Health* 2011 Nov;101(11):2102-2110. [doi: [10.2105/AJPH.2010.196030](https://doi.org/10.2105/AJPH.2010.196030)] [Medline: [21330588](https://pubmed.ncbi.nlm.nih.gov/21330588/)]
46. 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 Jun 22;9(1):78. [doi: [10.1186/1479-5868-9-78](https://doi.org/10.1186/1479-5868-9-78)] [Medline: [22726453](https://pubmed.ncbi.nlm.nih.gov/22726453/)]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research

JARS-Qual: Journal Article Reporting Standards for Qualitative Research

PA: physical activity

SDT: Self-Determination Theory

WHO: World Health Organization

Edited by S Brini; submitted 30.Oct.2025; peer-reviewed by KM Cheng, L Zhu; revised version received 19.Jan.2026; accepted 20.Jan.2026; published 25.Feb.2026.

Please cite as:

Li S, Liu G, Qiao L, Yu HY, Zhang YY, Pei R, Lin YT, Ma HM, Piao MH

Design Requirements for Gamified Exercise Apps for Adults With Prehypertension Based on the Octalysis Framework and Self-Determination Theory: Qualitative Interview Study

JMIR Serious Games 2026;14:e86793

URL: <https://games.jmir.org/2026/1/e86793>

doi: [10.2196/86793](https://doi.org/10.2196/86793)

© Sijia Li, Genying Liu, Li Qiao, Hong-Yu Yu, Yu-Yang Zhang, Runyuan Pei, Yue-Tao Lin, Hao-Ming Ma, Mei-Hua Piao. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 25.Feb.2026. 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.

The Mechanism and Design Principles of Serious Games in Enhancing Adolescents' Internet Adaptability

Shijie Gao¹, BS; Min Jia^{1*}, MA; Weijun Wang², PhD; Jianping Hu¹, MA; Shihao Ma^{1*}, PhD

¹School of Psychology, Central China Normal University, No.152 Luoyu Road, Hongshan District, Wuhan, Hubei, China

²Key Laboratory of Adolescent Cyberpsychology and Behavior, Ministry of Education, Wuhan, Hubei, China

*these authors contributed equally

Corresponding Author:

Jianping Hu, MA

School of Psychology, Central China Normal University, No.152 Luoyu Road, Hongshan District, Wuhan, Hubei, China

Abstract

Adolescents' internet adaptability (IA) is crucial for their online behavior and mental health. Serious games (SGs), as an emerging educational tool, hold promise for enhancing this adaptability through engaging, goal-oriented learning. Yet, direct evidence in this area remains limited. This viewpoint aims to clarify the mechanisms through which SGs enhance adolescents' IA and to derive corresponding design principles that can inform educational practice and game development. Drawing on insights from both Chinese and international studies, this study adopts a cross-contextual perspective to explore how SGs can foster IA in varied educational environments. Beyond synthesizing existing findings, this viewpoint provides an integrated account of why IA is essential in contemporary digital life and how SGs can support its development. It proposes a 3-stage framework, illustrating how contextualized design, real-time feedback, and dynamic tasks promote experiential learning, self-regulation, and the transfer of online skills. Based on this framework, the study further articulates 6 core design principles: clear goal definition, interaction diversity, contextual authenticity, immediate, scaffolding and explanatory feedback, a dynamically adaptive learning environment, and safety-by-design for digital well-being. These principles translate the core characteristics and mechanisms of SGs into actionable guidance for developing effective IA interventions. By synthesizing theoretical insights with practical considerations, this viewpoint highlights how SGs can serve as accessible and scalable tools to support adolescents in navigating increasingly complex digital environments. Together, these insights provide practical implications for educators, curriculum designers, and digital game developers seeking to foster adolescents' safe, responsible, and adaptive engagement in online environments.

(*JMIR Serious Games* 2026;14:e82505) doi:[10.2196/82505](https://doi.org/10.2196/82505)

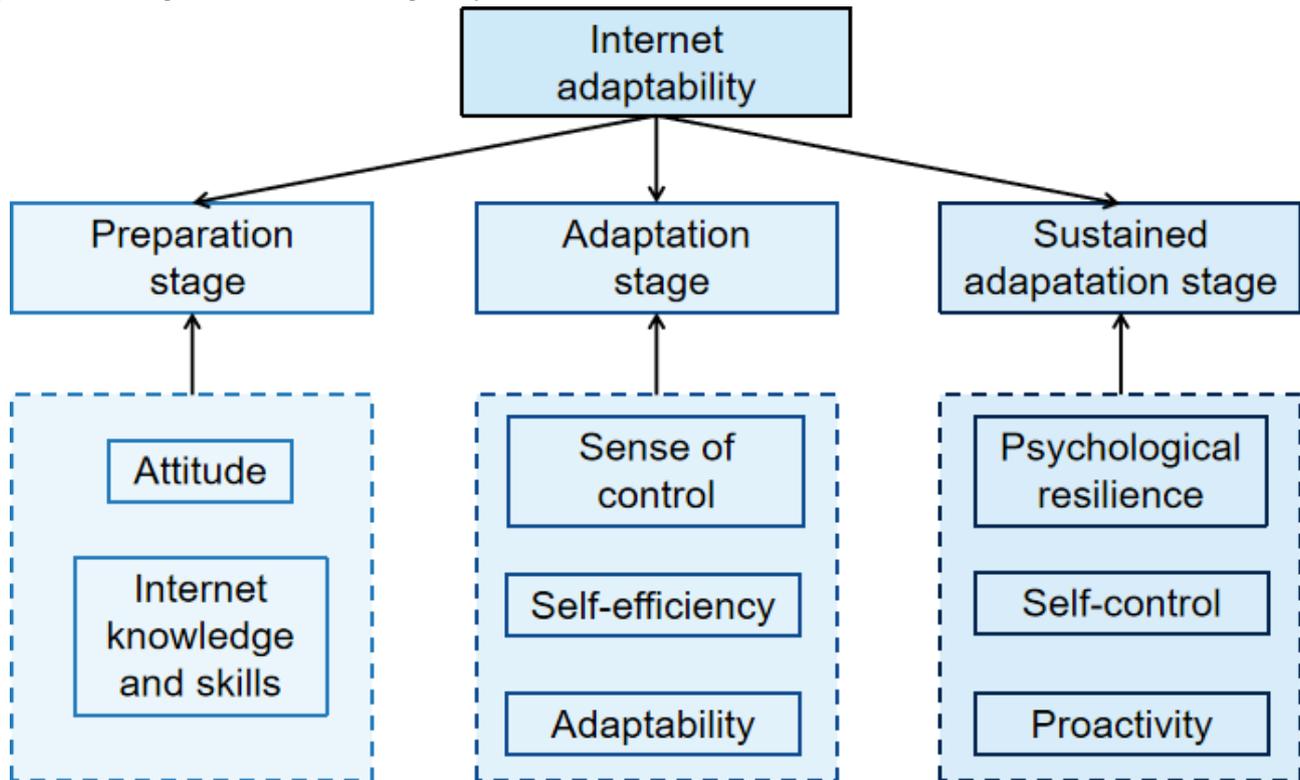
KEYWORDS

game design mechanism; adolescents; internet adaptability; game based learning; game design principles

Introduction

Internet adaptability (IA) refers to the individual capacities that develop through interactions with the online environment [1]. It is a core component of digital engagement and develops across three stages: (1) the preparation stage, in which adolescents

acquire relevant knowledge and psychological readiness; (2) the adaptation stage, involving their evaluation of the digital environment and behavioral adjustment; and (3) the sustainability stage, reflecting their ability to cope with technological change and digital stressors [1,2]. The following conceptualization illustrates these 3 stages and the psychological factors that characterize each [1] (Figure 1).

Figure 1. The conceptualization of internet adaptability.

To clarify its conceptual boundaries, IA must be distinguished from related constructs. Digital literacy or competence emphasizes relatively stable technical skills such as information processing and tool use [3-5]. Digital resilience focuses on recovering from adverse online experiences and is primarily reactive [6]. Digital well-being concerns maintaining a healthy and balanced relationship with technology [7]. In contrast, IA

represents a dynamic psychological adjustment process, grounded in self-efficacy, sense of control, resilience, and proactive regulation, that supports sustained functional engagement across evolving digital contexts [1,8]. A comparative overview of these constructs is presented in Table 1.

Table . Internet adaptability and related concepts.

Concepts	Definitions	Static or dynamic	Main dimensions	Relationship with IA ^a
Internet adaptability [1]	Individual capacities that develop through interactions with the online environment.	Dynamic and process-oriented	Attitude, online knowledge and skills, sense of control, self-efficacy, adaptability, self-control, psychological resilience, and proactivity.	IA integrates elements from digital competence, resilience, and well-being but distinguishes itself by focusing on long-term adaptation, proactive regulation, and psychological behavioral flexibility in a changing digital environment.
Digital literacy or competence [3,4]	Competence in using digital technologies to access, evaluate, create, and communicate information; ability to use ICT ^b tools appropriately.	More on stable skill set or capacity (relatively static)	Information and data literacy, communication or collaboration, content creation, safety or cybersecurity, problem solving, critical thinking, and ethical or reflective use.	IA differs by emphasizing ongoing adaptation and psychological-behavioral regulation, not just ability to use tools.
Digital resilience [6]	The capacity to cope with, recover from, and learn after adverse or risky online events or digital pressures; ability to bounce back, maintain well-being, and continue functioning after negative experiences.	More reactive or recovery-oriented (although may include learning after adversity)	Coping strategies, recovery, learning from negative online experiences, psychological well-being, and safe behavior online.	IA encompasses but extends beyond DR ^c : while DR addresses response to adversity, IA emphasizes proactive, continuous adaptation and regulation across changing digital contexts.
Digital well-being [7]	Individual's ability to establish a healthy, balanced, and purposeful relationship with digital technologies, thereby supporting their overall quality of life and personal development.	Ongoing state-oriented or stability-oriented	Digital habits, healthy use, ethical awareness, balanced usage, mental health, life-digital balance, and critical media use.	IA emphasizes adaptive capabilities (cognitive, behavioral, and psychological) to manage digital environments; well-being is an important outcome, but IA focuses more on adaptive functioning, not only well-being.

^aIA: internet adaptability.

^bICT: information and communication technology.

^cDR: digital resilience.

A key question is whether IA applies across different cultural internet environments. From a social adaptation perspective, development reflects the dynamic balance between personal characteristics and environmental demands [9]. Similarly, IA describes the psychological process through which individuals regulate and adjust behavior in response to changing online conditions. Although digital ecosystems differ across cultures, the core adaptive task remains constant: maintaining functional and psychological balance in a dynamic digital environment [1,10]. Thus, IA captures a universal adaptive mechanism, even as specific challenges and adaptive behaviors vary across societies.

IA plays an important role in adolescents' psychological well-being and digital engagement [1,11]. However, current approaches to fostering IA, including classroom instruction, family education, campus initiatives, and internet use regulations, often lack contextual relevance, interactivity, and flexibility, limiting their effectiveness in addressing adolescents' evolving online challenges [12].

Serious games (SGs) have emerged as a promising alternative. Originating from board games, they integrate education with

entertainment to provide personalized and interactive learning experiences [13]. With advances in digital technology and learning theory, SGs have evolved into formats such as video games, simulations, and virtual reality environments. This study focuses on educational video games, defined as games designed for purposes beyond entertainment [14,15]. In this context, SGs refer to educational video games that integrate gaming elements to transmit knowledge, train skills, and promote emotional or behavioral development in a context-rich and engaging way [16,17]. Well-designed SGs typically exhibit several core characteristics: clear educational objectives [18], disciplinary rigor [19], contextual authenticity [20], and interactivity with immediate formative feedback [21]. When integrated effectively, these features support engagement while promoting meaningful learning and behavioral development.

These immersive and interactive affordances make SGs particularly suitable for fostering adolescents' IA. By embedding learning goals within gameplay and simulating realistic online scenarios, SGs enable adolescents to develop cognitive, emotional, and behavioral regulation in controlled virtual environments, supporting safe, responsible, and flexible digital engagement. Empirical studies support these benefits. For

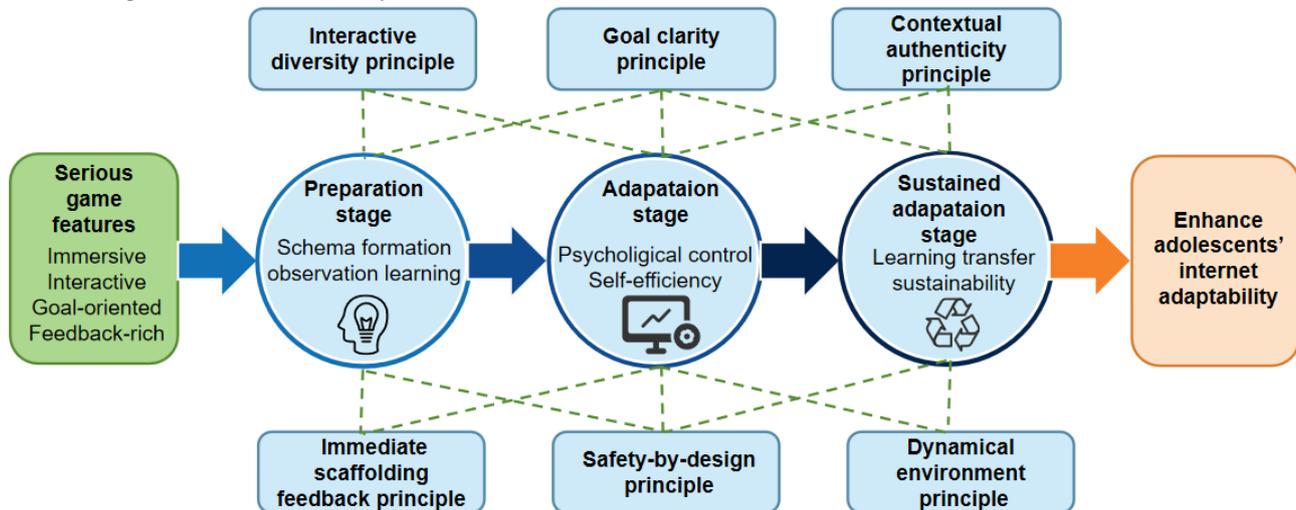
example, SGs targeting cyberbullying prevention enhance empathy and emotional regulation [22], and the cybersecurity game *Riskio* strengthens awareness and perceived control under uncertain conditions [23-25]. Extending beyond isolated dimensions, *InterWeb Action*, a SG grounded in situated learning theory, significantly improved overall IA and multiple behavioral indicators [26].

Despite these emerging findings, research that directly and systematically examines how SGs cultivate IA remains limited. Therefore, rather than conducting a conventional literature

review, this viewpoint synthesizes existing evidence to articulate the mechanisms through which SGs may enhance IA and to propose practical design principles for effective intervention development.

Building on this aim, the viewpoint advances a theoretical framework that examines the psychological processes involved in IA across 3 stages and derives 6 core design principles that translate theory into actionable guidance for educators, curriculum designers, and game developers seeking to promote safe, responsible, and adaptive online behavior (Figure 2).

Figure 2. Conceptual framework of the study.



The Unique Potential of SGs for Adolescents' IA

The Universality of SGs in Educational Practice

Although research directly examining SGs for enhancing adolescents' IA remains limited, their educational effectiveness has been widely demonstrated across diverse learning contexts [27,28].

Meta-analytic evidence indicates that SGs enhance cognitive performance while simultaneously promoting positive emotions and well-being during learning. Empirical studies further show that SGs improve academic achievement and learning motivation across disciplines such as science, language learning, and business education [29,30].

Beyond cognitive outcomes, SGs also support the development of soft skills. For example, the SG *FLIGBY* has been shown to foster leadership, conflict resolution, and critical thinking, highlighting the role of SGs in holistic skill development [31]. Compared with traditional instructional approaches, SGs offer greater flexibility and learner autonomy, aligning well with contemporary educational needs and contributing to more positive learning attitudes [28,32]. Collectively, these findings establish SGs as an effective and versatile educational tool, providing a strong foundation for their application in promoting adolescents' IA.

The Relevance of SG Features to Adolescents' IA

As adolescents face increasingly complex digital environments, SGs offer particular advantages due to their educational orientation, immersive contexts, and interactive design. Although direct empirical evidence remains limited, existing studies suggest several pathways through which SGs may support IA.

Purpose Fit: Precise Alignment of Educational Goals and Needs

SGs enable a close alignment between educational objectives and learners' needs. By embedding learning goals within game tasks and challenges, SGs effectively convey internet-related knowledge while fostering cognitive skills such as critical thinking and problem solving [27,28]. This goal-oriented design makes SGs well suited for strengthening adolescents' online knowledge and competencies.

Environmental Adaptation: A Bridge Between Virtual Scenarios and Real-World Issues

SGs provide simulated digital environments that bridge virtual experiences and real-world online challenges. Through scenario-based gameplay, adolescents can practice identifying cybersecurity risks, managing information overload, and protecting personal privacy in a safe and controlled setting. Such contextualized experiences enhance immersion and facilitate the transfer of learned skills to real online environments [33].

Formal Innovation: Dual Enhancement of Engagement and Learning Outcomes

The formal design of SGs enhances both engagement and learning outcomes. Rich audiovisual elements, narrative structures, and interactive feedback mechanisms increase adolescents' motivation and sustained participation [19]. The combination of contextual authenticity and interactivity further strengthens experiential learning and supports behavioral adaptation [26].

In summary, SGs integrate knowledge acquisition, emotional engagement, and behavioral practice within interactive digital contexts, closely aligning with the multidimensional nature of IA. Despite these theoretical and practical advantages, empirical research explicitly examining how SGs foster adolescents' IA remains scarce [26]. To address this gap, the following section examines the psychological mechanisms through which SGs support IA across the stages of preparation, adaptation, and sustained adaptation, thereby laying the foundation for the proposed design principles.

A 3-Phase Dynamic Mechanism Model: How SGs Facilitate Adolescents' IA

Building on the alignment between SGs and IA, we propose a 3-phase dynamic mechanism explaining how SGs support adolescents' adaptive development online. The model views IA as a trait shaped through ongoing interaction with digital environments, reflecting a reciprocal individual–context relationship [1]. It includes 3 stages: preparation, adaptation, and sustained adaptation, through which SGs promote continuous improvement in adolescents' ability to navigate the online world.

The Role of SGs in the Preparation Stage

In the preparation stage of IA, SGs primarily support adolescents in building foundational internet knowledge and skills while cultivating constructive attitudes toward digital environments. These early cognitive and affective structures form the basis for subsequent adaptive behaviors.

Adolescents need to develop a positive internet attitude, including cognitive, emotional, and behavioral tendencies, which influence responsible online behavior [34,35]. Simultaneously, internet knowledge and skills, such as information search, online communication, and cybersecurity practices, are essential for online activities [1]. These constructs interact: a positive attitude motivates learning, while improved skills strengthen engagement and perception of the internet [36].

SGs are well positioned to facilitate this preparatory learning process. Their immersive, interactive, and narrative-rich environments enable learners to construct initial cognitive frameworks about the online world [32]. By integrating technology, pedagogy, and interactive mechanisms [37], these games provide contextualized learning experiences that make abstract concepts concrete and engaging [38]. From the perspective of schema theory, such contextualized learning environments are particularly effective for supporting the formation and refinement of cognitive schema [39]. Schemas

function as early cognitive structures that organize new information, shape expectations, and influence attitudes. Some studies emphasize that learners encode and organize information more efficiently when it is embedded in meaningful, situational contexts that can be connected to prior knowledge [40–42]. Contextualized learning allows adolescents to connect new information to prior knowledge, promoting deeper understanding and stable orientations toward online environments.

Observational learning further reinforces early learning. According to Bandura's social cognitive theory, individuals acquire knowledge, attitudes, and behavioral strategies by attending to models and evaluating the consequences of their actions [43]. This process is supported by four subprocesses (attention, retention, reproduction, and motivation), which together determine how modeled behaviors are internalized [43]. Within SGs, these subprocesses are activated through designed modeling features. For example, non-player characters, scripted demonstrations, or peer-like avatars provide examples of digital norms, information-seeking strategies, and responses to common online dilemmas [21]. Empirical studies indicate that observing modeled responses in interactive digital environments can improve learners' ability to identify risks [23], foster more prosocial online attitudes [44,45], and support the development of normative expectations for digital interactions [46].

During the preparation stage, observational learning shapes early cognitive and affective orientations. Modeled experiences provide structured exposure to desirable online behaviors, allowing adolescents to internalize norms before independently navigating online environments.

Through the integrated mechanisms of schema formation and observational learning, SGs provide a safe, immersive learning environment. This enables adolescents to effectively construct internet knowledge and skills, cultivate a positive attitude, and actively explore the online environment, thereby laying a solid cognitive and affective foundation for the development of robust IA.

The Role of SGs in the Adaptation Process Stage

Beyond mastering internet skills, adolescents need a sense of control over the online environment. During adaptation, SGs help develop this psychological control and enhance self-efficacy, fostering confidence in handling online challenges.

Sense of control refers to one's belief in influencing outcomes, and this is a crucial factor in environmental adaptation [47]. Adolescents with higher control perceptions show greater agency and more positive evaluations of their abilities and influence on external events [48,49]. This perception is closely linked to self-efficacy, as believing in one's influence over outcomes enhances confidence in personal capabilities [50]. In a rapidly evolving digital space, adolescents with stronger control perceptions are more resilient to challenges. SGs, with their immersive and risk-free environments, allow adolescents to explore cyberspace, reinforcing their psychological control and self-efficacy.

Feedback mechanisms and progress visualization are central to this process. In SGs, players receive immediate feedback for

their actions, which strengthens their sense of control and helps establish a sense of order in digital environments. When unexpected outcomes occur, such as errors, timely prompts and guidance support the acquisition of internet knowledge and skills [51]. Through iterative trial and error, adolescents refine internet skills, ultimately applying them with confidence in real life.

SGs also provide structured environments that support the accumulation of mastery experiences, the most influential source of self-efficacy according to Bandura [43]. Research confirms that robust internet knowledge and experience significantly boost self-efficacy [52]. Therefore, SGs, by providing simulated online environments, allow adolescents to accumulate mastery experiences as they complete tasks, effectively strengthening their sense of online control and self-efficacy. Additionally, SGs set challenging tasks and achievement systems that gradually enhance players' self-efficacy. In these games, players overcome a series of challenges, complete tasks, and receive rewards, and these positive experiences further enhance their self-efficacy [53]. When adolescents perceive their abilities and accomplishments in the online space, they are more likely to transfer this confidence and motivation to real life, facing challenges in the online environment with greater composure.

In summary, SGs enhance IA through scenario simulations, real-time feedback, and achievement systems. They strengthen adolescents' sense of control and self-efficacy, equipping them to navigate the digital world with confidence and resilience.

The Role of SGs in the Sustained Adaptation Stage

In the sustained adaptation stage, SGs consolidate earlier learning and support the transfer of in-game experiences to real-world online contexts, reinforcing adolescents' self-regulation, control, resilience, and proactive engagement.

SGs use scenario-based instructional design to promote learning transfer, defined as the application of acquired knowledge and skills to new situations [54]. Contextualization involves embedding learning content within a scenario that closely mirrors real-world situations or practical problems [1]. Constructivist theory emphasizes that transfer depends on the alignment between learning contexts and real-world application contexts [55,56], a principle operationalized in SGs through realistic online simulations. For instance, cybersecurity games place players in expert roles, requiring them to manage cyberattacks and digital threats, thereby strengthening practical skill acquisition [57]. Activity theory further suggests that knowledge transfer efficiency depends on how well virtual activities align with real-world contexts [58]. Accordingly, the high contextual similarity between SGs and authentic online environments facilitates the efficient application of in-game learning to real-life digital situations.

Beyond supporting transfer, SGs offer scalable and flexible learning opportunities. Unlike traditional educational models, they do not depend on continuous teacher guidance or fixed learning spaces. Learners can engage in game-based learning activities at their own pace and according to their personal schedules, significantly enhancing learning continuity, flexibility, and autonomy [18]. Additionally, SGs can be updated

in real time to reflect emerging technologies and online risks, ensuring that adolescents' knowledge and skills remain aligned with evolving digital environments [22,23]. By transforming from static tools into dynamic platforms, SGs adapt to the ever-changing online landscape, offering a deeper and more practical experience to enhance IA.

In sum, by combining contextualized simulation, flexible access, and continuous updating, SGs support the sustained development of adolescents' IA and promote durable, transferable learning outcomes.

The Dynamic Enhancement Role of SGs Across the 3 Stages

SGs enhance adolescents' IA not only through the distinct functions of preparation, adaptation, and sustained adaptation, but also through their cyclical and progressive integration. Together, these stages form a dynamic learning mechanism that supports continuous reflection, application, and long-term adaptability.

Self-regulated learning (SRL) theory helps explain this process. Zimmerman [59] conceptualizes SRL as a cycle of planning, monitoring, and reflection that enables learners to regulate their behavior through goal setting, process evaluation, and experience review [53]. These processes align closely with the 3 stages of IA. In the preparation stage, SGs stimulate motivation and initiate the planning phase by encouraging adolescents to set learning goals, such as improving information search or cybersecurity awareness. In the adaptation stage, games support monitoring and strategy adjustment through immediate feedback and adaptive task design, requiring learners to apply prior knowledge in increasingly complex situations [59]. This iterative practice enhances focus, strengthens psychological control, and promotes self-efficacy. In the sustained adaptation stage, SGs consolidate learning outcomes through task progression and knowledge transfer. Reflection becomes especially prominent, as summary modules and performance tracking support evaluation of learning strategies and outcomes [60].

Across these stages, SGs create a spiral learning process. Goal setting in preparation lays the foundation for practice, experiences in adaptation deepen understanding, and reflection in sustained adaptation initiates the next learning cycle [59]. By integrating SRL processes with progressive game design, SGs not only enhance adolescents' current IA but also cultivate long-term SRL abilities. These skills enable adolescents to continually update and apply knowledge in an evolving digital environment, supporting sustained adaptability.

From Theory to Practice: Guiding Design Principles

Although SGs offer potential benefits, evidence shows their effectiveness is not guaranteed. Some studies report limited behavioral change or real-world transfer, including health behavior [61], safety knowledge [62], and cognitive or metacognitive skills [63]. In educational settings, games can boost engagement without significant learning gains [27,64],

or produce variable effects depending on design [65]. These findings highlight that outcomes depend more on design quality than on the game format itself [66].

The same applies to IA. Games promote adaptive development only when their goals, mechanics, and feedback support the cognitive, emotional, and behavioral processes essential for navigating digital environments. Translating theoretical mechanisms into practical design is therefore crucial. The following section presents key design principles for developing SGs that enhance adolescents' IA.

Principle of Clear Goal Definition

Clear goal definition is the foundational principle of IA-oriented SG design, as without explicit objectives, embedded mechanisms cannot reliably translate into meaningful learning outcomes. Within the SRL theory, the clear and specific goals guide learners' reflection, evaluation, and behavioral regulation throughout the learning process [67]. Constructivist learning theory similarly emphasizes defining the instructional theme for overall design and each unit [67]. However, some designers adopt existing game types without aligning them to specific educational goals, risking misalignment [18,68].

In the context of IA, goal clarity is particularly critical. IA comprises 8 core dimensions [1]: attitude, online knowledge and skills, sense of control, self-efficacy, adaptability, self-control, psychological resilience, and proactivity. These dimensions differ in importance across developmental stages: the preparation stage emphasizes attitudes and foundational skills, whereas the adaptation stage prioritizes sense of control and self-efficacy [1]. Accordingly, game objectives should be sequenced to reflect these developmental priorities.

Additionally, objectives must match adolescents' cognitive development. Overly complex or insufficiently scaffolded goals can increase cognitive load and hinder schema formation [69,70]. Finally, objectives should also be measurable to support progress monitoring. Standardized tests, surveys, and in-game analytics allow developers and educators to track learning and refine design [19]. Measurable goals also reinforce the SRL cycle by enabling players to compare current performance with desired outcomes and adjust strategies accordingly.

In sum, clear goal definition is not merely an administrative step; it operationalizes the psychological mechanisms underlying IA, transforming abstract adaptability constructs into actionable design elements that provide structural guidance for all subsequent design principles.

Principle of Interaction Diversity

Following clear goal definition, interaction diversity is the second major principle in IA-oriented SG design. While it supports all 3 stages of IA, its psychological functions align most closely with the preparation stage, where adolescents develop foundational knowledge, constructive attitudes, and early exploratory confidence [1]. Interaction diversity refers to offering multiple meaningful ways to engage with digital content, encouraging learners to interpret and respond to online situations from different perspectives [21,71]. It functions not merely as an interface feature but as a mechanism for deepening

cognitive processing, activating schemas, and strengthening early adaptive dispositions.

Varied interactions encourage adolescents to move beyond passive exposure and instead engage in increasingly elaborated cycles of exploration, interpretation, and decision-making [21]. Narrative-driven interactions, for instance, allow learners to engage with unfolding digital events in ways that highlight causality and perspective-taking [20]. Such narrative encounters support the activation and restructuring of social-cognitive schemas, enabling adolescents to develop more nuanced expectations regarding online behavior, interpersonal dynamics, and risk cues [22-24]. Procedural interactions, such as simulated online operations or information-handling tasks, enable the formation of procedural knowledge through repeated practice [23-25], while reinforcing early perceptions of competence and control. Socially oriented interactions situate learners within conversational or collaborative exchanges, modeling constructive digital communication and promoting empathy, negotiation, and responsible participation [45]. Reflective interactions, including prompts and consequence-based feedback, encourage learners to examine action outcomes, strengthening metacognitive monitoring and early risk evaluation [21,63]. Such reflective engagements are especially valuable in the preparation stage, where adolescents are forming their earliest internal standards for online judgment and behavior.

Taken together, interaction diversity enables adolescents to engage with digital scenarios through narrative, procedural, social, and reflective pathways, supporting early IA while laying the cognitive and affective foundation for later adaptation and sustained functioning.

Principle of Contextual Authenticity

Building on interaction diversity, contextual authenticity is a key principle for SGs enhancing adolescents' IA. Learning in environments resembling real digital contexts helps learners perceive the relevance of skills and decisions [72]. In such authentic scenarios, adolescents engage in experiential learning: they encounter realistic problems, make decisions, observe consequences, and reflect on outcomes [73,74]. This cycle of action, feedback, and reflection fosters mastery experiences, a primary source of self-efficacy [43,72,73], which in turn supports proactive coping and adaptive decision-making in real online environments [75].

However, far transfer from low-risk game settings to high-stakes real-world situations is often limited [63,65]. Realistic scenarios alone do not guarantee effective application of game-learned strategies. To bridge this gap, SGs should include structured scaffolding that supports cognitive and metacognitive transfer [76]. Increasing in-game task complexity and emotional demands allows learners to practice under conditions closer to real digital pressures [77]. This exposure strengthens stress tolerance and decision stability, allowing adolescents to form resilient strategy–response mappings that can generalize beyond the game. Simultaneously, the metacognitive prompts and reflective exercises encourage learners to articulate their reasoning, examine potential biases, and connect in-game experiences with real life [63]. This process transforms

gameplay experiences into conceptual knowledge, an essential condition for far transfer, which requires that learners not merely remember actions but understand underlying mechanisms [60]. Embedding negative feedback or conflict scenarios trains emotion regulation while maintaining effective decision-making [78]. Moreover, exposure to multiple roles and scenarios across digital contexts reinforces strategy application and consolidation, increasing the likelihood of successful transfer [45,66].

In this way, contextual authenticity, when supported by structured scaffolding, functions not merely as a realistic setting but as an integrated mechanism that deepens comprehension, strengthens self-efficacy, and facilitates the flexible application of adaptive strategies.

Principle of Immediate, Scaffolding, and Explanatory Feedback

In addition to the diversity and authenticity of interaction processes, the results of each interaction are equally important for improving adolescents' learning experience and motivation in SGs [58]. In SGs focused on IA, feedback should be immediate, explanatory, and scaffolded according to developmental stages, directly fostering adaptive self-efficacy, a strong sense of control, and metacognitive regulation [43,79]. Effective feedback establishes a transparent loop between digital behaviors and their consequences, allowing adolescents to rapidly form accurate mental models of online causal relationships [80].

To cultivate adaptability, feedback must go beyond correctness judgments and act as a dynamic scaffold. During the preparation stage, feedback should be highly supportive and formative, linking actions to positive outcomes to reinforce initial schemata and build foundational confidence [81,82]. For example, successfully verifying a website might trigger feedback that affirms the behavior and explains how it enhances digital safety [83]. As learners progress into adaptation and sustained adaptation stages, feedback should increasingly include corrective and explanatory components following not optimal choices. By simulating plausible negative outcomes and providing guided analysis, these interventions transform errors into low-stakes mastery experiences, a process that is critical for developing resilience and flexible problem-solving skills [84]. From the perspective of self-efficacy theory, these scaffolded feedback cycles are the engine for generating mastery experiences, the source of confidence in managing digital challenges [57,82]. Crucially, by explaining the "why" behind both successes and failures, feedback fosters metacognitive awareness, enabling adolescents to understand not only what to do but how to think about digital situations [63]. This lays the foundation for self-regulation and the strategic transfer of skills to new, real-world online contexts [74].

In sum, this principle redefines feedback from a mere informational tool to a key driver of adaptive psychological development, with in-game interaction contributing to both skill acquisition and the cultivation of beliefs and self-regulatory capacities that sustain long-term IA.

Principle of Dynamically Adaptive Learning Environment

The principles of clear goals, contextual authenticity, interaction diversity, and immediate feedback provide the framework for game-based learning [85]. To extend this foundation into sustainable adaptability, IA-oriented games should function as dynamically adaptive learning environments, operating at both content and individualized learner levels.

At the content level, dynamic adaptability ensures that the game continually integrates new digital challenges reflecting emerging norms, technologies, and risks [86,87]. Rather than presenting fixed modules, the game introduces novel scenarios over time, supporting repeated experiential learning and adaptive transfer [60]. Updates act as pedagogical extensions of emerging digital realities. For example, modules may present artificial intelligence (AI)-related challenges, such as identifying deepfake content, evaluating algorithmically curated information, or managing interactions with AI-driven platforms, helping adolescents respond to evolving online risks [87]. As players confront these tasks, they are encouraged to revisit earlier strategies, refine their understanding, and rebuild adaptive approaches. This iterative process strengthens metacognitive awareness, self-regulation, and resilience [88,89]. Embedded assessments can further personalize progression by adjusting difficulty or focus based on performance, cognitive style, or competencies. For example, a player who excels in online communication but lacks cybersecurity skills may receive additional tasks focused on privacy management or digital ethics [90].

At the individual level, dynamic adaptability is supported through an intuitive and well-organized interface that matches learners' cognitive habits and operational routines [70,91]. Such design reduces extraneous cognitive load and enables learners to concentrate on meaningful problem-solving rather than procedural navigation. By lowering technical barriers and offering clear guidance, the interface functions as a scaffold that facilitates efficient information processing, responsible decision-making, and reflective engagement [92,93]. When combined with dynamically evolving content, this learner-centered support helps players apply adaptive strategies in real time and internalize the self-regulatory and metacognitive skills required for navigating complex digital environments.

In sum, the principle highlights how SGs can integrate evolving content with accessible, learner-centered design. This dual adaptability nurtures reflective, self-regulated, and transferable skills, equipping adolescents with the cognitive and behavioral capacities needed for sustainable IA.

Principle of Safety-by-Design for Digital Well-Being

While the preceding principles focus on fostering motivation, cognitive engagement, and effective skill acquisition, Safety-by-Design ensures that the learning environment itself does not inadvertently reproduce the addictive or compulsive engagement patterns associated with negative forms of internet use [94]. Adolescents are particularly sensitive to reward-seeking cues, variable reinforcement schedules, and emotionally arousing feedback structures, all of which can

amplify screen fixation or habitual checking behaviors [95]. To counter these risks, safety-by-design establishes boundaries that promote regulated, purposeful engagement rather than hedonic immersion. For instance, gameplay cycles are intentionally time-bounded and punctuated by reflective pauses that encourage players to monitor their internal states, evaluate the usefulness of the strategies they are using, and recognize the distinction between intentional learning and habitual digital consumption [96]. Moreover, the reinforcement mechanisms embedded in the game prioritize informational over affective feedback: success is communicated through clarity of consequences, improvement indicators, or strategy-relevant insights rather than through escalating stimuli such as points, streaks, or attention-grabbing animations [97]. By avoiding variable structures and other persuasive design elements known to increase compulsive use [98], the game models a healthier digital ecology that aligns with the broader objective of cultivating adaptive self-regulation.

Additionally, Safety-by-Design further supports IA by fostering metacognitive awareness of one's digital habits [96]. While the structural boundaries of the game help reduce the likelihood of compulsive use, the next step is enabling adolescents to understand and actively manage their own digital habits. Break reminders, self-assessment checkpoints, and reflective prompts, which have been shown to foster self-regulation and transfer in digital learning environments [63,96]. These opportunities for reflection help adolescents identify moments when online interactions may compromise their well-being and consider how the strategies practiced in the game could inform their responses in everyday digital contexts.

Together, the structural safeguards and metacognitive supports enable Safety-by-Design to move beyond simply preventing harmful engagement, actively fostering adolescents' self-regulation and reflective capacities for adaptive, intentional, and sustainable digital behavior.

Conclusions and Future Research

This review examined how SGs enhance adolescents' IA through the 3 stages of preparation, adaptation, and sustained

adaptation. By leveraging contextualized design, real-time feedback, and dynamic tasks, these games support experiential learning, psychological control, skill acquisition, and resilience to online risks. Based on these insights, we proposed design principles to guide the development of IA-oriented SGs.

However, this work remains theoretical and requires empirical validation. Future studies should develop and evaluate SGs targeting age-specific adaptability traits across diverse online environments, considering both user experience and educational effectiveness. Additionally, future research could investigate how IA manifests across culturally and digitally diverse contexts. Drawing on ecological and social-ecological frameworks [99], researchers can examine how nested environmental systems, including family, school, peer networks, and broader community and sociocultural contexts, shape both the types of online challenges adolescents encounter and the strategies they use to adapt. These studies would clarify how IA is shaped by different environments and inform culturally sensitive interventions and SG design.

With rapid advances in AI, adolescents face increasing demands, such as evaluating information authenticity and responding to novel digital risks. Future game designs can integrate AI-driven situational simulations and real-time adaptive feedback to personalize learning, simulate AI-generated risks, and help adolescents practice verification, problem-solving, and adaptive strategies [86,87,100].

In conclusion, SGs, as an innovative educational tool, hold great potential for improving adolescents' IA. Future research should further combine technological innovations with empirical studies to explore their application across varied educational settings. At the same time, effective implementation requires collaboration among policymakers, schools, and parents to build a supportive internet ecosystem. With ongoing efforts, SGs can become an essential tool in adolescent internet education, providing sustained support for their learning and development.

Acknowledgments

Artificial intelligence tools (specifically ChatGPT) were used only for language polishing to improve grammar, clarity, and readability. All intellectual content and manuscript organization were solely developed and verified by the authors.

JH and SM contributed equally to this work and are co-corresponding authors.

Funding

This work was supported by the National Natural Science Foundation of China (NNSFC, Grant No. 71974072) and the Fundamental Research Funds for the Central Universities (CCNU23KYZHSY17; CCNU24ZZ042).

Authors' Contributions

Methodology, writing-original draft, writing – review and editing: SG

Methodology, validation, writing – review and editing: MJ

Conceptualization: WW

Conceptualization: JH

Supervision, writing – review and editing: SM

All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Wang W, Dong R, Niu G, Zhou Z. Network adaptation: concepts and models. *J Nanchang Univ (Humanities Soc Sci Ed)* 2021;52(2):45-54 [[FREE Full text](#)]
2. Uchino BN. Understanding the links between social support and physical health: a life-span perspective with emphasis on the separability of perceived and received support. *Perspect Psychol Sci* 2009 May;4(3):236-255. [doi: [10.1111/j.1745-6924.2009.01122.x](#)] [Medline: [26158961](#)]
3. Tinmaz H, Lee YT, Fanea-Ivanovici M, Baber H. A systematic review on digital literacy. *Smart Learn Environ* 2022;9(1):21. [doi: [10.1186/s40561-022-00204-y](#)] [Medline: [40478098](#)]
4. Yeşilyurt E, Vezne R. Digital literacy, technological literacy, and internet literacy as predictors of attitude toward applying computer-supported education. *Educ Inf Technol (Dordr)* 2023 Jan 18;28(8):1-27. [doi: [10.1007/s10639-022-11311-1](#)] [Medline: [36688220](#)]
5. Carretero Gomez S, Vuorikari R, Punie Y. DigComp 2.1: the digital competence framework for citizens with eight proficiency levels and examples of use. : Publications Office of the European Union; 2017. [doi: [10.2760/38842](#)]
6. Pan Q, Lan M, Tan CY, Tao S, Liang Q, Law N. Protective factors contributing to adolescents' multifaceted digital resilience for their wellbeing: a socio-ecological perspective. *Comput Human Behav* 2024 Jun;155:108164. [doi: [10.1016/j.chb.2024.108164](#)]
7. Arkan Z, Bal M. The relationship between school-age students' literacy skills and digital well-being: a systematic review. *BMC Psychol* 2025 Nov 10;13(1):1240. [doi: [10.1186/s40359-025-03573-4](#)] [Medline: [41214781](#)]
8. Wang W, Ma S, Han X, Zhao X. The impact of internet adaptability on internet addiction: the serial mediation effect of meaning in life and anxiety. *Front Psychiatry* 2023;14:1268539. [doi: [10.3389/fpsy.2023.1268539](#)] [Medline: [38148745](#)]
9. Chen X, Wang L, Wang Z. Shyness-sensitivity and social, school, and psychological adjustment in rural migrant and urban children in China. *Child Dev* 2009;80(5):1499-1513. [doi: [10.1111/j.1467-8624.2009.01347.x](#)] [Medline: [19765014](#)]
10. Jin CC, Wang BC, Ji AT. The relationship between the dark triad and internet adaptation among adolescents in China: internet use preference as a mediator. *Front Psychol* 2019;10:2023. [doi: [10.3389/fpsyg.2019.02023](#)] [Medline: [31543856](#)]
11. Dong WH, Zhang J, Meng SJ, Jia M, Wang WJ. The topological structure of adolescents' internet adaptation: a longitudinal tracking study. *Acta Psychologica Sinica* 2025;57(3):415. [doi: [10.3724/SP.J.1041.2025.0415](#)]
12. Livingstone S, Helsper EJ. Balancing opportunities and risks in teenagers' use of the internet: the role of online skills and internet self-efficacy. *New Media & Society* 2010 Mar;12(2):309-329. [doi: [10.1177/1461444809342697](#)]
13. Abt CC. *Serious Games*: Viking Press; 1970. URL: <https://www.worldcat.org/title/58218> [accessed 2026-02-28]
14. Franco Vega I, Eleftheriou A, Graham C. Using video games to improve the sexual health of young people aged 15 to 25 years: rapid review. *JMIR Serious Games* 2022 May 19;10(2):e33207. [doi: [10.2196/33207](#)] [Medline: [35587873](#)]
15. Zyda M. From visual simulation to virtual reality to games. *Computer (Long Beach Calif)* 2005;38(9):25-32. [doi: [10.1109/MC.2005.297](#)]
16. Lameris P, Arnab S, Dunwell I, Stewart C, Clarke S, Petridis P. Essential features of serious games design in higher education: linking learning attributes to game mechanics. *Brit J Educational Tech* 2017 Jun;48(4):972-994. [doi: [10.1111/bjet.12467](#)]
17. Li F, Sun Y. Functional games: definition, value exploration, and development suggestions. *Educ Media Res* 2019;1:65-68 [[FREE Full text](#)]
18. Tay J, Goh YM, Safiena S, Bound H. Designing digital game-based learning for professional upskilling: a systematic literature review. *Comput Educ* 2022 Jul;184:104518. [doi: [10.1016/j.compedu.2022.104518](#)]
19. Caserman P, Hoffmann K, Müller P, et al. Quality criteria for serious games: serious part, game part, and balance. *JMIR Serious Games* 2020 Jul 24;8(3):e19037. [doi: [10.2196/19037](#)] [Medline: [32706669](#)]
20. Naul E, Liu M. Why story matters: a review of narrative in serious games. *J Educ Comput Res* 2020 Jun;58(3):687-707. [doi: [10.1177/0735633119859904](#)]
21. Fonseca X, Slingerland G, Lukosch S, Brazier F. Designing for meaningful social interaction in digital serious games. *Entertain Comput* 2021 Jan;36:100385. [doi: [10.1016/j.entcom.2020.100385](#)]
22. Calvo-Morata A, Alonso-Fernández C, Freire M, Martínez-Ortiz I, Fernández-Manjón B. Serious games to prevent and detect bullying and cyberbullying: a systematic serious games and literature review. *Comput Educ* 2020 Nov;157:103958. [doi: [10.1016/j.compedu.2020.103958](#)]
23. Hart S, Margheri A, Paci F, Sassone V. Riskio: a serious game for cyber security awareness and education. *Computers & Security* 2020 Aug;95:101827. [doi: [10.1016/j.cose.2020.101827](#)]
24. Ahmead M, El Sharif N, Abuiram I. Risky online behaviors and cybercrime awareness among undergraduate students at Al Quds University: a cross sectional study. *Crime Sci* 2024;13(1):29. [doi: [10.1186/s40163-024-00230-w](#)]

25. Booc NBB, Budiongan K, Carballo R. Cybersecurity awareness, and cybersecurity behavior of high school students in Davao city: a mediation role of perceived behavioral control. *Eur J Appl Sci Eng Technol* 2024;2(3):4-9. [doi: [10.59324/ejaset.2024.2\(3\).01](https://doi.org/10.59324/ejaset.2024.2(3).01)]
26. Wang W, Li J, Liu S, Ye J, Ma S. Development and validation of an adolescent internet adaptability educational game based on situated learning theory. *Educ Inf Technol* 2025 Oct;30(15):21573-21608. [doi: [10.1007/s10639-025-13608-3](https://doi.org/10.1007/s10639-025-13608-3)]
27. Riopel M, Nenciovici L, Potvin P, et al. Impact of serious games on science learning achievement compared with more conventional instruction: an overview and a meta-analysis. *Studies in Science Education* 2019 Jul 3;55(2):169-214. [doi: [10.1080/03057267.2019.1722420](https://doi.org/10.1080/03057267.2019.1722420)]
28. Behl A, Jayawardena N, Pereira V, Islam N, Giudice MD, Choudrie J. Gamification and e-learning for young learners: a systematic literature review, bibliometric analysis, and future research agenda. *Technol Forecast Soc Change* 2022 Mar;176:121445. [doi: [10.1016/j.techfore.2021.121445](https://doi.org/10.1016/j.techfore.2021.121445)]
29. Ullah H, Afzal S, Khan IU. Perceptual quality assessment of panoramic stitched contents for immersive applications: a prospective survey. *VRH* 2022 Jun;4(3):223-246. [doi: [10.1016/j.vrih.2022.03.004](https://doi.org/10.1016/j.vrih.2022.03.004)]
30. Yu Z. Learning outcomes, motivation, and satisfaction in gamified english vocabulary learning. *Sage Open* 2023 Apr;13(2). [doi: [10.1177/21582440231158332](https://doi.org/10.1177/21582440231158332)]
31. Almeida F, Buzady Z. Development of soft skills competencies through the use of FLIGBY. *Technol Pedagog Educ* 2022 Aug 8;31(4):417-430. [doi: [10.1080/1475939X.2022.2058600](https://doi.org/10.1080/1475939X.2022.2058600)]
32. tefan IA, Hauge JB, Hasse F, tefan A. Using serious games and simulations for teaching co-operative decision-making. *Procedia Comput Sci* 2019;162:745-753. [doi: [10.1016/j.procs.2019.12.046](https://doi.org/10.1016/j.procs.2019.12.046)]
33. Bakhanova E, Garcia JA, Raffe WL, Voinov A. Targeting social learning and engagement: what serious games and gamification can offer to participatory modeling. *Environ Model Softw* 2020 Dec;134:104846. [doi: [10.1016/j.envsoft.2020.104846](https://doi.org/10.1016/j.envsoft.2020.104846)]
34. Tsai CC, Lin SSJ, Tsai MJ. Developing an Internet Attitude Scale for high school students. *Comput Educ* 2001 Aug;37(1):41-51. [doi: [10.1016/S0360-1315\(01\)00033-1](https://doi.org/10.1016/S0360-1315(01)00033-1)]
35. Tu W, Wang M, Li Y, Nie Y. The relationship between college students' network attitudes and internet addiction: a chain mediation effect. *Chin J Health Psychol* 2022;30(6):905-909 [FREE Full text]
36. Brandtzæg PB. Towards a unified media-user typology (MUT): a meta-analysis and review of the research literature on media-user typologies. *Comput Human Behav* 2010 Sep;26(5):940-956. [doi: [10.1016/j.chb.2010.02.008](https://doi.org/10.1016/j.chb.2010.02.008)]
37. Deng Y, Zhang Y, Zhao R. Investigating learning behaviors in desktop-based simulated and VR headset-based immersive 3D learning environments: a cross-media comparative study. *Educ Inf Technol* 2025 Jul;30(11):15441-15465. [doi: [10.1007/s10639-025-13411-0](https://doi.org/10.1007/s10639-025-13411-0)]
38. Makransky G, Petersen GB. The cognitive affective model of immersive learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. *Educ Psychol Rev* 2021 Sep;33(3):937-958. [doi: [10.1007/s10648-020-09586-2](https://doi.org/10.1007/s10648-020-09586-2)]
39. Anderson RC. Role of the reader's schema in comprehension, learning, and memory. In: Alvermann DE, Unrau NJ, Sailors M, Ruddell RB, editors. *Theoretical Models and Processes of Literacy*: Routledge; 2018:136-145. [doi: [10.4324/9781315110592-9](https://doi.org/10.4324/9781315110592-9)]
40. Jung E, Lim R, Kim D. A schema-based instructional design model for self-paced learning environments. *Education Sciences* 2022;12(4):271. [doi: [10.3390/educsci12040271](https://doi.org/10.3390/educsci12040271)]
41. Ye L, Zhou X, Yang S, Hang Y. Serious game design and learning effect verification supporting traditional pattern learning. *Interactive Learning Environments* 2023 Dec 15;31(10):6573-6587. [doi: [10.1080/10494820.2022.2042032](https://doi.org/10.1080/10494820.2022.2042032)]
42. Petersen GB, Petkakis G, Makransky G. A study of how immersion and interactivity drive VR learning. *Comput Educ* 2022 Apr;179:104429. [doi: [10.1016/j.compedu.2021.104429](https://doi.org/10.1016/j.compedu.2021.104429)]
43. Bandura A. *Self-efficacy: The Exercise of Control*: W. H. Freeman; 1997. [doi: [10.1037/10522-094](https://doi.org/10.1037/10522-094)]
44. Ferreira PC, Simão AMV, Paiva A, Martinho C, Prada R, Rocha J. Serious game-based psychosocial intervention to foster prosociality in cyberbullying bystanders. *Psychosoc Interv* 2022 May;31(2):83-96. [doi: [10.5093/pi2022a5](https://doi.org/10.5093/pi2022a5)] [Medline: [37360059](https://pubmed.ncbi.nlm.nih.gov/37360059/)]
45. Bachen CM, Hernández-Ramos P, Raphael C, Waldron A. How do presence, flow, and character identification affect players' empathy and interest in learning from a serious computer game? *Comput Human Behav* 2016 Nov;64:77-87. [doi: [10.1016/j.chb.2016.06.043](https://doi.org/10.1016/j.chb.2016.06.043)]
46. Gentile DA, Anderson CA, Yukawa S, et al. The effects of prosocial video games on prosocial behaviors: international evidence from correlational, longitudinal, and experimental studies. *Pers Soc Psychol Bull* 2009 Jun;35(6):752-763. [doi: [10.1177/0146167209333045](https://doi.org/10.1177/0146167209333045)] [Medline: [19321812](https://pubmed.ncbi.nlm.nih.gov/19321812/)]
47. Skinner EA. *Perceived Control, Motivation, & Coping*: Sage; 1996. [doi: [10.4135/9781483327198](https://doi.org/10.4135/9781483327198)]
48. 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](https://doi.org/10.1037//0003-066x.55.1.68)] [Medline: [11392867](https://pubmed.ncbi.nlm.nih.gov/11392867/)]
49. Frazier P, Keenan N, Anders S, Perera S, Shallcross S, Hintz S. Perceived past, present, and future control and adjustment to stressful life events. *J Pers Soc Psychol* 2011 Apr;100(4):749-765. [doi: [10.1037/a0022405](https://doi.org/10.1037/a0022405)] [Medline: [21299308](https://pubmed.ncbi.nlm.nih.gov/21299308/)]
50. Li Y. The relationship between college students' locus of control, academic self-efficacy, and learning autonomy [master's thesis]. : Hunan Normal University; 2016 URL: <https://kns.cnki.net/kcms2/article/>

- <https://doi.org/10.1111/jcc4.12009> [accessed 2026-03-03]
51. Hamari J, Koivisto J, Sarsa H. Does gamification work? -- A literature review of empirical studies on gamification. Presented at: 47th Hawaii International Conference on System Sciences (HICSS); Jan 6-9, 2014; Waikoloa, HI p. 3025-3034. [doi: [10.1109/HICSS.2014.377](https://doi.org/10.1109/HICSS.2014.377)]
 52. Anderson AA, Brossard D, Scheufele DA, Xenos MA, Ladwig P. The “Nasty Effect:” online incivility and risk perceptions of emerging technologies. *J Comput-Mediat Comm* 2014 Apr;19(3):373-387. [doi: [10.1111/jcc4.12009](https://doi.org/10.1111/jcc4.12009)]
 53. Arnab S, Lim T, Carvalho MB, et al. Mapping learning and game mechanics for serious games analysis. *Brit J Educational Tech* 2015 Mar;46(2):391-411. [doi: [10.1111/bjet.12113](https://doi.org/10.1111/bjet.12113)]
 54. Prenzel M, Mandl H. Transfer of learning: considerations for the design of learning environments. In: Frensch PA, Funke J, editors. *Complex Problem Solving: The European Perspective*: Lawrence Erlbaum Associates, Inc; 1993:403-420 URL: https://link.springer.com/chapter/10.1007/978-3-642-78069-1_16 [accessed 2026-02-28]
 55. Brown JS, Collins A, Duguid P. Situated cognition and the culture of learning. *Educ Res* 1989 Jan;18(1):32-42. [doi: [10.3102/0013189X018001032](https://doi.org/10.3102/0013189X018001032)]
 56. Barnett SM, Ceci SJ. When and where do we apply what we learn? A taxonomy for far transfer. *Psychol Bull* 2002 Jul;128(4):612-637. [doi: [10.1037/0033-2909.128.4.612](https://doi.org/10.1037/0033-2909.128.4.612)] [Medline: [12081085](https://pubmed.ncbi.nlm.nih.gov/12081085/)]
 57. Roque C, Moodley G, Mandal S. Cybersafe: gamifying cybersecurity training with a training app. *iccws* 2024;19(1):299-307. [doi: [10.34190/iccws.19.1.2198](https://doi.org/10.34190/iccws.19.1.2198)]
 58. Hamari J, Shernoff DJ, Rowe E, Coller B, Asbell-Clarke J, Edwards T. Challenging games help students learn: an empirical study on engagement, flow and immersion in game-based learning. *Comput Human Behav* 2016 Jan;54:170-179. [doi: [10.1016/j.chb.2015.07.045](https://doi.org/10.1016/j.chb.2015.07.045)]
 59. Zimmerman BJ. Becoming a self-regulated learner: an overview. *Theory Pract* 2002 May;41(2):64-70. [doi: [10.1207/s15430421tip4102_2](https://doi.org/10.1207/s15430421tip4102_2)]
 60. Sailer M, Homner L. The gamification of learning: a meta-analysis. *Educ Psychol Rev* 2020 Mar;32(1):77-112. [doi: [10.1007/s10648-019-09498-w](https://doi.org/10.1007/s10648-019-09498-w)]
 61. DeSmet A, Van Ryckeghem D, Compennolle S, et al. A meta-analysis of serious digital games for healthy lifestyle promotion. *Prev Med* 2014 Dec;69:95-107. [doi: [10.1016/j.ypmed.2014.08.026](https://doi.org/10.1016/j.ypmed.2014.08.026)] [Medline: [25172024](https://pubmed.ncbi.nlm.nih.gov/25172024/)]
 62. Dankbaar MEW, Richters O, Kalkman CJ, et al. Comparative effectiveness of a serious game and an e-module to support patient safety knowledge and awareness. *BMC Med Educ* 2017 Feb 2;17(1):30. [doi: [10.1186/s12909-016-0836-5](https://doi.org/10.1186/s12909-016-0836-5)] [Medline: [28148296](https://pubmed.ncbi.nlm.nih.gov/28148296/)]
 63. Zumbach J, Rammerstorfer L, Deibl I. Cognitive and metacognitive support in learning with a serious game about demographic change. *Comput Human Behav* 2020 Feb;103:120-129. [doi: [10.1016/j.chb.2019.09.026](https://doi.org/10.1016/j.chb.2019.09.026)]
 64. Tsai YL, Tsai CC. A meta - analysis of research on digital game - based science learning. *Computer Assisted Learning* 2020 Jun;36(3):280-294. [doi: [10.1111/jcal.12430](https://doi.org/10.1111/jcal.12430)]
 65. Gui Y, Cai Z, Yang Y, Kong L, Fan X, Tai RH. Effectiveness of digital educational game and game design in STEM learning: a meta-analytic review. *IJ STEM Ed* 2023;10(1):36. [doi: [10.1186/s40594-023-00424-9](https://doi.org/10.1186/s40594-023-00424-9)]
 66. Clark DB, Tanner-Smith EE, Killingsworth SS. Digital games, design, and learning: a systematic review and meta-analysis. *Rev Educ Res* 2016 Mar;86(1):79-122. [doi: [10.3102/0034654315582065](https://doi.org/10.3102/0034654315582065)] [Medline: [26937054](https://pubmed.ncbi.nlm.nih.gov/26937054/)]
 67. Jonassen DH. Designing constructivist learning environments. In: Reigeluth CM, editor. *Instructional-Design Theories and Models: A New Paradigm of Instructional Theory*: Lawrence Erlbaum Associates; 1999:215-239. [doi: [10.4324/9781410603784-12](https://doi.org/10.4324/9781410603784-12)]
 68. Connolly TM, Boyle EA, MacArthur E, Hainey T, Boyle JM. A systematic literature review of empirical evidence on computer games and serious games. *Comput Educ* 2012 Sep;59(2):661-686. [doi: [10.1016/j.compedu.2012.03.004](https://doi.org/10.1016/j.compedu.2012.03.004)]
 69. Plass JL, Homer BD, Kinzer CK. Foundations of game-based learning. *Educ Psychol* 2015 Oct 2;50(4):258-283. [doi: [10.1080/00461520.2015.1122533](https://doi.org/10.1080/00461520.2015.1122533)]
 70. Sweller J, Ayres P, Kalyuga S. *Cognitive Load Theory*: Springer; 2011. [doi: [10.1007/978-1-4419-8126-4](https://doi.org/10.1007/978-1-4419-8126-4)]
 71. Chen CW. Low-tech serious games in higher education: bridging the digital divide and enhancing student thinking and performance. *Humanit Soc Sci Commun* 2025;12(1):111. [doi: [10.1057/s41599-024-04341-2](https://doi.org/10.1057/s41599-024-04341-2)]
 72. Merchant Z, Goetz ET, Cifuentes L, Keeney-Kennicutt W, Davis TJ. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: a meta-analysis. *Comput Educ* 2014 Jan;70:29-40. [doi: [10.1016/j.compedu.2013.07.033](https://doi.org/10.1016/j.compedu.2013.07.033)]
 73. Kolb DA. *Experiential Learning: Experience as the Source of Learning and Development*, 2nd edition: Pearson Education; 2015. URL: <https://dl.icdst.org/pdfs/files3/c960fd4ba0e6e861d2af44103707fc96.pdf> [accessed 2026-02-24]
 74. Ferchaud A, Beth Oliver M. It's my choice: the effects of moral decision-making on narrative game engagement. *J Gaming Virtual Worlds* 2019 Sep 1;11(2):101-118. [doi: [10.1386/jgvw.11.2.101_1](https://doi.org/10.1386/jgvw.11.2.101_1)]
 75. Lee AY, Hancock JT. Developing digital resilience: an educational intervention improves elementary students' response to digital challenges. *Computers and Education Open* 2023 Dec;5:100144. [doi: [10.1016/j.caeo.2023.100144](https://doi.org/10.1016/j.caeo.2023.100144)]
 76. Nietfeld JL, Anderson CA, Yukawa S. Predicting transfer from a game-based learning environment. *Comput Educ* 2020 Mar;146:103780. [doi: [10.1016/j.compedu.2019.103780](https://doi.org/10.1016/j.compedu.2019.103780)]

77. Richter K, Kickmeier-Rust M. The role of gamification in learning transfer: does early skill learning predict performance in complex tasks? *ECGBL* 2025 Sep 26;19(2):725-732. [doi: [10.34190/ecgbl.19.2.4009](https://doi.org/10.34190/ecgbl.19.2.4009)]
78. López-Serrano A, McGowan N, Moreno-Ger P, Burgos D. Teaching soft skills in higher education through serious games: validation of the compete! gamification. *Smart Learn Environ* 2025;12(1):49. [doi: [10.1186/s40561-025-00401-5](https://doi.org/10.1186/s40561-025-00401-5)]
79. van der Kleij FM, Eggen T, Timmers CF, Veldkamp BP. Effects of feedback in a computer-based assessment for learning. *Comput Educ* 2012 Jan;58(1):263-272. [doi: [10.1016/j.compedu.2011.07.020](https://doi.org/10.1016/j.compedu.2011.07.020)]
80. Maxim RI, Arnedo-Moreno J. Identifying key principles and commonalities in digital serious game design frameworks: scoping review. *JMIR Serious Games* 2025 Mar 5;13:e54075. [doi: [10.2196/54075](https://doi.org/10.2196/54075)] [Medline: [40053743](https://pubmed.ncbi.nlm.nih.gov/40053743/)]
81. Andersson C, Granberg C, Palmberg B, Palm T. Basic psychological needs satisfaction as a mediator of the effects of a formative assessment practice on behavioral engagement and autonomous motivation. *Front Educ* 2025;10:1523124. [doi: [10.3389/educ.2025.1523124](https://doi.org/10.3389/educ.2025.1523124)]
82. Guay F. Applying self-determination theory to education: regulations types, psychological needs, and autonomy supporting behaviors. *Can J Sch Psychol* 2022 Mar;37(1):75-92. [doi: [10.1177/08295735211055355](https://doi.org/10.1177/08295735211055355)]
83. Ravysse WS, Seugnet Bignaut A, Leendertz V, Woolner A. Success factors for serious games to enhance learning: a systematic review. *Virtual Real* 2017 Mar;21(1):31-58. [doi: [10.1007/s10055-016-0298-4](https://doi.org/10.1007/s10055-016-0298-4)]
84. Yasin A, Liu L, Li T, Fatima R, Jianmin W. Improving software security awareness using a serious game. *IET Software* 2019 Apr;13(2):159-169. [doi: [10.1049/iet-sen.2018.5095](https://doi.org/10.1049/iet-sen.2018.5095)]
85. Dabbous M, Kawtharani A, Fahs I, et al. The role of game-based learning in experiential education: tool validation, motivation assessment, and outcomes evaluation among a sample of pharmacy students. *Education Sciences* 2022;12(7):434. [doi: [10.3390/educsci12070434](https://doi.org/10.3390/educsci12070434)]
86. Pistono ADA, dos Santos AMP, Baptista RJV, Mamede HS. Framework for adaptive serious games. *Comp Applic In Engineering* 2024 Jul;32(4):e22731. [doi: [10.1002/cae.22731](https://doi.org/10.1002/cae.22731)]
87. Tolks D, Schmidt JJ, Kuhn S. The role of AI in serious games and gamification for health: scoping review. *JMIR Serious Games* 2024 Jan 15;12:e48258. [doi: [10.2196/48258](https://doi.org/10.2196/48258)] [Medline: [38224472](https://pubmed.ncbi.nlm.nih.gov/38224472/)]
88. Checa-Romero M, Gimenez-Lozano JM. Video games and metacognition in the classroom for the development of 21st century skills: a systematic review. *Front Educ* 2025;9:1485098. [doi: [10.3389/educ.2024.1485098](https://doi.org/10.3389/educ.2024.1485098)]
89. Zheng R, New J. The role of metacognition in learning via serious games. In: *Handbook of Research on Serious Games for Educational Environments*: IGI Global; 2017:19-40. [doi: [10.4018/978-1-5225-0513-6.ch002](https://doi.org/10.4018/978-1-5225-0513-6.ch002)]
90. Vanbecelaere S, Demedts F, Reynvoet B, Depaepe F. Toward a framework for analyzing adaptive digital games' research effectiveness. *IJSG* 2023;10(4):77-91. [doi: [10.17083/ijsg.v10i4.618](https://doi.org/10.17083/ijsg.v10i4.618)]
91. Carmichael L, Poirier SM, Coursaris CK, Léger PM, Sénécal S. Users' information disclosure behaviors during interactions with chatbots: the effect of information disclosure nudges. *Appl Sci (Basel)* 2022;12(24):12660. [doi: [10.3390/app122412660](https://doi.org/10.3390/app122412660)]
92. Carrión-Toro M, Morales-Martínez D, Santórum M, et al. Rethinking usability in serious games: designing an instrument that evaluates what really matters in learning contexts. *Appl Sci (Basel)* 2025;15(18):10084. [doi: [10.3390/app151810084](https://doi.org/10.3390/app151810084)]
93. Sweller J, van Merriënboer JGG, Paas F. Cognitive architecture and instructional design: 20 years later. *Educ Psychol Rev* 2019 Jun;31(2):261-292. [doi: [10.1007/s10648-019-09465-5](https://doi.org/10.1007/s10648-019-09465-5)]
94. Spahl W, Motta V, Woodcock K, Rubeis G. Gamified digital mental health interventions for young people: scoping review of ethical aspects during development and implementation. *JMIR Serious Games* 2024 Nov 28;12:e64488. [doi: [10.2196/64488](https://doi.org/10.2196/64488)] [Medline: [39607995](https://pubmed.ncbi.nlm.nih.gov/39607995/)]
95. Batten SR, Bang D, Kopell BH, et al. Dopamine and serotonin in human substantia nigra track social context and value signals during economic exchange. *Nat Hum Behav* 2024 Apr;8(4):718-728. [doi: [10.1038/s41562-024-01831-w](https://doi.org/10.1038/s41562-024-01831-w)] [Medline: [38409356](https://pubmed.ncbi.nlm.nih.gov/38409356/)]
96. Shaheen A, Ali S, Fotaris P. Assessing the efficacy of reflective game design: a design-based study in digital game-based learning. *Education Sciences* 2023;13(12):1204. [doi: [10.3390/educsci13121204](https://doi.org/10.3390/educsci13121204)]
97. Karhulahti VM. Vitality structures in "addictive" game design. *Open Res Eur* 2024;4:47. [doi: [10.12688/openreseurope.17177.2](https://doi.org/10.12688/openreseurope.17177.2)] [Medline: [38854673](https://pubmed.ncbi.nlm.nih.gov/38854673/)]
98. Saini N, Hodgins DC. Investigating gaming structural features associated with gaming disorder and proposing a revised taxonomical model: a scoping review. *J Behav Addict* 2023 Jun 29;12(2):352-374. [doi: [10.1556/2006.2023.00019](https://doi.org/10.1556/2006.2023.00019)] [Medline: [37256740](https://pubmed.ncbi.nlm.nih.gov/37256740/)]
99. Bronfenbrenner U, editor. *Making Human Beings Human: Bioecological Perspectives on Human Development*: Sage Publications Ltd; 2005. URL: <https://us.sagepub.com/en-us/nam/making-human-beings-human/book225589#features> [accessed 2026-02-24]
100. Kowal M, Conroy E, Ramsbottom N, Smithies T, Toth A, Campbell M. Gaming your mental health: a narrative review on mitigating symptoms of depression and anxiety using commercial video games. *JMIR Serious Games* 2021 Jun 16;9(2):e26575. [doi: [10.2196/26575](https://doi.org/10.2196/26575)] [Medline: [34132648](https://pubmed.ncbi.nlm.nih.gov/34132648/)]

Abbreviations

AI: artificial intelligence

IA: internet adaptability

SG: serious game

SRL: self-regulated learning

Edited by S Brini; submitted 16.Aug.2025; peer-reviewed by D Morales-Martinez, Y Asada; revised version received 30.Jan.2026; accepted 12.Feb.2026; published 09.Mar.2026.

Please cite as:

Gao S, Jia M, Wang W, Hu J, Ma S

The Mechanism and Design Principles of Serious Games in Enhancing Adolescents' Internet Adaptability

JMIR Serious Games 2026;14:e82505

URL: <https://games.jmir.org/2026/1/e82505>

doi: [10.2196/82505](https://doi.org/10.2196/82505)

© Shijie Gao, Min Jia, Weijun Wang, Jianping Hu, Shihao Ma. Originally published in JMIR Serious Games (<https://games.jmir.org>), 9.Mar.2026. 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.

Is Gamification the New Panacea for Health Behavioral Changes? Implications for the Health and Life Insurance Industry

Abbas Salami¹, PhD; Tasos Papastylianou^{1,2}, MBChB, DPhil; Marvellous Adeoye¹, MSc; John Ronayne³, PhD; Honor Bixby⁴, PhD; Robert S Stawski⁵, PhD; Bernard Fromson³, MBBS, MRes; Matt Doltis³, BSc; Osama Mahmoud^{6,7}, PhD, SFHEA; Mariachiara Di Cesare¹, PhD

¹Institute of Public Health and Wellbeing, University of Essex, Wivenhoe Park, Colchester, United Kingdom

²School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester, United Kingdom

³YuLife, London, United Kingdom

⁴School of Health and Medical Sciences, City St George's, University of London, London, United Kingdom

⁵Emma Eccles Jones College of Education and Human Services, Utah State University, Utah, UT, United States

⁶School of Mathematics, Statistics and Actuarial Science, University of Essex, Colchester, United Kingdom

⁷Department of Applied Statistics, Helwan University, Cairo, Egypt

Corresponding Author:

Tasos Papastylianou, MBChB, DPhil

Institute of Public Health and Wellbeing, University of Essex, Wivenhoe Park, Colchester, United Kingdom

Abstract

Chronic health conditions impose substantial financial and operational burdens on the public health sector and insurance providers in the United Kingdom. While gamification demonstrates the potential for enhancing health behavior, a structured analysis linking to established behavioral frameworks is missing. We provide a viewpoint on whether, as health and life insurers transition from traditional risk assessment toward proactive risk reduction strategies, gamification offers an innovative mechanism to strengthen their prevention initiatives and insurer-insured relationships. We examine how gamification aligns with key theoretical models, including the Behavior Change Wheel and Behavior Change Techniques, and how gamification elements can be mapped onto them. This enables combining multiple Behavior Change Techniques into effective interventions, which provide engaging user experiences and promote intrinsic motivation. We distinguish gamification from mere incentivization, highlighting its potential for sustained health outcomes. We also explore the ethical and practical considerations of gamification in the insurance sector. We highlight the need for a robust ethical framework that preserves an individual's ability to make free and informed decisions, while ensuring inclusivity and absence of discrimination based on personal characteristics that may affect their capacity to engage in healthy behaviors. Similarly, we highlight how privacy, transparency, and accountability need to be prioritized in the governance structure of gamification programs in the sector. Our analysis emphasizes that gamification has the potential to represent the new panacea for the insurance sector, if effective gamified interventions incorporate inclusive design principles, theoretical grounding, ethical accountability, and continuous refinement to ensure alignment with long-term public and individual health objectives. This viewpoint is the first to map gamification and behavioral change frameworks into a unified model for insurer-led health behavior interventions and encourage greater investment in gamified wellness products and the use of theory-driven behavioral science in insurance-led digital health tools.

(*JMIR Serious Games* 2026;14:e80684) doi:[10.2196/80684](https://doi.org/10.2196/80684)

KEYWORDS

gamification; behavior change interventions; physical activity; health and life insurance; mobile phone; digital health

Introduction

In the United Kingdom, the increasing prevalence of chronic health conditions is placing a growing financial and systemic burden on the public health sector and insurance providers. An estimated 15 million people in the country live with long-term conditions such as heart disease, cancer, renal disease, and diabetes, which cost the National Health Service (NHS) £73.3 billion in 2023 alone [1,2] (for comparison, at the time of writing, the conversion rate was approximately GBP £1=US

\$1.32). A significant factor contributing to this issue is physical inactivity [3]. In the United Kingdom, 32% of men and 40% of women aged above 18 years did not meet physical activity (PA) recommendations in 2022; for adults aged above 70 years, this rose to 47% and 56%, respectively, and for adolescents aged 11 to 17 years, this rose to a staggering 70% and 85%, respectively [4], despite strong evidence that regular PA reduces the risk of mortality and morbidity [5]. Even in otherwise phenotypically healthy individuals, sedentary lifestyles are more likely to be associated with early physiological changes

indicating metabolic stress, which may predate overt metabolic disease [6,7]. In 2022, physical inactivity cost the NHS an estimated £0.9 billion [8]. Beyond direct medical costs, the overall economic burden, including diminished productivity and quality of life, was estimated at £7.4 billion per year [8]. Globally, the annual cost of preventable chronic conditions associated with physical inactivity amounted to approximately US \$27 billion [4]. These trends intensify pressures on the NHS and contribute to rising insurance claims and costs, especially within an aging and chronically ill population.

Health behaviors are actions taken by individuals that affect their health and encompass a wide range of actions such as PA, smoking, or diet [9]. Therefore, behavior change is a critical component of public health initiatives aimed at improving individual and community well-being. However, multiple barriers, such as lack of motivation, limited knowledge, and insufficient support, can prevent individuals—particularly middle-aged adults—from engaging in and maintaining healthy behaviors [10].

Traditional approaches to health behaviors often rely on education and awareness campaigns. While valuable, these often fail to achieve sustained behavior change [11], as individuals often struggle to adapt their daily routines accordingly [12]. This knowledge-action gap highlights the need for innovative interventions that can effectively engage individuals and facilitate lasting change.

In this context, gamification, defined as “the use of game design elements in nongame contexts” [13], has emerged as a novel and promising approach [14]. It leverages the intrinsic enjoyment associated with games to engage users, making health-related tasks more appealing and rewarding [15]. By incorporating elements such as points, badges, leaderboards, and challenges, gamification can transform healthy activities into engaging experiences that encourage sustained participation [16]. Leveraging social interaction is a key component of gamified interventions, fostering a sense of community and competition through elements such as leaderboards, team challenges, and shared goals [15,17]. These features engage individuals’ intrinsic motivations for social connection, recognition, competition, and belonging [18,19]. Similarly, collaborative tasks or group challenges encourage participants to coordinate their efforts with others, transforming individual behaviors into socially reinforced habits [20,21].

While gamification is often paired with incentivization [22], these reflect fundamentally different strategies. Incentivization encourages behavior change primarily through extrinsic motivators, such as financial rewards [23,24]. Gamification focuses more on creating engaging, immersive experiences beyond mere incentivization by additionally using intrinsic motivators such as enjoyment, mastery, and autonomy [25]. For instance, while incentivization might reward improved fitness with monetary discounts, gamification could further motivate through level unlocking and badge earning. Similarly, gamification transcends mere incentivization by embedding social mechanisms and engaging deeper motivational drivers related to personal identity, mutual support, and social recognition [26]. Therefore, although both approaches can be

effective [22], gamification uniquely fosters sustained engagement by making the behavior change process inherently enjoyable rather than simply driven by external reward prospects [27].

The rapid advancement of computing technology has revolutionized the landscape of health behavior change, resulting in a diverse array of consumer applications aimed at monitoring and maintaining individual health [28,29]. This has given rise to the cultural phenomenon known as the “quantified self” [30-32], whereby individuals use technology, such as activity trackers or sleep monitors, to track and quantify various aspects of their daily lives. These technological advancements have provided policymakers with new avenues to implement effective behavioral interventions, catering to both individual- and population-level health initiatives [33]. Digital health technologies have provided a fertile ground for integrating gamification into behavior interventions, offering scalable solutions with real-time monitoring, feedback, and personalization to sustain user engagement [34,35]. The global gamification market is projected to grow from US \$22.01 billion in 2024 to US \$73.66 billion by 2029, with the mobile segment leading the market [14,36]. In addition, gamification has attracted interest in sectors such as health and life insurance (Textbox S1 in [Multimedia Appendix 1](#) [22,37-90]), where it holds the potential to drive positive health behaviors among policyholders while reducing risks and costs for insurers.

This paper aims to explore the role of gamification in supporting health behavior change, with a particular focus on its emerging use in the health and life insurance industry and the implications for the sector. To support this discussion and inform the design of gamified products in insurance settings, we outline the theoretical foundations of gamification and review examples of its application in health behavior change interventions (BCIs), particularly in gamified digital health applications. We also discuss potential benefits and challenges, as well as ethical and practical considerations involved in designing such interventions within this setting. While the primary audience of this work is the health insurance sector, the discussion is also relevant to other stakeholders involved in designing health behavior-oriented interventions more generally.

It should be noted that, more broadly, the scope of health apps extends beyond behavior-oriented applications; apps addressing treatment, education, and advice are also common [91-93] and have implications for the health and life insurance industry. However, these are generally unrelated to gamification and, therefore, beyond the scope of this paper.

The Evolving Role of Health and Life Insurers: From Risk Prediction to Risk Reduction Through Health Behavior Change

Traditionally, health insurance underwriting involves point-in-time evaluations using demographic, medical, behavioral, and sometimes financial data to assign individuals to risk classes, which then determine premiums [94,95]. For employer-sponsored group cover, traditional underwriting relies

on broad rating factors, such as age, occupation, and group-level claims experience. This approach spreads risk across a company's workforce, enabling employees below specified benefit thresholds to receive cover without individual medical underwriting [96]. While this model is administratively efficient and equitable at scale, the data collected during underwriting remain static until policy renewal. In addition, insurers typically refine their pricing models only periodically, drawing on population-level experience. However, in recent years, there has been a shift toward more dynamic approaches enabled by technology and artificial intelligence [97]. Insurers may leverage data from smartphone or wearable health apps to monitor risk-related behavioral and biometric data in real time [98]. Rather than simply predict claims risk, the goal is to reduce risk by promoting healthier behaviors, shifting insurers from being passive underwriters to partners in health management.

Several health insurers in the United States (eg, UnitedHealth, Humana, Cigna, and Aetna) and the United Kingdom (eg, Vitality, YuLife, Aviva, and AXA Health) now offer mobile apps and wellness programs (refer to Textbox S1 in [Multimedia Appendix 1](#) for specific examples) to encourage healthier lifestyles among policyholders. These initiatives aim to create a healthier policyholder pool and support broader organizational outcomes, such as reduced absenteeism and improved productivity. This is particularly relevant as emerging risks such as mental health represent a sizable share of group income protection claims in the United Kingdom, accounting for 21% of new claims in 2023 [99].

These insurer-led programs are often primarily reward based, relying on financial incentives or discounts to motivate engagement. As rewards also constitute a core element of gamification, it is important to clarify how such mechanisms operate within gamified health interventions, as opposed to nongamified approaches that rely solely on monetary incentives. Addressing this distinction requires an understanding of how intrinsic and extrinsic motivations interact in this context. Extrinsic motivation, including points or prizes granted upon completing specific tasks, forms one facet of gamification. In contrast, intrinsic motivation is driven by the pleasure and satisfaction derived from immersive, enjoyable game participation, such as eagerness to compete [37]. Although external rewards can catalyze short-term motivation spurts, overreliance on them can diminish intrinsic motivation [100]. Furthermore, depending solely on extrinsic motivators may hinder sustainable engagement or result in negative consequences, such as cheating or disengagement [101].

Therefore, understanding when and how gamification supports sustainable behavior change requires grounding these design elements within established theories and frameworks of behavior change.

Theoretical Foundations of Health Behavior Change and the Role of Gamification

Changing health behaviors presents a significant challenge across populations [102]. To address its complexity, several

theories, collectively referred to as *theories of behavioral change*, have been developed to provide deeper insights into the processes underlying changes in health behavior and facilitate the design and evaluation of BCIs [102]. These theories vary in scope, from models focusing on cognitive risk appraisal to those emphasizing motivation, capability, and environmental influences (Textbox S2 in [Multimedia Appendix 1](#)).

From Theory to Intervention Design

Overview

While these theories provide valuable insights into why people engage or fail to engage in health behaviors, highlighting the roles of beliefs, perceived capability, motivational quality, and situational triggers, they offer limited guidance on how to systematically design interventions that can modify these determinants in practice [103]. To bridge this gap, comprehensive frameworks such as the Behavior Change Wheel (BCW) and the Behavior Change Technique (BCT) taxonomy offer systematic guidance on translating theory into practice.

The BCW Framework

The BCW framework [87] has been developed to design interventions by linking them to a detailed analysis of the behaviors they aim to influence. Consisting of 19 distinct behavior change frameworks, the BCW effectively captures mechanisms (internal and external) potentially involved in changing behaviors. For example, internal factors may involve motivation or physical ability, while external factors can include social norms or access to resources. Three essential components—capability, opportunity, and motivation—at the center of the BCW framework (known as the capability, opportunity, motivation, and behavior [COM-B] system) are needed for change to occur (Textbox S3 in [Multimedia Appendix 1](#)). At the outer level, the BCW framework outlines 9 distinct categories of BCIs, each targeting 1 or more components to drive behavior change [104] (Table S1 in [Multimedia Appendix 1](#)). The BCW has emerged as a widely accepted framework for designing effective health behavior interventions across various settings. The UK public health guidance encourages local authorities to adopt this framework to systematically develop population-level interventions [105]. Similarly, international organizations such as the United Nations Children's Fund emphasize the importance of the BCW in ensuring evidence-based and context-sensitive programming in global health initiatives [106]. For example, the BCW has been used in adults with obesity to create tailored home-based exercise programs [107]; in the workplace to address sedentary behavior by identifying key influences on employee behavior and linking them to appropriate intervention strategies [108]; in maternal health to guide the co-design of the Healthy Gut Diet initiative, which aims to prevent gestational diabetes through changes in diet [109]; and to develop resources for enhancing PA conversations in health care settings [110].

The BCT Taxonomy

Health behavior interventions tend to use multiple interconnected components [111]. Consequently, more nuanced theoretical constructs have been developed to inform the design of effective behavior change strategies. Notably, the BCT

taxonomy version 1 [90] identifies 93 distinct “active ingredients” or techniques that can potentially be integrated into an intervention. These techniques are observable, replicable, and irreducible components designed to modify or redirect the causal processes that govern behavior [112,113]. The taxonomy serves as a comprehensive, hierarchical classification system of intervention components using 93 BCTs, categorized into 16 clusters (Table S2 in [Multimedia Appendix 1](#)). For instance, the concept of incentives is classified as a BCT under the “reward and threat” cluster. This taxonomy effectively codes interventions across various behavioral domains [114], such as PA and dietary habits [115,116].

Several of the BCW-based interventions discussed earlier provide clear examples of the BCT approach. For instance, Power et al [107] integrated 24 BCTs into their home-based exercise program for adults with overweight and obesity, including *goal setting*, *action planning*, and *reward systems* to promote sustained engagement. Ojo et al [108] identified 39 BCTs to reduce workplace sitting, targeting both motivational and contextual barriers through techniques such as *environmental restructuring*, *prompts and cues*, and *social support*. Meloncelli et al [109] incorporated 40 BCTs into the Healthy Gut Diet intervention from several BCT clusters, including *goals and planning*, *feedback and monitoring*, and *reward and threat*, to encourage healthier eating during pregnancy. These examples demonstrate how BCW-guided interventions can also systematically embed multiple BCTs to address complex health behaviors effectively.

Linking Gamification to Theories of Behavioral Change

Earlier, we introduced the 9 intervention functions of the BCW framework and discussed how they address deficits in the components of the COM-B system. We also argued that the techniques outlined in the BCT taxonomy are linked to the BCW’s intervention functions. In essence, intervention functions can consist of 1 or more BCTs, and a single BCT can serve multiple intervention functions [87]. As game elements can be straightforwardly mapped onto corresponding BCTs, gamification can serve as a driver for effective behavioral change.

For instance, points and badges serve as *nonspecific incentives* and *rewards*, respectively, offering a sense of achievement, while points can also become *material incentives* if redeemable [117]. Leaderboards encourage *social comparison* and *competition* [19], motivating users by letting them track their progress against others. Levels and challenges reflect BCTs related to *goal setting* and *graded tasks*. Each level typically sets a specific, achievable goal (reaching a daily step count or avoiding smoking for a day), supporting goal setting techniques. As users advance, tasks gradually become more difficult, exemplifying the use of *graded tasks* to encourage behavior

change in manageable steps. Feedback and self-monitoring tools are also central to gamification, providing users with immediate, engaging updates such as progress bars, activity logs, or performance summaries [15,17]. These features align with BCTs such as *self-monitoring of behavior* and *feedback on outcomes* and often incorporate rewards such as badges for meeting goals.

One significant advantage of using gamification in BCIs is that it allows for a wider variety of BCTs to be incorporated into a single intervention function, effectively acting as a toolkit to enrich the behavioral intervention applied [118]. One example is incentivization, which aims to influence motivation and behavior through reward mechanisms. Without gamification, mere incentivization-based interventions often rely on a narrow subset of BCTs within the “reward and threat” cluster, typically *material incentive (behavior)* or *material reward (behavior)*, such as vouchers or redeemable points. However, gamification facilitates the inclusion of a broader spectrum of BCTs within the same cluster, such as *social reward*, through badges or avatar upgrades, expanding the ways rewards can be delivered, moving beyond mere monetary incentives to include symbolic, social, and intrinsic motivators [119].

Empirical evidence supports this distinction. A review of gamified health apps found that the median number of BCTs implemented per app was 14, with some apps using up to 22 techniques [120], while nongamified digital interventions often include only a few techniques [121].

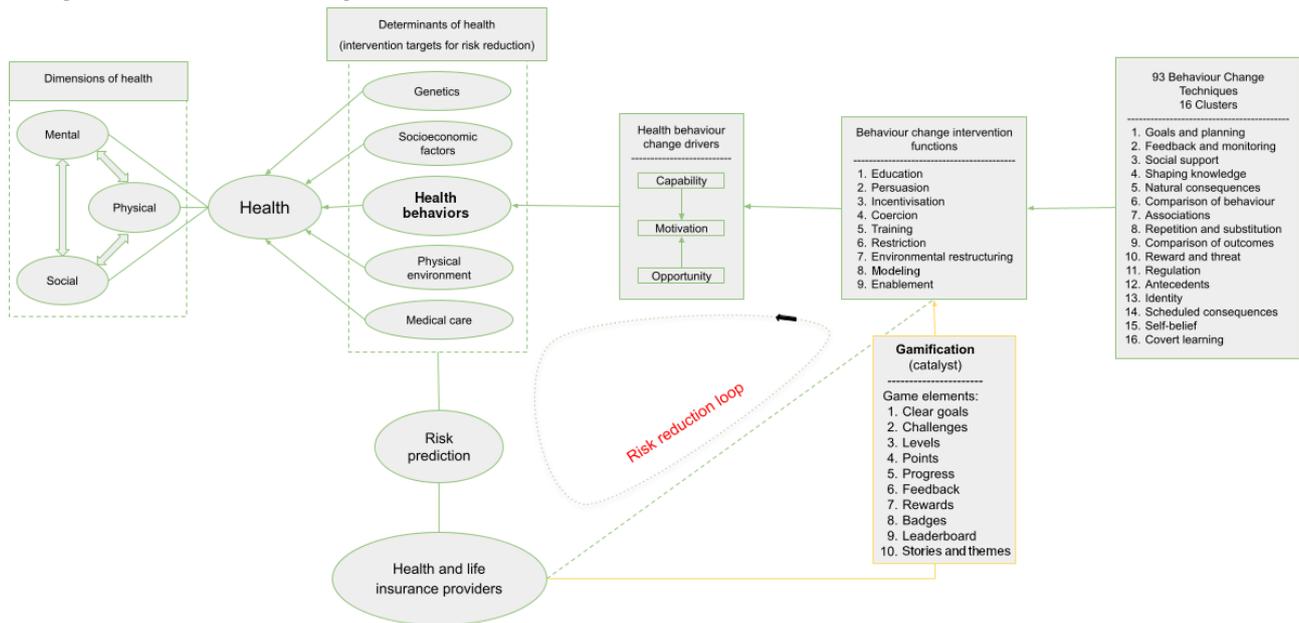
Designing and Implementing Gamified Interventions in Insurance Settings: Opportunities and Challenges

The existing evidence suggests that gamification can positively impact health and well-being [122] by promoting healthy behaviors [14,123] and sustaining user engagement in personal informatics applications designed for the quantified-self movement [124]. Gamification leverages people’s desire for competition and social interaction, serving as a powerful catalyst for encouraging lifestyle changes [125]. These effects form the basis for understanding how gamified interventions may generate value for both policyholders and insurers.

Value Creation for Insurers and Policyholders

The true power of gamification lies in the simultaneous use of multiple game elements. [Figure 1](#) illustrates how gamification can enable the integration of BCTs into risk reduction initiatives and improve health outcomes (and thus insurer costs). A total of 93 BCTs comprise the foundational components of 9 broader intervention functions defined in the BCW framework. These functions influence the key drivers of behavior (COM-B system), which in turn shape individual health behaviors. These behaviors can be integrated into existing actuarial practices for more comprehensive risk assessment.

Figure 1. Overview of theories of behavioral change in health, their intersection with gamification, and the role of gamified interventions in supporting insurance providers' risk reduction strategies.



Additionally, as policyholders adopt healthier lifestyles, insurers experience lower health care claims, allowing them to offer reduced premiums and financial incentives. Similarly, engaged policyholders benefit from better health outcomes and lower insurance costs. While peer-reviewed studies in this area are lacking, a 5-year impact study of approximately 40,000 insured individuals reported that engaged members averaged US \$462 less in annual medical claims, corresponding to a 4% reduction in claims costs with an estimated return on investment of approximately 180% [126].

Early evidence indicates that incentive-driven interventions can significantly boost engagement. Roughly two-thirds of consumers in a 2019 study said they would be willing to enroll in insurance wellness programs with wearables [127], with rates increasing further when rewards such as premium reductions were offered. While incentives can drive initial participation, gamified environments can increase user engagement further by helping satisfy basic psychological needs [128], such as facilitating a sense of community and support, and providing motivation and, thus, creating positive associations with the insurance provider. Such sustained engagement (often higher than industry standards [129,130]) can, in turn, help reduce customer lapse rates for insurance providers [131].

Engagement, Data, and Personalization as Strategic Assets

High engagement through gamification provides insurers with strategic advantages, particularly richer data collection. Real-time behavioral data from gamified interactions can improve the accuracy of risk assessments, highlight emerging health risks before they manifest in claims, and enable personalized services, including dynamic premium pricing and tailored product offerings. Employers benefit from additional insights into the health of their employee population, enabling targeted wellness initiatives. Brokers can use behavioral profiles to identify optimal supplementary health products, enhancing cross-selling effectiveness. Reinsurers gain increased portfolio

confidence, improving risk management and solvency assessments.

Design and Implementation Challenges for Insurers

Designing effective products delivering gamified BCIs is a challenge. Gamification should strike a balance between extrinsic and intrinsic motivators, with intrinsic factors informed by theories of behavioral change, to ensure sustainable engagement and meaningful health improvements [132]. For insurers in the employer-sponsored group-cover market, any health gains arising from such interventions must be credibly reflected in underwriting, according to 3 core principles: risk pooling, actuarial fairness, and regulatory compliance, while also guarding against “2-tier market” selection effects [133], and implementing strict data-governance measures to build and maintain user trust [134]. Successfully implementing gamification strategies requires ongoing investment in advanced analytics, iterative product refinement based on user feedback, and collaborative integration across the broader insurance ecosystem to maximize collective benefits [135].

Ethical and Practical Considerations of Gamification in Insurance Settings

Gamification in health insurance raises several ethical challenges. If engagement and behavioral data are analyzed to directly inform insurance policies, this can compromise people's ability to engage with such risk-reducing interventions freely [38]. Interventions require design considerations and adequate monitoring to mitigate potential adverse effects (eg, on mental health or social well-being) [38,39]. This may be further problematic in employer-sponsored insurance, compromising autonomy and the ability to provide informed consent, particularly with regard to how personal behavioral data are handled [40]. In the absence of appropriate data agreements, function creep (ie, the use of data for purposes other than the original intent) could result in harm, for example, interpreting

periods of illness as noncompliance or relying on out-of-scope data to counter insurance claims. Finally, inclusivity may be a concern, as standard game elements may not be equally accessible to all, resulting in a form of digital divide [136]. This could lead to adverse effects on insurance premiums and health outcomes in the affected populations, for example, through perceptions of nonadherence or indirect discrimination regarding rewards [41]. Therefore, ethical evaluation of gamified interventions should also address how data could be misused, rather than focus solely on privacy considerations. For a more detailed discussion on ethical and practical considerations, including practical and data security challenges, refer to Textbox S4 in [Multimedia Appendix 1](#).

Empirical Evidence From Systematic Reviews and Meta-Analyses

Overview

While there is a sound theoretical basis for applying gamification in BCIs, empirical evidence addressing improvements in such initiatives should also be considered. However, studies evaluating gamification specifically within insurance settings are scarce. Therefore, to inform this viewpoint, we focused on evidence from broader digital health BCIs, which has been assessed in numerous health behavior change programs [15]. To inform this viewpoint, we draw on findings from published systematic reviews and meta-analyses examining the impact of gamification in BCIs, with a primary focus on PA (due to its prominence in the literature and relevance to population-level digital health and prevention strategies). [Table 1](#) summarizes the key findings.

Table . Summary of systematic reviews and meta-analyses examining gamified behavior change interventions.

Study	Review type	Number of studies (participants)	Target population	Main findings
Mazeas et al [42] (2022)	Systematic review and meta-analysis	16 RCTs ^a (n=2407)	General population (9 - 73 y)	Small-to-moderate positive effect on PA ^b (Hedges' $g=0.42$; 95% CI 0.14 - 0.69), step count increase of +1609 steps per day (95% CI 372-2847), and reduced long-term effect at follow-up (Hedges' $g=0.15$)
Nishi et al [75] (2024)	Systematic review and meta-analysis	36 RCTs (n>10,000)	General adult populations	+489 steps per day (95% CI 64-914), BMI -0.28 kg/m ² , weight -0.70 kg, waist circumference -1.16 cm, and body fat -1.92%
Xu et al [76] (2022)	Systematic review	50 studies (n=9977)	Mixed populations (predominantly younger adults)	Mixed evidence and modest improvements in PA and engagement
Chen et al [77] (2025)	Systematic review	8 trials (n=1454)	Older adults (≥ 60 y)	Generally positive effects on steps and MVPA ^c , stronger effects in theory-driven and hybrid interventions, and effect sizes ranged from small to moderate
Wang et al [86] (2025)	Systematic review and meta-analysis	16 RCTs (n=7472)	Children and adolescents (6 - 18 y)	Small increase in MVPA (standardized mean difference=0.15), BMI reduction (standardized mean difference=0.11), and no effect on steps or sedentary behavior
Alzghoul [88] (2024)	Systematic review and meta-analysis	30 RCTs (n=11,558)	Mixed populations	Significant pooled effect on behavior change (odds ratio 1.27, 95% CI 1.20-1.33) and no evidence of publication bias
Yu et al [89] (2025)	Systematic review and meta-analysis	6 RCTs (n=1109)	Adults with cardiovascular disease	Short-term PA effect (Hedges' $g=0.32$), maintenance effect at follow-up (Hedges' $g=0.20$), and +697 steps per day

^aRCT: randomized controlled trial.

^bPA: physical activity.

^cMVPA: moderate-to-vigorous physical activity.

Overall, the studies summarized in [Table 1](#) suggest that gamified interventions were generally associated with small-to-moderate PA improvements, mostly reflecting daily step counts, but also, to a lesser extent, “moderate-to-vigorous PA,” across diverse populations (eg, young, old, and chronically ill). However, individual studies vary considerably in terms of intervention design, outcome measures, and combinations of game elements, making comparisons across the spectrum somewhat challenging. This is further compounded by common methodological constraints, such as short interventions, limited follow-up, and challenges in isolating individual game component contributions. Brief summaries of individual reviews and additional references providing relevant context [78-85] are provided in [Textbox S5](#) in [Multimedia Appendix 1](#).

Benefits and Limitations Identified in the Literature

Collectively, evidence from systematic reviews and meta-analyses indicates that gamified interventions are associated with positive effects on PA. Meta-analytic findings suggest small-to-moderate improvements in overall PA (eg, step count, moderate-to-vigorous PA, and active minutes), with statistically significant increases in daily step counts reported across several reviews. Beyond PA, additional benefits have been observed in other clinical and behavioral domains [137], including modest improvements in cardiometabolic risk markers [138], medication adherence [139], dietary behaviors [140,141], and short-term smoking abstinence [142]. Taken together, these findings suggest that gamification can successfully initiate behavior change and enhance engagement across diverse populations and health contexts.

However, a substantial proportion of studies report mixed or neutral outcomes. One key limitation is the dominance of extrinsic motivators, most notably financial incentives, which characterize many interventions labeled as gamified. While rewards are a legitimate game-design element and may take nonmonetary forms (eg, virtual items), the frequent coupling of gamification with direct financial incentives blurs the distinction between gamification and mere incentivization. This overlap makes it difficult to disentangle behavioral effects attributable to game mechanics from those driven primarily by monetary rewards, particularly when both are implemented concurrently.

Methodological limitations further constrain interpretation. Considerable heterogeneity in intervention design, target populations, and outcome measures, combined with the widespread use of multiple game elements simultaneously, limits the ability to isolate the effects of individual components and contributes to highly variable effect sizes across studies. Many trials are underpowered, rely on short intervention durations, or lack appropriate nongamified comparison groups, reducing confidence in causal attribution. Sustained effectiveness also remains a central challenge. Although short-term improvements are frequently observed, engagement and behavioral effects often diminish over time, with follow-up analyses consistently showing smaller and more volatile maintenance effects. Effectiveness also depends on alignment between game design and user characteristics. Features that motivate novice users may disengage more experienced users, while certain elements may be poorly suited to sensitive contexts such as mental health or mindfulness. Usability problems (such as complex interfaces or high interaction burden), declining engagement over time, privacy concerns, and limited personalization further constrain the scalability and long-term sustainability of gamified health interventions [15].

Discussion

This paper explored the role of gamification in enhancing health BCIs, with particular emphasis on its application within the health and life insurance sector. We have highlighted the potential of gamification as a powerful tool for increasing engagement and promoting sustained behavior change. We showed that gamification enables the deployment and integration of multiple BCTs, targeting essential drivers of health behavior change. By mapping game elements to specific BCTs, designers can ensure that fun and engaging interactions also serve meaningful behavioral purposes. However, there are still some limitations to address. Gamification studies often report short-term outcomes but fail to address the sustainability of behavior change over time. As the novelty of gamified features wears off, maintaining long-term engagement becomes a challenge [14]. Moreover, many interventions lack appropriate control groups, making it difficult to isolate the specific

contribution of gamification to the observed outcomes. Future research must prioritize long-term evaluations to investigate sustained behavior change and examine to what extent personalization can enhance gamification's impact [143], as one-size-fits-all approaches may not be adequate in health contexts [144,145].

Practical opportunities for gamification extend beyond public health interventions and hold value for the health and life insurance sector. The increasing prevalence of chronic conditions, primarily driven by unhealthy behaviors such as sedentary lifestyles [146], places a growing financial burden on health care systems and insurance providers [147]. Gamification enables insurers to encourage healthier lifestyles among policyholders, potentially improving health outcomes and reducing claim costs. Moreover, gamification can enhance health literacy among younger adults [148], supporting better decision-making regarding insurance coverage and preventive behaviors.

One main finding emerging from this review is the lack of studies specific to the insurance sector, despite the emerging use of gamification in the sector. For insurers seeking to meet the varied needs and preferences of different demographics, leveraging a diverse range of game mechanics and BCTs, grounded in sound evidence, is essential. Programs must be inclusive, accessible, and continuously refined based on participant feedback and health outcomes. Establishing clear links between game elements, BCTs, and behavior change determinants can strengthen program evaluation and ensure interventions are underpinned by established behavioral science. In doing so, insurers encourage healthier behavior among policyholders and also create more sustainable, financially efficient insurance models.

Conclusions

As digital health continues to advance, gamification emerges as a promising approach for enhancing health behavior interventions. When applied thoughtfully and evaluated rigorously, gamified strategies can significantly contribute to healthier populations, more proactive health care, and innovative service models in industries such as health and life insurance, having the potential to be a new panacea for the sector and for population health more generally. However, while gamification represents a powerful catalyst for advancing health behavior change, its application must be guided by careful, theory-based design and ongoing evaluation [149]. Ethical, practical, and regulatory considerations must also be considered to ensure that gamification is used responsibly. Looking ahead, the challenge lies not just in gamifying health, but in doing so in ways that are equitable, grounded in solid theoretical principles and evidence from the literature, and aligned with long-term public and individual well-being.

Acknowledgments

The authors declare limited use of generative artificial intelligence (GenAI) in the research and writing process. In accordance with the Generative AI Delegation Taxonomy (2025), selected tasks were delegated to GenAI tools under full human supervision,

including the evaluation of research novelty and the identification of high-level trends and potential gaps in the literature. ChatGPT (versions 4.5 and 5.2) and NotebookLM were used exclusively for narrowly defined exploratory purposes, following initial author-led screening and curation of relevant systematic reviews and meta-analyses. These tools were used solely to support ideation and organization. All outputs were critically reviewed, verified against original sources, and curated by the authors. No artificial intelligence (AI)-generated references or factual claims were included without independent verification. GenAI was not used for study conceptualization, data analysis, interpretation of findings, or substantive authorship of the manuscript text, and no AI-generated text appears verbatim in the manuscript. Responsibility for the accuracy, integrity, and conclusions of the manuscript rests entirely with the authors. GenAI tools are not listed as authors and bear no responsibility for the final work. This declaration was submitted by AS.

Funding

This work was supported by a Knowledge Transfer Partnership (KTP) grant from Innovate UK awarded to the University of Essex and YuLife Ltd (grant 10059363; KTP reference 13536), which facilitates collaboration between industry and academic institutions. Three authors (JR, BF, and MD) are employed by the industry partner, YuLife Ltd. The manuscript was drafted entirely by the academic team; JR, BF, and MD contributed to content review and verification and provided supervisory and administrative support within the broader project. The academic authors declare no financial interests related to YuLife Ltd beyond the scope of the KTP grant.

Authors' Contributions

Conceptualization: AS, TP, HB, RS, OM, MDC

Funding acquisition: TP, JR, HB, RS, MD, MDC

Investigation: AS, TP, JR, MDC

Project administration: TP, JR, MD, MDC

Supervision: TP, JR, BF, OM, MDC

Validation: JR, MDC

Writing – original draft: AS, TP, MA

Writing – review & editing: AS, TP, MA, JR, HB, RS, BF, MD, OM, MDC

Conflicts of Interest

None declared.

Multimedia Appendix 1

Tables and textboxes providing an overview of gamification in the UK insurance sector, core psychological theories of behavior change, ethical and practical considerations, evidence-based summaries from systematic reviews, and standardized taxonomies for behavior change interventions and techniques.

[[DOCX File, 33 KB - games_v14i1e80684_app1.docx](#)]

References

1. Prevention before cure: securing the long-term sustainability of the NHS. British Medical Association. 2017. URL: <https://www.bma.org.uk/media/2107/bma-securing-the-long-term-sustainability-of-the-nhs.pdf> [accessed 2026-03-28]
2. Healthcare expenditure, UK Health Accounts: 2023 and 2024. : Office for National Statistics; 2025 URL: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/bulletins/ukhealthaccounts/2023and2024> [accessed 2026-03-28]
3. Strain T, Flaxman S, Guthold R, et al. National, regional, and global trends in insufficient physical activity among adults from 2000 to 2022: a pooled analysis of 507 population-based surveys with 5·7 million participants. *Lancet Glob Health* 2024 Aug;12(8):e1232-e1243. [doi: [10.1016/S2214-109X\(24\)00150-5](https://doi.org/10.1016/S2214-109X(24)00150-5)] [Medline: [38942042](https://pubmed.ncbi.nlm.nih.gov/38942042/)]
4. Global status report on physical activity 2022. : World Health Organization; 2022 URL: <https://www.who.int/publications/i/item/9789240059153> [accessed 2026-03-28]
5. Warburton DE, Bredin SS. Health benefits of physical activity: a systematic review of current systematic reviews. *Curr Opin Cardiol* 2017 Sep;32(5):541-556. [doi: [10.1097/HCO.0000000000000437](https://doi.org/10.1097/HCO.0000000000000437)] [Medline: [28708630](https://pubmed.ncbi.nlm.nih.gov/28708630/)]
6. San-Millan I, Martinez JL, Sparagna GC, D' Alessandro A, Stefanoni D, Nemkov T, et al. Metabolic and cellular differences between sedentary and active individuals at rest and during exercise. *bioRxiv*. Preprint posted online on Aug 19, 2024. [doi: [10.1101/2024.08.19.608601](https://doi.org/10.1101/2024.08.19.608601)]
7. Hernández-López OA, Murillo-Ortíz B, Luna-Marco C, et al. Cardiorespiratory fitness as a key predictor of metabolic, inflammatory, and oxidative stress biomarkers in adults with different physical activity levels. *Free Radic Biol Med* 2025 Dec 1;240:735-744. [doi: [10.1016/j.freeradbiomed.2025.08.058](https://doi.org/10.1016/j.freeradbiomed.2025.08.058)] [Medline: [40889726](https://pubmed.ncbi.nlm.nih.gov/40889726/)]

8. Physical activity: applying All Our Health. Office for Health Improvement & Disparities, United Kingdom Government. 2022. URL: <https://www.gov.uk/government/publications/physical-activity-applying-all-our-health/physical-activity-applying-all-our-health> [accessed 2026-03-28]
9. Short SE, Mollborn S. Social determinants and health behaviors: conceptual frames and empirical advances. *Curr Opin Psychol* 2015 Oct;5:78-84. [doi: [10.1016/j.copsyc.2015.05.002](https://doi.org/10.1016/j.copsyc.2015.05.002)] [Medline: [26213711](https://pubmed.ncbi.nlm.nih.gov/26213711/)]
10. Kelly S, Martin S, Kuhn I, Cowan A, Brayne C, Lafortune L. Barriers and facilitators to the uptake and maintenance of healthy behaviours by people at mid-life: a rapid systematic review. *PLoS One* 2016;11(1):e0145074. [doi: [10.1371/journal.pone.0145074](https://doi.org/10.1371/journal.pone.0145074)] [Medline: [26815199](https://pubmed.ncbi.nlm.nih.gov/26815199/)]
11. Marteau TM, Hollands GJ, Fletcher PC. Changing human behavior to prevent disease: the importance of targeting automatic processes. *Science* 2012 Sep 21;337(6101):1492-1495. [doi: [10.1126/science.1226918](https://doi.org/10.1126/science.1226918)] [Medline: [22997327](https://pubmed.ncbi.nlm.nih.gov/22997327/)]
12. Schlaff AL. Behavior change in America: public health, medicine, and individual counseling. *Virtual Mentor* 2013 Apr 1;15(4):353-361. [doi: [10.1001/virtualmentor.2013.15.4.mhst1-1304](https://doi.org/10.1001/virtualmentor.2013.15.4.mhst1-1304)] [Medline: [23566786](https://pubmed.ncbi.nlm.nih.gov/23566786/)]
13. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining “gamification”. Presented at: MindTrek '11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; Sep 28-30, 2011. [doi: [10.1145/2181037.2181040](https://doi.org/10.1145/2181037.2181040)]
14. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games* 2013 Dec 12;1(1):e3. [doi: [10.2196/games.3139](https://doi.org/10.2196/games.3139)] [Medline: [25658754](https://pubmed.ncbi.nlm.nih.gov/25658754/)]
15. Johnson D, Deterding S, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. *Internet Interv* 2016 Nov;6:89-106. [doi: [10.1016/j.invent.2016.10.002](https://doi.org/10.1016/j.invent.2016.10.002)] [Medline: [30135818](https://pubmed.ncbi.nlm.nih.gov/30135818/)]
16. Uechi H, Tan N, Honda Y. Effects of gamification-based intervention for promoting health behaviors. *J Phys Fitness Sports Med* 2018;7(3):185-192. [doi: [10.7600/jpfs.7.185](https://doi.org/10.7600/jpfs.7.185)]
17. Lister C, West JH, Cannon B, Sax T, Brodegard D. Just a fad? Gamification in health and fitness apps. *JMIR Serious Games* 2014 Aug 4;2(2):e9. [doi: [10.2196/games.3413](https://doi.org/10.2196/games.3413)] [Medline: [25654660](https://pubmed.ncbi.nlm.nih.gov/25654660/)]
18. Hamari J, Koivisto J. Social motivations to use gamification: an empirical study of gamifying exercise. Presented at: ECIS 2013: Proceedings of the 21st European Conference on Information Systems; Jun 5-8, 2013 URL: https://www.researchgate.net/publication/236269293_Social_motivations_to_use_gamification_An_empirical_study_of_gamifying_exercise [accessed 2026-03-28]
19. Hamari J, Koivisto J. “Working out for likes”: an empirical study on social influence in exercise gamification. *Comput Human Behav* 2015 Sep;50:333-347. [doi: [10.1016/j.chb.2015.04.018](https://doi.org/10.1016/j.chb.2015.04.018)]
20. Wing RR, Jeffery RW. Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. *J Consult Clin Psychol* 1999;67(1):132-138. [doi: [10.1037/0022-006X.67.1.132](https://doi.org/10.1037/0022-006X.67.1.132)] [Medline: [10028217](https://pubmed.ncbi.nlm.nih.gov/10028217/)]
21. Graupensperger S, Gottschall JS, Benson AJ, Eys M, Hastings B, Evans MB. Perceptions of groupness during fitness classes positively predict recalled perceptions of exertion, enjoyment, and affective valence: an intensive longitudinal investigation. *Sport Exerc Perform Psychol* 2019 Aug;8(3):290-304. [doi: [10.1037/spy0000157](https://doi.org/10.1037/spy0000157)] [Medline: [31548915](https://pubmed.ncbi.nlm.nih.gov/31548915/)]
22. Fanaroff AC, Patel MS, Chokshi N, et al. Effect of gamification, financial incentives, or both to increase physical activity among patients at high risk of cardiovascular events: the BE ACTIVE randomized controlled trial. *Circulation* 2024 May 21;149(21):1639-1649. [doi: [10.1161/CIRCULATIONAHA.124.069531](https://doi.org/10.1161/CIRCULATIONAHA.124.069531)] [Medline: [38583084](https://pubmed.ncbi.nlm.nih.gov/38583084/)]
23. Giles EL, Robalino S, McColl E, Sniehotta FF, Adams J. The effectiveness of financial incentives for health behaviour change: systematic review and meta-analysis. *PLoS One* 2014;9(3):e90347. [doi: [10.1371/journal.pone.0090347](https://doi.org/10.1371/journal.pone.0090347)] [Medline: [24618584](https://pubmed.ncbi.nlm.nih.gov/24618584/)]
24. Promberger M, Marteau TM. When do financial incentives reduce intrinsic motivation? Comparing behaviors studied in psychological and economic literatures. *Health Psychol* 2013 Sep;32(9):950-957. [doi: [10.1037/a0032727](https://doi.org/10.1037/a0032727)] [Medline: [24001245](https://pubmed.ncbi.nlm.nih.gov/24001245/)]
25. Dahlstrøm C. Impacts of gamification on intrinsic motivation [Master’s thesis]. : Norwegian University of Science and Technology; 2017 URL: https://www.ntnu.edu/documents/139799/1279149990/04%2BArticle%2BFinal_camildah_fors%25C3%25B8k_2017-12-06-13-53-55_TPD4505.Camilla.Dahlstr%25C3%25B8m.pdf [accessed 2026-03-28]
26. Sardi L, Idri A, Fernández-Alemán JL. A systematic review of gamification in e-Health. *J Biomed Inform* 2017 Jul;71:31-48. [doi: [10.1016/j.jbi.2017.05.011](https://doi.org/10.1016/j.jbi.2017.05.011)] [Medline: [28536062](https://pubmed.ncbi.nlm.nih.gov/28536062/)]
27. Bardach L, Murayama K. The role of rewards in motivation—beyond dichotomies. *Learn Instr* 2025 Apr;96:102056. [doi: [10.1016/j.learninstruc.2024.102056](https://doi.org/10.1016/j.learninstruc.2024.102056)]
28. Wongvibulsin S, Martin SS, Saria S, Zeger SL, Murphy SA. An individualized, data-driven digital approach for precision behavior change. *Am J Lifestyle Med* 2020;14(3):289-293. [doi: [10.1177/1559827619843489](https://doi.org/10.1177/1559827619843489)] [Medline: [32477031](https://pubmed.ncbi.nlm.nih.gov/32477031/)]
29. Glanz K, Rimer BK, Viswanath K. Health Behavior and Health Education: Theory, Research, and Practice: John Wiley & Sons; 2008.
30. Zhang X, Gao P, Snyder MP. The exposome in the era of the quantified self. *Annu Rev Biomed Data Sci* 2021 Jul 20;4:255-277. [doi: [10.1146/annurev-biodatasci-012721-122807](https://doi.org/10.1146/annurev-biodatasci-012721-122807)] [Medline: [34465170](https://pubmed.ncbi.nlm.nih.gov/34465170/)]
31. Swan M. The quantified self: fundamental disruption in big data science and biological discovery. *Big Data* 2013 Jun;1(2):85-99. [doi: [10.1089/big.2012.0002](https://doi.org/10.1089/big.2012.0002)] [Medline: [27442063](https://pubmed.ncbi.nlm.nih.gov/27442063/)]

32. Ossenbrink L, Haase T, Timpel P, et al. Effectiveness of digital health interventions containing game components for the self-management of type 2 diabetes: systematic review. *JMIR Serious Games* 2023 Jun 1;11:e44132. [doi: [10.2196/44132](https://doi.org/10.2196/44132)] [Medline: [37261900](https://pubmed.ncbi.nlm.nih.gov/37261900/)]
33. Chimatapu SN, Mittelman SD, Habib M, Osuna-Garcia A, Vidmar AP. Wearable devices beyond activity trackers in youth with obesity: summary of options. *Child Obes* 2024 Apr;20(3):208-218. [doi: [10.1089/chi.2023.0005](https://doi.org/10.1089/chi.2023.0005)] [Medline: [37023409](https://pubmed.ncbi.nlm.nih.gov/37023409/)]
34. Murray E, Hekler EB, Andersson G, et al. Evaluating digital health interventions: key questions and approaches. *Am J Prev Med* 2016 Nov;51(5):843-851. [doi: [10.1016/j.amepre.2016.06.008](https://doi.org/10.1016/j.amepre.2016.06.008)] [Medline: [27745684](https://pubmed.ncbi.nlm.nih.gov/27745684/)]
35. Son YS, Kwon KH. Utilization of smart devices and the evolution of customized healthcare services focusing on big data: a systematic review. *Mhealth* 2024;10:7. [doi: [10.21037/mhealth-23-24](https://doi.org/10.21037/mhealth-23-24)] [Medline: [38323151](https://pubmed.ncbi.nlm.nih.gov/38323151/)]
36. Gamification market size & share analysis - growth trends and forecast (2026 - 2031). Mordor Intelligence. URL: <https://www.mordorintelligence.com/industry-reports/gamification-market> [accessed 2026-03-28]
37. Ryan RM, Deci EL. Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp Educ Psychol* 2000 Jan;25(1):54-67. [doi: [10.1006/ceps.1999.1020](https://doi.org/10.1006/ceps.1999.1020)] [Medline: [10620381](https://pubmed.ncbi.nlm.nih.gov/10620381/)]
38. Kim TW, Werbach K. More than just a game: ethical issues in gamification. *Ethics Inf Technol* 2016 Jun;18(2):157-173. [doi: [10.1007/s10676-016-9401-5](https://doi.org/10.1007/s10676-016-9401-5)]
39. Yang H, Li D. Understanding the dark side of gamification health management: a stress perspective. *Inf Process Manag* 2021 Sep;58(5):102649. [doi: [10.1016/j.ipm.2021.102649](https://doi.org/10.1016/j.ipm.2021.102649)]
40. Durieux BN, DeCamp M, Lindvall C. 21st Century Cures Act: ethical recommendations for new patient-facing products. *J Am Med Inform Assoc* 2022 Sep 12;29(10):1818-1822. [doi: [10.1093/jamia/ocac112](https://doi.org/10.1093/jamia/ocac112)] [Medline: [35876830](https://pubmed.ncbi.nlm.nih.gov/35876830/)]
41. Arora C, Razavian M. Ethics of gamification in health and fitness-tracking. *Int J Environ Res Public Health* 2021 Oct 21;18(21):11052. [doi: [10.3390/ijerph182111052](https://doi.org/10.3390/ijerph182111052)] [Medline: [34769570](https://pubmed.ncbi.nlm.nih.gov/34769570/)]
42. Mazeas A, Duclos M, Pereira B, Chalabaev A. Evaluating the effectiveness of gamification on physical activity: systematic review and meta-analysis of randomized controlled trials. *J Med Internet Res* 2022 Jan 4;24(1):e26779. [doi: [10.2196/26779](https://doi.org/10.2196/26779)] [Medline: [34982715](https://pubmed.ncbi.nlm.nih.gov/34982715/)]
43. Rewards partners. Vitality. URL: <https://www.vitality.co.uk/rewards/> [accessed 2025-06-17]
44. Vitality to transform client engagement with evolution of vitality programme with launch of pick and play. Vitality. URL: <https://www.vitality.co.uk/media/evolution-of-vitality-programme-with-launch-of-pick-and-play/> [accessed 2025-06-17]
45. From steps to science: how gamification is transforming health and life insurance. Yu Life. 2025. URL: <https://yulife.com/blog/how-gamification-transforms-group-health-life-insurance/> [accessed 2026-06-17]
46. YuLife partners with MetLife UK to launch pre-early intervention group income protection solution. Yu Life. 2024. URL: <https://yulife.com/news-and-events/metlife-yulife-launch-group-income-protection/> [accessed 2025-06-17]
47. YuLife launches group health insurance provided by Bupa: here's how it works. Yu Life. 2023. URL: <https://yulife.com/blog/how-yulife-business-health-insurance-works/> [accessed 2025-06-17]
48. Does gamification still work, or is it a trend from the past? Yu Life. 2025. URL: <https://yulife.com/blog/does-gamification-still-work/> [accessed 2025-6-17]
49. Business health insurance. Aviva. URL: <https://www.aviva.co.uk/business/workplace-wellbeing/wellbeing/aviva-wellbeing/> [accessed 2025-06-17]
50. Find out everything your AXA health membership offers you. AXA Health. URL: <https://www.axahealth.co.uk/members/member-benefits-explained/> [accessed 2025-06-17]
51. Alyafei A, Easton-Carr R. The health belief model of behavior change. In: *StatPearls*: StatPearls Publishing; 2024.
52. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. *Health Educ Q* 1988;15(2):175-183. [doi: [10.1177/109019818801500203](https://doi.org/10.1177/109019818801500203)] [Medline: [3378902](https://pubmed.ncbi.nlm.nih.gov/3378902/)]
53. Holloway A, Watson HE. Role of self-efficacy and behaviour change. *Int J Nurs Pract* 2002 Apr;8(2):106-115. [doi: [10.1046/j.1440-172x.2002.00352.x](https://doi.org/10.1046/j.1440-172x.2002.00352.x)] [Medline: [11993577](https://pubmed.ncbi.nlm.nih.gov/11993577/)]
54. Schwarzer R, Warner LM. Perceived self-efficacy and its relationship to resilience. In: Prince-Embury S, Saklofske DH, editors. *Resilience in Children, Adolescents, and Adults: Translating Research Into Practice*: Springer Science + Business Media; 2013. [doi: [10.1007/978-1-4614-4939-3_10](https://doi.org/10.1007/978-1-4614-4939-3_10)]
55. 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](https://doi.org/10.1007/s12160-014-9593-z)] [Medline: [24648017](https://pubmed.ncbi.nlm.nih.gov/24648017/)]
56. Prestwich A, Kellar I, Parker R, et al. How can self-efficacy be increased? Meta-analysis of dietary interventions. *Health Psychol Rev* 2014;8(3):270-285. [doi: [10.1080/17437199.2013.813729](https://doi.org/10.1080/17437199.2013.813729)] [Medline: [25053214](https://pubmed.ncbi.nlm.nih.gov/25053214/)]
57. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act* 2013 Mar 3;10:29. [doi: [10.1186/1479-5868-10-29](https://doi.org/10.1186/1479-5868-10-29)] [Medline: [23452345](https://pubmed.ncbi.nlm.nih.gov/23452345/)]
58. Newby K, Teah G, Cooke R, et al. Do automated digital health behaviour change interventions have a positive effect on self-efficacy? A systematic review and meta-analysis. *Health Psychol Rev* 2021 Mar;15(1):140-158. [doi: [10.1080/17437199.2019.1705873](https://doi.org/10.1080/17437199.2019.1705873)] [Medline: [31847702](https://pubmed.ncbi.nlm.nih.gov/31847702/)]

59. Ashford S, Edmunds J, French DP. What is the best way to change self-efficacy to promote lifestyle and recreational physical activity? A systematic review with meta-analysis. *Br J Health Psychol* 2010 May;15(Pt 2):265-288. [doi: [10.1348/135910709X461752](https://doi.org/10.1348/135910709X461752)] [Medline: [19586583](#)]
60. 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](https://doi.org/10.1037//0003-066x.55.1.68)] [Medline: [11392867](#)]
61. Ng JY, Ntoumanis N, Thøgersen-Ntoumani C, et al. Self-determination theory applied to health contexts: a meta-analysis. *Perspect Psychol Sci* 2012 Jul;7(4):325-340. [doi: [10.1177/1745691612447309](https://doi.org/10.1177/1745691612447309)] [Medline: [26168470](#)]
62. Fogg BJ. A behavior model for persuasive design. Presented at: Persuasive '09: Proceedings of the 4th International Conference on Persuasive Technology; Apr 26-29, 2009. [doi: [10.1145/1541948.1541999](https://doi.org/10.1145/1541948.1541999)]
63. Ryan RM. Psychological needs and the facilitation of integrative processes. *J Pers* 1995 Sep;63(3):397-427. [doi: [10.1111/j.1467-6494.1995.tb00501.x](https://doi.org/10.1111/j.1467-6494.1995.tb00501.x)] [Medline: [7562360](#)]
64. Deci EL. *Intrinsic Motivation*: Springer Science & Business Media; 2012. URL: https://books.google.co.in/books/about/Intrinsic_Motivation.html?id=Ut-BgAAQBAJ&redir_esc=y [accessed 2026-04-02]
65. Seligman ME. *Helplessness: On Depression, Development, and Death*: W. H. Freeman; 1975. URL: https://books.google.co.in/books/about/Helplessness.html?id=De12AQAAAJ&source=kp_book_description&redir_esc=y [accessed 2026-04-02]
66. Ntoumanis N, Ng JY, Prestwich A, et al. A meta-analysis of self-determination theory-informed intervention studies in the health domain: effects on motivation, health behavior, physical, and psychological health. *Health Psychol Rev* 2021 Jun;15(2):214-244. [doi: [10.1080/17437199.2020.1718529](https://doi.org/10.1080/17437199.2020.1718529)] [Medline: [31983293](#)]
67. Oudeyer PY, Kaplan F. What is intrinsic motivation? A typology of computational approaches. *Front Neurobot* 2007;1:6. [doi: [10.3389/neuro.12.006.2007](https://doi.org/10.3389/neuro.12.006.2007)] [Medline: [18958277](#)]
68. Equality Act 2010. [legislation.gov.uk](https://www.legislation.gov.uk/ukpga/2010/15/notes/contents). 2010. URL: <https://www.legislation.gov.uk/ukpga/2010/15/notes/contents> [accessed 2025-12-17]
69. What are the rules on special category data? Information Commissioner's Office. URL: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/special-category-data/what-are-the-rules-on-special-category-data/> [accessed 2025-12-17]
70. What are the conditions for processing? Information Commissioner's Office. URL: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/special-category-data/what-are-the-conditions-for-processing/> [accessed 2025-12-17]
71. A guide to the data protection principles. Information Commissioner's Office. URL: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-protection-principles/a-guide-to-the-data-protection-principles/> [accessed 2025-12-17]
72. PS22/9: a new consumer duty. Financial Conduct Authority. URL: <https://www.fca.org.uk/publications/policy-statements/ps22-9-new-consumer-duty> [accessed 2025-12-17]
73. Automated decision-making, including profiling. Information Commissioner's Office. URL: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/individual-rights/automated-decision-making-and-profiling/> [accessed 2025-12-17]
74. Explaining decisions made with AI. Information Commissioner's Office. URL: <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/> [accessed 2025-12-17]
75. Nishi SK, Kavanagh ME, Ramboanga K, et al. Effect of digital health applications with or without gamification on physical activity and cardiometabolic risk factors: a systematic review and meta-analysis of randomized controlled trials. *EclinicalMedicine* 2024 Oct;76:102798. [doi: [10.1016/j.eclinm.2024.102798](https://doi.org/10.1016/j.eclinm.2024.102798)] [Medline: [39764571](#)]
76. Xu L, Shi H, Shen M, et al. The effects of mHealth-based gamification interventions on participation in physical activity: systematic review. *JMIR Mhealth Uhealth* 2022 Feb 3;10(2):e27794. [doi: [10.2196/27794](https://doi.org/10.2196/27794)] [Medline: [35113034](#)]
77. Chen L, Jang F, Li M, Zong W, Yu H. Effectiveness of mHealth-based gamified interventions on physical activity in older adults: systematic review. *JMIR Aging* 2025 Oct 31;8:e78686. [doi: [10.2196/78686](https://doi.org/10.2196/78686)] [Medline: [41171624](#)]
78. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk handbook of the fundamentals of financial decision making. In: MacLean LC, Ziemba WT, editors. *Handbook of the Fundamentals of Financial Decision Making: Part I*: World Scientific Publishing; 2013:99-127. [doi: [10.1142/9789814417358_0006](https://doi.org/10.1142/9789814417358_0006)]
79. Greysen SR, Oon AL, Harkins K, et al. Effect of gamification with a support partner to increase physical activity in older adults at risk for Alzheimer's disease: the STEP 4Life randomized clinical trial. *Alzheimers Dement* 2024 Aug;20(8):5450-5459. [doi: [10.1002/alz.14058](https://doi.org/10.1002/alz.14058)] [Medline: [38962958](#)]
80. Martinho D, Crista V, Carneiro J, Matsui K, Corchado JM, Marreiros G. Effects of a gamified agent-based system for personalized elderly care: pilot usability study. *JMIR Serious Games* 2023 Nov 23;11:e48063. [doi: [10.2196/48063](https://doi.org/10.2196/48063)] [Medline: [37995116](#)]
81. Santos LH, Okamoto K, Hiragi S, et al. Pervasive game design to evaluate social interaction effects on levels of physical activity among older adults. *J Rehabil Assist Technol Eng* 2019;6:2055668319844443. [doi: [10.1177/2055668319844443](https://doi.org/10.1177/2055668319844443)] [Medline: [31285836](#)]

82. Santos LH, Okamoto K, Funghetto SS, et al. Effects of social interaction mechanics in pervasive games on the physical activity levels of older adults: quasi-experimental study. *JMIR Serious Games* 2019;7(3):e13962. [doi: [10.2196/13962](https://doi.org/10.2196/13962)] [Medline: [31333202](https://pubmed.ncbi.nlm.nih.gov/31333202/)]
83. Santos LH, Okamoto K, Otsuki R, et al. Promoting physical activity in Japanese older adults using a social pervasive game: randomized controlled trial. *JMIR Serious Games* 2021 Jan 6;9(1):e16458. [doi: [10.2196/16458](https://doi.org/10.2196/16458)] [Medline: [33404507](https://pubmed.ncbi.nlm.nih.gov/33404507/)]
84. Randriambelonoro M, Perrin Franck C, Herrmann F, et al. Gamified physical rehabilitation for older adults with musculoskeletal issues: pilot noninferiority randomized clinical trial. *JMIR Rehabil Assist Technol* 2023 Mar 6;10:e39543. [doi: [10.2196/39543](https://doi.org/10.2196/39543)] [Medline: [36877563](https://pubmed.ncbi.nlm.nih.gov/36877563/)]
85. Kawaguchi K, Nakagomi A, Ide K, Kondo K. Effects of a mobile app to promote social participation on older adults: randomized controlled trial. *J Med Internet Res* 2024 Sep 30;26:e64196. [doi: [10.2196/64196](https://doi.org/10.2196/64196)] [Medline: [39348180](https://pubmed.ncbi.nlm.nih.gov/39348180/)]
86. Wang M, Xu J, Zhou X, Li X, Zheng Y. Effectiveness of gamification interventions to improve physical activity and sedentary behavior in children and adolescents: systematic review and meta-analysis. *JMIR Serious Games* 2025 Sep 18;13:e68151. [doi: [10.2196/68151](https://doi.org/10.2196/68151)] [Medline: [40966596](https://pubmed.ncbi.nlm.nih.gov/40966596/)]
87. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011;6:42. [doi: [10.1186/1748-5908-6-42](https://doi.org/10.1186/1748-5908-6-42)] [Medline: [21513547](https://pubmed.ncbi.nlm.nih.gov/21513547/)]
88. Alzghoul B. The effectiveness of gamification in changing health-related behaviors: a systematic review and meta-analysis. *Open Public Health J* 2024 Sep;17:e18749445234806. [doi: [10.2174/0118749445234806240206094335](https://doi.org/10.2174/0118749445234806240206094335)]
89. Yu T, Parry M, Yu T, et al. Effectiveness of mobile health-based gamification interventions for improving physical activity in individuals with cardiovascular diseases: systematic review and meta-analysis of randomized controlled trials. *JMIR Serious Games* 2025 Jan 24;13:e64410. [doi: [10.2196/64410](https://doi.org/10.2196/64410)] [Medline: [39854099](https://pubmed.ncbi.nlm.nih.gov/39854099/)]
90. Michie S, Richardson M, Johnston M, 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](https://doi.org/10.1007/s12160-013-9486-6)] [Medline: [23512568](https://pubmed.ncbi.nlm.nih.gov/23512568/)]
91. Denecke K, Schmid N, Nüssli S. Implementation of cognitive behavioral therapy in e-mental health apps: literature review. *J Med Internet Res* 2022 Mar 10;24(3):e27791. [doi: [10.2196/27791](https://doi.org/10.2196/27791)] [Medline: [35266875](https://pubmed.ncbi.nlm.nih.gov/35266875/)]
92. Timmers T, Janssen L, Kool RB, Kremer JA. Educating patients by providing timely information using smartphone and tablet apps: systematic review. *J Med Internet Res* 2020 Apr 13;22(4):e17342. [doi: [10.2196/17342](https://doi.org/10.2196/17342)] [Medline: [32281936](https://pubmed.ncbi.nlm.nih.gov/32281936/)]
93. Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ. Mobile apps for increasing treatment adherence: systematic review. *J Med Internet Res* 2019 Jun 18;21(6):e12505. [doi: [10.2196/12505](https://doi.org/10.2196/12505)] [Medline: [31215517](https://pubmed.ncbi.nlm.nih.gov/31215517/)]
94. Maier M, Carlotto H, Sanchez F, Balogun S, Merritt S. Transforming underwriting in the life insurance industry. *Proc AAAI Conf Artif Intell* 2019;33(1):9373-9380. [doi: [10.1609/aaai.v33i01.33019373](https://doi.org/10.1609/aaai.v33i01.33019373)]
95. Brackenridge RD, Croxson RS, Mackenzie R. Brackenridge's Medical Selection of Life Risks: Palgrave Macmillan; 2016.
96. Castaneda MA, Marton J. Employer-provided health insurance and the adverse selection problem. *Public Finan Rev* 2013 Jan;41(1):3-36. [doi: [10.1177/1091142111434594](https://doi.org/10.1177/1091142111434594)]
97. Malali N. Artificial intelligence in life insurance underwriting: a risk assessment and ethical implications. *Int J Interdiscip Res Methods* 2025;12(1):36-49. [doi: [10.37745/ijirm.14/vol12n13649](https://doi.org/10.37745/ijirm.14/vol12n13649)]
98. Kang HS, Exworthy M. Wearing the future-wearables to empower users to take greater responsibility for their health and care: scoping review. *JMIR Mhealth Uhealth* 2022 Jul 13;10(7):e35684. [doi: [10.2196/35684](https://doi.org/10.2196/35684)] [Medline: [35830222](https://pubmed.ncbi.nlm.nih.gov/35830222/)]
99. The UK Group Risk industry paid out a record £2.49bn in claims in 2023: £6.82m a day. GRiD. 2024. URL: <https://grouprisk.org.uk/2024/05/21/the-uk-group-risk-industry-paid-out-a-record-2-49bn-in-claims-in-2023-6-82m-a-day/> [accessed 2026-03-28]
100. Dah J, Hussin N, Zaini MK, Isaac Helda L, Senanu Ametefe D, Adozuka Aliu A. Gamification is not working: why? *Games Cult* 2025 Nov;20(7):934-957. [doi: [10.1177/15554120241228125](https://doi.org/10.1177/15554120241228125)]
101. Manzoni JF. Motivation through incentives: a cross-disciplinary review of the evidence. In: Epstein MJ, Manzoni JF, Davila A, editors. *Performance Measurement and Management Control: Innovative Concepts and Practices*: Emerald Group Publishing; 2010, Vol. 20. [doi: [10.1108/S1479-3512\(2010\)0000020005](https://doi.org/10.1108/S1479-3512(2010)0000020005)]
102. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychol Rev* 2015;9(3):323-344. [doi: [10.1080/17437199.2014.941722](https://doi.org/10.1080/17437199.2014.941722)] [Medline: [25104107](https://pubmed.ncbi.nlm.nih.gov/25104107/)]
103. Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. *Health Psychol* 2010 Jan;29(1):1-8. [doi: [10.1037/a0016939](https://doi.org/10.1037/a0016939)] [Medline: [20063930](https://pubmed.ncbi.nlm.nih.gov/20063930/)]
104. Stevely AK, Buykx P, Brown J, et al. Exposure to revised drinking guidelines and “COM-B” determinants of behaviour change: descriptive analysis of a monthly cross-sectional survey in England. *BMC Public Health* 2018 Feb 14;18(1):251. [doi: [10.1186/s12889-018-5129-y](https://doi.org/10.1186/s12889-018-5129-y)] [Medline: [29444647](https://pubmed.ncbi.nlm.nih.gov/29444647/)]
105. Achieving behaviour change: a guide for local government and partners. : Public Health England; 2019 URL: https://assets.publishing.service.gov.uk/media/5e7b4e85d3bf7f133c923435/PHEBI_Achieving_Behaviour_Change_Local_Government.pdf [accessed 2026-03-28]
106. Evidence-based intervention design for behaviour change during a health emergency. UNICEF. 2023. URL: <https://www.unicef.org/innocenti/media/3801/file/Evidence-based-Intervention-Design-Behaviour-Change-2023.pdf> [accessed 2026-03-28]

107. Power S, Broom D, Duncan M, Biddle S, Rowley N. Using the behavior change wheel to design a novel home-based exercise program for adults living with overweight and obesity: comprehensive reporting of intervention development. *Obes Sci Pract* 2024 Jun;10(3):e774. [doi: [10.1002/osp4.774](https://doi.org/10.1002/osp4.774)] [Medline: [38903853](https://pubmed.ncbi.nlm.nih.gov/38903853/)]
108. Ojo SO, Bailey DP, Brierley ML, Hewson DJ, Chater AM. Breaking barriers: using the behavior change wheel to develop a tailored intervention to overcome workplace inhibitors to breaking up sitting time. *BMC Public Health* 2019 Aug 16;19(1):1126. [doi: [10.1186/s12889-019-7468-8](https://doi.org/10.1186/s12889-019-7468-8)] [Medline: [31420033](https://pubmed.ncbi.nlm.nih.gov/31420033/)]
109. Meloncelli N, O'Connor H, de Jersey S, et al. Designing a behaviour change intervention using COM - B and the Behaviour Change Wheel: co - designing the Healthy Gut Diet for preventing gestational diabetes. *J Hum Nutr Diet* 2024 Dec;37(6):1391-1406. [doi: [10.1111/jhn.13355](https://doi.org/10.1111/jhn.13355)] [Medline: [39054768](https://pubmed.ncbi.nlm.nih.gov/39054768/)]
110. Reid H, Smith R, Williamson W, et al. Use of the behaviour change wheel to improve everyday person-centred conversations on physical activity across healthcare. *BMC Public Health* 2022 Sep 20;22(1):1784. [doi: [10.1186/s12889-022-14178-6](https://doi.org/10.1186/s12889-022-14178-6)] [Medline: [36127688](https://pubmed.ncbi.nlm.nih.gov/36127688/)]
111. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008 Sep 29;337:a1655. [doi: [10.1136/bmj.a1655](https://doi.org/10.1136/bmj.a1655)] [Medline: [18824488](https://pubmed.ncbi.nlm.nih.gov/18824488/)]
112. Michie S, Abraham C, Eccles MP, Francis JJ, Hardeman W, Johnston M. Strengthening evaluation and implementation by specifying components of behaviour change interventions: a study protocol. *Implement Sci* 2011;6:10. [doi: [10.1186/1748-5908-6-10](https://doi.org/10.1186/1748-5908-6-10)] [Medline: [21299860](https://pubmed.ncbi.nlm.nih.gov/21299860/)]
113. Marques MM, Wright AJ, Corker E, et al. The behaviour change technique ontology: transforming the Behaviour Change Technique Taxonomy v1. *Wellcome Open Res* 2023;8:308. [doi: [10.12688/wellcomeopenres.19363.2](https://doi.org/10.12688/wellcomeopenres.19363.2)] [Medline: [37593567](https://pubmed.ncbi.nlm.nih.gov/37593567/)]
114. Michie S, Carey RN, Johnston M, et al. From theory-inspired to theory-based interventions: a protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action. *Ann Behav Med* 2018 May 18;52(6):501-512. [doi: [10.1007/s12160-016-9816-6](https://doi.org/10.1007/s12160-016-9816-6)] [Medline: [27401001](https://pubmed.ncbi.nlm.nih.gov/27401001/)]
115. Young MD, Collins CE, Callister R, Plotnikoff RC, Doran CM, Morgan PJ. The SHED-IT weight loss maintenance trial protocol: a randomised controlled trial of a weight loss maintenance program for overweight and obese men. *Contemp Clin Trials* 2014 Jan;37(1):84-97. [doi: [10.1016/j.cct.2013.11.004](https://doi.org/10.1016/j.cct.2013.11.004)] [Medline: [24246820](https://pubmed.ncbi.nlm.nih.gov/24246820/)]
116. Po'e EK, Heerman WJ, Mistry RS, Barkin SL. Growing Right Onto Wellness (GROW): a family-centered, community-based obesity prevention randomized controlled trial for preschool child-parent pairs. *Contemp Clin Trials* 2013 Nov;36(2):436-449. [doi: [10.1016/j.cct.2013.08.013](https://doi.org/10.1016/j.cct.2013.08.013)] [Medline: [24012890](https://pubmed.ncbi.nlm.nih.gov/24012890/)]
117. Balci S, Secaur JM, Morris BJ. Comparing the effectiveness of badges and leaderboards on academic performance and motivation of students in fully versus partially gamified online physics classes. *Educ Inf Technol (Dordr)* 2022;27(6):8669-8704. [doi: [10.1007/s10639-022-10983-z](https://doi.org/10.1007/s10639-022-10983-z)] [Medline: [35308640](https://pubmed.ncbi.nlm.nih.gov/35308640/)]
118. Milne-Ives M, Homer SR, Andrade J, Meinert E. Potential associations between behavior change techniques and engagement with mobile health apps: a systematic review. *Front Psychol* 2023;14:1227443. [doi: [10.3389/fpsyg.2023.1227443](https://doi.org/10.3389/fpsyg.2023.1227443)] [Medline: [37794916](https://pubmed.ncbi.nlm.nih.gov/37794916/)]
119. Kurtzman GW, Day SC, Small DS, et al. Social incentives and gamification to promote weight loss: the LOSE IT randomized, controlled trial. *J Gen Intern Med* 2018 Oct;33(10):1669-1675. [doi: [10.1007/s11606-018-4552-1](https://doi.org/10.1007/s11606-018-4552-1)] [Medline: [30003481](https://pubmed.ncbi.nlm.nih.gov/30003481/)]
120. Edwards EA, Lumsden J, Rivas C, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. *BMJ Open* 2016 Oct 4;6(10):e012447. [doi: [10.1136/bmjopen-2016-012447](https://doi.org/10.1136/bmjopen-2016-012447)] [Medline: [27707829](https://pubmed.ncbi.nlm.nih.gov/27707829/)]
121. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. *Int J Behav Nutr Phys Act* 2014 Jul 25;11:97. [doi: [10.1186/s12966-014-0097-9](https://doi.org/10.1186/s12966-014-0097-9)] [Medline: [25059981](https://pubmed.ncbi.nlm.nih.gov/25059981/)]
122. Litvin S, Saunders R, Maier MA, Lüttke S. Gamification as an approach to improve resilience and reduce attrition in mobile mental health interventions: a randomized controlled trial. *PLoS One* 2020 Sep 2;15(9):e0237220. [doi: [10.1371/journal.pone.0237220](https://doi.org/10.1371/journal.pone.0237220)] [Medline: [32877425](https://pubmed.ncbi.nlm.nih.gov/32877425/)]
123. Czerska I. Gamification in health promotion in the context of creating the concept of Health 2.0 - a comparative analysis of selected health mobile applications. *Eur Res Stud J* 2024;XXVII(4):985-1007. [doi: [10.35808/ersj/3641](https://doi.org/10.35808/ersj/3641)]
124. Morschheuser BS, Rivera-Pelayo V, Mazarakis A, Zacharias V. Interaction and reflection with quantified self and gamification: an experimental study. *J Lit Technol* 2014;15(2):136-156 [FREE Full text]
125. Patel MS, Small DS, Harrison JD, et al. Effect of behaviorally designed gamification with social incentives on lifestyle modification among adults with uncontrolled diabetes: a randomized clinical trial. *JAMA Netw Open* 2021 May 3;4(5):e2110255. [doi: [10.1001/jamanetworkopen.2021.10255](https://doi.org/10.1001/jamanetworkopen.2021.10255)] [Medline: [34028550](https://pubmed.ncbi.nlm.nih.gov/34028550/)]
126. Impact study. Vitality Group. 2024. URL: https://insights.vitalitygroup.com/viewer/65c559efc9c9802b5fd713c0?_gl=1*1rvp8k7*_gcl_au*MTA0NzU2NjczNC4xNzQ5MTYyMTg3#1 [accessed 2026-03-29]
127. Soliño-Fernandez D, Ding A, Bayro-Kaiser E, Ding EL. Willingness to adopt wearable devices with behavioral and economic incentives by health insurance wellness programs: results of a US cross-sectional survey with multiple consumer health vignettes. *BMC Public Health* 2019 Dec 16;19(1):1649. [doi: [10.1186/s12889-019-7920-9](https://doi.org/10.1186/s12889-019-7920-9)] [Medline: [31839006](https://pubmed.ncbi.nlm.nih.gov/31839006/)]
128. Bitrián P, Buil I, Catalán S. Enhancing user engagement: the role of gamification in mobile apps. *J Bus Res* 2021 Aug;132:170-185. [doi: [10.1016/j.jbusres.2021.04.028](https://doi.org/10.1016/j.jbusres.2021.04.028)]

129. Roomer J. How YuLife uses game mechanics to drive 3x the average benefits engagement rates. YuLife. 2022 Apr 20. URL: <https://yulife.com/blog/how-yulife-uses-game-mechanics-to-drive-3-times-the-average-benefits-engagement-rates/> [accessed 2026-03-29]
130. Engagement white paper. Vitality. 2022. URL: <https://view-su2.highspot.com/viewer/62e7fa0dc9c48ad0468c8b65> [accessed 2026-03-29]
131. Adviser FAQs. AIA Australia. URL: https://www.aia.com.au/content/dam/au/en/docs/Adviser/aia-vitality/Adviser_FAQs.pdf [accessed 2026-03-29]
132. Seifert CM, Chapman LS, Hart JK, Perez P. Enhancing intrinsic motivation in health promotion and wellness. *Am J Health Promot* 2012;26(3):TAHP1-TAHP12. [doi: [10.4278/ajhp.26.3.tahp](https://doi.org/10.4278/ajhp.26.3.tahp)] [Medline: [22208425](https://pubmed.ncbi.nlm.nih.gov/22208425/)]
133. Rothschild C, Thistle PD. Supply, demand, and selection in insurance markets: theory and applications in pictures. *Risk Manag Insur Rev* 2022;25(4):419-444. [doi: [10.1111/rmir.12233](https://doi.org/10.1111/rmir.12233)]
134. Ho CW, Ali J, Caals K. Ensuring trustworthy use of artificial intelligence and big data analytics in health insurance. *Bull World Health Organ* 2020 Apr 1;98(4):263-269. [doi: [10.2471/BLT.19.234732](https://doi.org/10.2471/BLT.19.234732)] [Medline: [32284650](https://pubmed.ncbi.nlm.nih.gov/32284650/)]
135. Chen Y. Exploring design guidelines of using user-centered design in gamification development: a delphi study. *Int J Hum Comput Interact* 2019 Aug 9;35(13):1170-1181. [doi: [10.1080/10447318.2018.1514823](https://doi.org/10.1080/10447318.2018.1514823)]
136. Lythreath S, Singh SK, El-Kassar AN. The digital divide: a review and future research agenda. *Technol Forecast Soc Change* 2022 Feb;175:121359. [doi: [10.1016/j.techfore.2021.121359](https://doi.org/10.1016/j.techfore.2021.121359)]
137. Yau KW, Tang TS, Görges M, et al. Effectiveness of mobile apps in promoting healthy behavior changes and preventing obesity in children: systematic review. *JMIR Pediatr Parent* 2022 Mar 28;5(1):e34967. [doi: [10.2196/34967](https://doi.org/10.2196/34967)] [Medline: [35343908](https://pubmed.ncbi.nlm.nih.gov/35343908/)]
138. Mitra S, Kroeger CM, Wang T, et al. The impact of gamified smartphone app interventions on behaviour and metabolic outcomes in individuals at risk of cardiovascular disease. *Stud Health Technol Inform* 2024 Sep 24;318:172-173. [doi: [10.3233/SHTI240913](https://doi.org/10.3233/SHTI240913)] [Medline: [39320202](https://pubmed.ncbi.nlm.nih.gov/39320202/)]
139. Tran S, Smith L, El-Den S, Carter S. The use of gamification and incentives in mobile health apps to improve medication adherence: scoping review. *JMIR Mhealth Uhealth* 2022 Feb 21;10(2):e30671. [doi: [10.2196/30671](https://doi.org/10.2196/30671)] [Medline: [35188475](https://pubmed.ncbi.nlm.nih.gov/35188475/)]
140. Yoshida-Montezuma Y, Ahmed M, Ezezika O. Does gamification improve fruit and vegetable intake in adolescents? A systematic review. *Nutr Health* 2020 Dec;26(4):347-366. [doi: [10.1177/0260106020936143](https://doi.org/10.1177/0260106020936143)] [Medline: [32703067](https://pubmed.ncbi.nlm.nih.gov/32703067/)]
141. Suleiman-Martos N, García-Lara RA, Martos-Cabrera MB, et al. Gamification for the improvement of diet, nutritional habits, and body composition in children and adolescents: a systematic review and meta-analysis. *Nutrients* 2021 Jul 20;13(7):2478. [doi: [10.3390/nu13072478](https://doi.org/10.3390/nu13072478)] [Medline: [34371989](https://pubmed.ncbi.nlm.nih.gov/34371989/)]
142. Rajani NB, Bustamante L, Weth D, Romo L, Mastellos N, Filippidis FT. Engagement with gamification elements in a smoking cessation app and short-term smoking abstinence: quantitative assessment. *JMIR Serious Games* 2023 Feb 1;11:e39975. [doi: [10.2196/39975](https://doi.org/10.2196/39975)] [Medline: [36724003](https://pubmed.ncbi.nlm.nih.gov/36724003/)]
143. Nurmi J, Knittle K, Ginchev T, et al. Engaging users in the behavior change process with digitalized motivational interviewing and gamification: development and feasibility testing of the precious app. *JMIR Mhealth Uhealth* 2020 Jan 30;8(1):e12884. [doi: [10.2196/12884](https://doi.org/10.2196/12884)] [Medline: [32003750](https://pubmed.ncbi.nlm.nih.gov/32003750/)]
144. Cheng VW, Davenport T, Johnson D, Vella K, Hickie IB. Gamification in apps and technologies for improving mental health and well-being: systematic review. *JMIR Ment Health* 2019 Jun 26;6(6):e13717. [doi: [10.2196/13717](https://doi.org/10.2196/13717)] [Medline: [31244479](https://pubmed.ncbi.nlm.nih.gov/31244479/)]
145. Cheng VWS. Recommendations for implementing gamification for mental health and wellbeing. *Front Psychol* 2020;11:586379. [doi: [10.3389/fpsyg.2020.586379](https://doi.org/10.3389/fpsyg.2020.586379)] [Medline: [33365001](https://pubmed.ncbi.nlm.nih.gov/33365001/)]
146. Huston P. A sedentary and unhealthy lifestyle fuels chronic disease progression by changing interstitial cell behaviour: a network analysis. *Front Physiol* 2022;13:904107. [doi: [10.3389/fphys.2022.904107](https://doi.org/10.3389/fphys.2022.904107)] [Medline: [35874511](https://pubmed.ncbi.nlm.nih.gov/35874511/)]
147. Heron L, O'Neill C, McAneney H, Kee F, Tully MA. Direct healthcare costs of sedentary behaviour in the UK. *J Epidemiol Community Health* 2019 Jul;73(7):625-629. [doi: [10.1136/jech-2018-211758](https://doi.org/10.1136/jech-2018-211758)] [Medline: [30910857](https://pubmed.ncbi.nlm.nih.gov/30910857/)]
148. Zikra AA. The impact of gamified interventions on youth health literacy: a systematic review of effectiveness and implementation challenges. *JRKPK* 2023;2(2):79-91. [doi: [10.61194/jrkpk.v2i2.666](https://doi.org/10.61194/jrkpk.v2i2.666)]
149. Schmidt-Kraepelin M, Warsinsky S, Thiebes S, Sunyaev A. The role of gamification in health behavior change: a review of theory-driven studies. Presented at: 53rd Hawaii International Conference on System Sciences; Jan 7-10, 2020. [doi: [10.24251/HICSS.2020.155](https://doi.org/10.24251/HICSS.2020.155)]

Abbreviations

- BCI:** behavior change intervention
- BCT:** Behavior Change Technique
- BCW:** Behavior Change Wheel
- COM-B:** capability, opportunity, motivation, and behavior
- NHS:** National Health Service
- PA:** physical activity

Edited by S Brini; submitted 15.Jul.2025; peer-reviewed by S Wang, S Göbel, T Lima; revised version received 06.Feb.2026; accepted 14.Feb.2026; published 13.Apr.2026.

Please cite as:

Salami A, Papastylianou T, Adeoye M, Ronayne J, Bixby H, Stawski RS, Fromson B, Doltis M, Mahmoud O, Di Cesare M
Is Gamification the New Panacea for Health Behavioral Changes? Implications for the Health and Life Insurance Industry
JMIR Serious Games 2026;14:e80684

URL: <https://games.jmir.org/2026/1/e80684>

doi: [10.2196/80684](https://doi.org/10.2196/80684)

© Abbas Salami, Tasos Papastylianou, Marvellous Adeoye, John Ronayne, Honor Bixby, Robert S Stawski, Bernard Fromson, Matt Doltis, Osama Mahmoud, Mariachiara Di Cesare. Originally published in JMIR Serious Games (<https://games.jmir.org>), 13.Apr.2026. 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.

Original Paper

A Haptic-Driven Serious Game for Cognitive Stimulation and Visual Impairment Mitigation in Older Adults Based on the Design-Play-Experience Framework: Cross-Sectional Mixed Methods Pilot Study

Xin Huang¹, MA; Nazlena Mohamad Ali¹, PhD; Shafrida Sahrani¹, PhD; Yue Zhang¹, MA

Institute of Visual Informatics (IVI), Universiti Kebangsaan Malaysia, Selangor, Malaysia

Corresponding Author:

Nazlena Mohamad Ali, PhD

Institute of Visual Informatics (IVI)

Universiti Kebangsaan Malaysia

Aras 1, Akademia Siber Teknopolis, Universiti Kebangsaan Malaysia

Selangor, 43600

Malaysia

Phone: 60 389272402 ext 2425

Fax: 60 389272402

Email: nazlena.ali@ukm.edu.my

Abstract

Background: In the context of global aging, cognitive decline among older adults has become a prevalent issue, significantly impacting their daily lives. Serious games have demonstrated potential in enhancing cognitive abilities in this population. However, most existing serious games designed for older adults rely heavily on visual interfaces, which are often potentially detrimental for those with pre-existing visual impairments.

Objective: This study had two primary objectives: (1) to design a theoretical prototype for a haptic-driven serious game for older adults based on the Design-Play-Experience (DPE) framework, aiming to enhance cognitive abilities, including attention, logical reasoning, and decision-making while simultaneously mitigating challenges associated with visual impairment, and (2) to conduct a pilot study evaluating the prototype's usability, accessibility, and user experience within the target population.

Methods: We used a cross-sectional, mixed methods pilot study with a single-group observational design, comprising a theoretical design and a pilot user study. First, the DPE framework was systematically applied to develop a game prototype by integrating haptic feedback technology (using built-in smartphone vibration motors) across its 3 core dimensions: design (haptic symbol system, accessible interface), play (dynamic difficulty adjustment), and experience (emotional engagement). Subsequently, a pilot study was conducted with 10 older adults recruited via convenience sampling (mean age 62.9, SD 3.35 years; 5 male, 5 female; all with self-reported mild visual impairments, such as presbyopia). Following interaction with the prototype, data were collected remotely using the System Usability Scale (SUS) and semistructured interviews administered via videoconferencing. Quantitative data from the SUS were analyzed using descriptive statistics, while qualitative data from the interviews were processed using thematic analysis.

Results: Pilot user studies showed that the game prototype had good usability, with an average SUS score of 89.5 (SD 2.72; 95% CI 87.6-91.4), which is considered "excellent." Thematic analysis of the interviews revealed three significant themes. The first theme was intuitive haptic feedback, which reflected that participants were able to quickly grasp and value the vibrational cues used to identify cards. The second theme was based on reduced eye strain, in which the combination of large fonts, high-contrast interfaces, and haptic feedback was praised for its effectiveness in relieving eye strain. The third theme was simplicity, where the simplified card game mechanics were considered both fun and challenging.

Conclusions: This study developed and validated a haptic, serious game for older adults. Its innovation lies in the systematic application of the DPE framework to achieve "haptic substitution for vision," which differs from previous research that focused on general immersive experiences. The main contribution of this study is providing a reusable design blueprint for creating easy-to-use cognitive training tools. These findings have practical implications in the real world, providing a feasible approach for deploying low visual load interventions in communities and care facilities.

KEYWORDS

serious games; haptic feedback; older adults; cognitive abilities; visual impairment; DPE framework; Design-Play-Experience; pilot study

Introduction

By 2050, two-thirds of the global population aged 60 years and older will live in low- and middle-income countries [1]. The population structure is undergoing a significant aging transition, and the most common cognitive diseases in older adults include mild cognitive impairment [2], dementia [3], and Parkinson disease [4]. Mild cognitive impairment is a cognitive disorder between dementia and health [5]. At this time, the patient's memory and cognitive function have already had problems, and one or more cognitive functions have declined, but the patient's daily life will not be significantly affected [6,7]. Dementia affects 36 million people worldwide [8]. Dementia is closely associated with age (aging), with nearly half of the population aged 85 years and older developing dementia [9]. The most common type of dementia is Alzheimer disease. The typical initial symptom is memory loss. Parkinson disease usually occurs in people aged 60 years and older. It is a chronic neurodegenerative disease that affects the central nervous system, with the most obvious early symptoms being tremors, limb stiffness, decreased movement, and gait abnormalities [10]. Parkinson disease may also cause cognitive and behavioral problems.

Serious games have had some success in helping older adults recover cognitive abilities [11]. Serious games are a set of solutions designed to make a range of rehabilitation courses more engaging and less boring [12]. MINWii (Samuel Benveniste, Pierre Jouvelot, and Renaud Pequignot) is a new serious video game tailored for people with Alzheimer disease and dementia, in which older gamers use Wiimotes to improvise or play predefined songs on a virtual keyboard displayed on the screen [13]. However, vision impairments in older adults may affect their interaction with serious games [14]. Older adults often experience sensory changes that affect their interaction with digital displays and games. These changes include decreased vision, impaired dark adaptation, reduced contrast sensitivity, limited visual accommodation range, reduced color sensitivity, and increased sensitivity to glare [14]. These impairments can make it difficult for older adults to perceive small elements on a screen, read text, or navigate complex interfaces [15]. The older adults said that the mobile phone screen was too small when interacting, which shows that the limitations of digital tools themselves are also one of the factors that affect the older adults' interaction with serious games [16]. These studies show that, in addition to the decline of the older adult's own sensory organs over time, serious games designed for older adults do not fully consider the older adult's user experience.

The addition of haptic feedback technology further expands the impact of serious games [17] and can reduce the visual damage caused by long-term serious game playing by older adults. The haptic experience provides a new way of perception for older

adults, thereby enhancing the user's cognitive ability. Haptic technology is a sensory feedback technology that simulates haptic perception through force, vibration, or motion. It captures the user's movements and interaction information through sensors and generates corresponding haptic feedback through actuators, allowing users to feel real physical interactions in a virtual environment [17]. Haptic feedback technology provides an immersive experience for older adults by combining their sense of touch with what they see, hear, or interact with [18]. At the same time, serious games combined with haptic feedback technology have shown the potential to significantly improve cognitive abilities [19]. For example, in mobile serious games, the interaction method developed by Deng et al [20] that combines eye tracking with haptic feedback also significantly improves the user's operation accuracy and task completion rate. By providing haptic feedback on mobile devices, the user's error rate in interactive tasks is significantly reduced, indicating that serious games combined with haptic feedback can effectively enhance the user's understanding and control of cognitive abilities [20]. Silva et al [21] reported a new haptic device under development, mainly for older adults, to stimulate and quantify the response of the older adults' nervous system through serious games.

The existing combination of serious games and haptic feedback technology mainly focuses on enhancing interactive immersion. Although it has improved the cognitive ability of some older people, there are few system frameworks specifically for older adults with visual impairment. There is a lack of deep integration of cognitive training and haptic feedback. Its design is mostly oriented to general users and lacks targeted adaptation for the visual and haptic coordination needs of older adults. There are 2 major limitations in the current research on the use of serious games by older adults, one being the problem of visual dependence. Most serious games rely on complex visual interfaces, which aggravate the interaction barriers caused by vision decline in older adults. The second is the one-sidedness of haptic feedback design. Although haptic technology can enhance immersion through force feedback, its application focuses on general scenarios and lacks a system framework for the cognitive decline and sensory coordination needs of older adults (such as "haptic substitution for vision").

To address these shortcomings, this study uses the Design-Play-Experience (DPE) framework [22] as its theoretical basis to develop a prototype of a haptic-driven serious card game. The DPE framework provides a structured approach to integrate technology, game mechanics, and user experience, making it well-suited for systematically incorporating haptic feedback to achieve specific cognitive and accessibility goals. The final prototype embodies this integration (1) in the design dimension, through a haptic symbol system and a high contrast interface; (2) in the play dimension, through haptic cues to dynamically adjust the game difficulty; and (3) in the experience

dimension, through haptic rewards to enhance emotional engagement. The objective of this study is to develop a haptic feedback serious card game theoretical prototype based on the DPE framework to reduce the visual burden of older adults and improve their cognitive ability. Second, to conduct a pilot study evaluating the prototype's usability, accessibility, and user experience within the target population. The contribution of this study is to combine haptic feedback with the card game mechanism and provide a reusable design blueprint for the design of older adult-friendly serious games through the system mapping of the DPE framework.

Methods

Quantitative Analysis

Inclusion and Exclusion

The inclusion criteria for participants were (1) age ≥ 60 years; (2) self-reported mild visual impairment, such as presbyopia; (3) basic smartphone usage skills; and (4) willingness to sign an informed consent form. The exclusion criteria included (1) being diagnosed with severe cognitive impairment, such as clinically diagnosed dementia, and (2) having severe hearing loss or motor impairment that significantly affects haptic perception or device operation.

Participant Characteristics

Participants had a mean age of 62.9 (SD 3.35) years, with an age range of 60-71 years, and comprised 5 males and 5 females. All participants reported having mild visual impairment, such as presbyopia, and possessed basic smartphone operation skills.

Sampling Procedures

Participants were recruited from Dangtu Old Age University, using convenience sampling [23]. Recruitment and data collection were conducted remotely from June 1 to July 1, 2025. The research team also collaborated with the institution's administration to disseminate recruitment information through employee referrals. Interested potential seniors contacted the research team by phone, and those who passed the initial screening were formally enrolled.

Sample Size, Power, and Precision

As a pilot study, the sample size (N=10) was determined primarily based on feasibility considerations rather than statistical power calculations. This sample size is typical for pilot studies and aims to provide preliminary data and process validation for subsequent large-scale efficacy trials.

Measures and Covariates

The primary outcome measure was usability of the game prototype, measured using the System Usability Scale (SUS). The scale's total score ranges from 0 to 100. The secondary outcome measure includes accessibility and user experience. In covariates and confounding factors, participants' baseline cognitive level, severity of visual impairment, technical experience, and age were recorded as potential confounding factors and qualitatively documented during interviews via background questions.

Data Collection

Quantitative data were collected through online survey platforms, such as Wenjuanxing [24]. Participants completed a SUS [25] online after interacting with the game prototype.

Quality of Measurements

To ensure data quality, all data collection procedures were standardized. Specifically, all participants used the same game prototype and completed the same SUS. Furthermore, a consistent interview process was followed, with semistructured interviews [26] conducted by the same researcher after participants had finished experiencing the game, using interview guidelines to ensure consistency across all interviews. The SUS [25], the primary quantitative tool in this study, has had its reliability and validity extensively confirmed in previous research [27,28].

Instrumentation

The main tools used in this study were the SUS and a haptic game prototype developed based on the DPE framework. This prototype runs on an Android smartphone and uses the device's built-in linear resonant motor to provide structured haptic feedback.

Masking

Given that this study used a single-group observational design aimed at directly obtaining users' subjective feedback on the prototype, no masking was applied to participants or researchers.

Psychometrics

This study used the widely validated SUS, whose reliability and validity are supported by literature [27]. SUS has demonstrated excellent psychometric properties, including good reliability (eg, high internal consistency, Cronbach alpha coefficient typically above 0.85, and test-retest reliability of approximately 0.80) and strong validity (convergent validity, discriminant validity, and construct validity) [25,27,28]. This study directly reports the SUS scores observed in the current sample.

Conditions and Design

This study uses a nonexperimental observational design, specifically a single-group, cross-sectional study. Reporting of the study was performed in accordance with the JARS (Journal Article Reporting Standards) guidelines [29].

Data Diagnostics

The study conducted diagnostic checks on the data. All 10 participants completed the entire study process, and there were no missing data for either the primary or secondary outcome measures. Therefore, no imputation was performed. Given the small sample size, the distribution of SUS scores was examined. No extreme outliers requiring intervention were found, and no data transformation was performed.

Analytic Strategy

Quantitative data analysis used descriptive statistics. The total score of the SUS was described using the mean and SD, and 95% CIs based on a *t* test distribution were calculated to provide more conservative interval estimates for small samples. All

analyses aimed to assess the initial usability of the game prototype.

Qualitative Analysis

Research Design Overview

The qualitative portion of this study used a thematic analysis [30] approach to explore older users' experiences, perceived accessibility, and emotional engagement with haptic-driven game prototypes. This design was well-suited for gathering detailed participant-centered perspectives from semistructured interviews.

Study Participants or Data Sources

The qualitative data came from the same 10 participants (mean age of 62.9, SD 3.35 years; 5 male and 5 female) as the quantitative data. The data were in the form of audio recordings of semistructured interviews with each participant and their verbatim transcripts.

Participant Recruitment

The recruitment process for participants was the same as that for the quantitative part. Convenience sampling was used to recruit from Dangtu Old Age University. The research team collaborated with the institution to make initial contact, and eligible volunteers were enrolled after confirmation by phone. All participants in the quantitative assessment completed subsequent qualitative interviews to ensure the continuity of data sources.

Data Collection

This study collected qualitative data through semistructured interviews conducted using videoconferencing software, such as Tencent Meeting [31]. We developed an interview guideline to explore participants' experiences in three areas: (1) the intuitiveness of haptic feedback, (2) the impact of haptic feedback on alleviating visual fatigue, and (3) the overall appeal and engagement of the game. This approach ensured comprehensive data collection. All interviews were conducted immediately after participants completed the game experience and SUSs to ensure the freshness of the experience. Data collection continued until thematic saturation was reached. All interviews were recorded with the participants' consent and transcribed by researchers for analysis. All interviews were recorded with the participants' consent and transcribed by researchers for analysis.

Ethical Considerations

The study was approved by the ethics committee before it was carried out, and it strictly adhered to the ethical guidelines for research involving human participants. The study protocol, including the data collection methods, was considered to comply with the ethical standards for research involving human participants (approval reference UKM.IVI.600/8/1-P136397). Written informed consent was obtained from all participants before participating in the study. The consent form clearly outlined the purpose, procedures, risks, benefits, and voluntary nature of participating in the study. Participants were informed of their right to withdraw at any time without penalty. For secondary data analysis, the original consent form explicitly allowed the use of deidentified data for research purposes without the need for additional consent. All participant data were anonymized and deidentified during collection and analysis. Personally identifiable information was removed from survey responses, and data were securely stored on a password-protected server accessible only to the research team. No identifiable information will be shared with external parties or institutions. Participants received no monetary or nonmonetary compensation for their participation in this study. Their contribution was voluntary and motivated by the potential societal benefits of the research. No identification of individual participants in any images of the paper or supplementary material is possible. All data reported were aggregated and anonymized to ensure participant confidentiality. This study conformed to the tenets of the Declaration of Helsinki and adhered to the ethical standards outlined by JMIR for research involving human participants.

Results

Participant Flow and Demographics

All 10 recruited older participants (5 male and 5 female; mean age 62.9, SD 3.35; 60-71 years) completed the entire research process, including interactive games, the SUS questionnaire, and interviews. No data were missing for either the primary or secondary outcome measures.

Quantitative Analysis

Quantitative analysis of the pilot study showed that the haptic game prototype received a very high usability rating. The mean score on the SUS was 89.5 (SD 2.72; 95% CI 87.6-91.4). According to the scale's evaluation criteria, this score falls into the "excellent" category, clearly indicating that the prototype possesses both high usability and high acceptability among the target older adults (Table 1).

Table 1. System Usability Scale (SUS) scores from a cross-sectional pilot study evaluating a haptic-driven serious game prototype for cognitive stimulation and visual impairment mitigation in older adults.

Metric	Value
SUS score, mean (SD; 95% CI)	89.5 (2.72; 87.6-91.4)
Adjective rating	Excellent
Score range	85.0-92.5

Qualitative Analysis

After conducting a thematic analysis of interviews, researchers identified three prominent themes regarding user experience and usability:

1. **Intuitive haptic feedback:** Participants quickly grasped and highly valued the vibrational cues used to identify cards. This haptic feedback was considered a natural and effective alternative to constant visual confirmation. One participant stated, “I quickly learned what the different vibrations represented; I didn’t have to constantly stare at the screen to remember what cards I was holding.”
2. **Effective reduction of eye strain:** The combination of haptic feedback and a high-contrast, large-font visual interface significantly reduced eye strain and made gameplay more comfortable. One participant commented, “The large font and vibrational cues combined made my eyes much more comfortable.”
3. **Simplicity:** The simplified game mechanics (a simplified 7-card Doudizhu) were considered both fun and challenging, lowering the learning curve while maintaining player

engagement through a dynamic difficulty adjustment system. One participant said, “It’s easy to pick up, but still requires thought. It’s fun without being frustrating.”

A Theoretical Prototype of Serious Card Games Based on the DPE Framework

Table 2 is a mapping table of the DPE framework in game design. It comprehensively elaborates on the DPE framework for serious games aimed at older adults from 3 vertical dimensions (design, play, and experience) and 4 horizontal dimensions (learning, storytelling, gameplay, and user experience). The paper designs a theoretical prototype of a serious game with haptic feedback, aiming to improve the cognitive abilities of older adults and alleviate their visual impairments. In the design dimension, we created a haptic symbol system to reduce reliance on vision. In the gameplay dimension, we designed haptic cues to facilitate dynamic adjustment of game difficulty, and in the experience dimension, we added haptic rewards to enhance the player’s emotional engagement.

Table 2. Prototype structure mapping for a haptic-driven serious card game based on the Design-Play-Experience (DPE) framework.

Core Pillar	Design	Play	Experience
Learning	<ul style="list-style-type: none"> • Content: Target cognitive abilities, such as attention, thinking, decision-making, and logical reasoning. • Teaching methods: digital cognition, simple mathematical operations, and decision-making. 	<ul style="list-style-type: none"> • Teaching: Roles: Players play community members and play card games with virtual opponents. • Setting: Community environment, build the community by winning games. • Narrative: Earn coins by winning games, and build the community until the community is complete. 	<ul style="list-style-type: none"> • Learning: Through gaming activities, older adults can practice and improve their cognitive abilities. • Self-efficacy: By successfully completing tasks through gaming, older adults can enhance their sense of self-efficacy and improve their self-confidence.
Storytelling	<ul style="list-style-type: none"> • Roles: Players play community members and play card games with virtual opponents. • Setting: Community environment where players build the community by winning games. • Narrative: Earn coins by winning games, and build the community until the community is complete. 	<ul style="list-style-type: none"> • Storytelling: Players win coins and build communities by winning games. • Interactive narrative: The storyline can have different development paths based on the player’s choices and performance. 	<ul style="list-style-type: none"> • Story: Provide a rich story experience through the game storyline to increase the appeal of the game.
Gameplay	<ul style="list-style-type: none"> • Mechanics: A simplified version of Doudizhu, with only 7 cards, smaller than playing cards. Haptic feedback prompts, identifying the size of the card based on the number of vibrations (card 1, slight vibration 1 time, and so on). 	<ul style="list-style-type: none"> • Dynamics: A smooth gameplay flow ensures a cohesive and engaging experience. • Difficulty Adjustment: Haptic feedback automatically adjusts the game difficulty based on player performance, maintaining the game’s challenge and fun. 	<ul style="list-style-type: none"> • Emotion: Design game elements that evoke positive emotional experiences, such as a sense of accomplishment and satisfaction. Provide unique haptic feedback as a reward when players achieve something or win, enhancing emotional engagement and a sense of accomplishment.
User experience	<ul style="list-style-type: none"> • User interface: Large fonts, high contrast, and a simple operation interface are suitable for older adults. • Navigation: Clear navigation and instructions help older adults easily find the required functions. 	<ul style="list-style-type: none"> • Interactivity: Encourage players to interact with virtual opponents through game tasks. • Feedback interaction: Players’ actions will receive immediate feedback, enhancing their sense of participation. 	<ul style="list-style-type: none"> • Engagement: Design game elements that can attract players to continue to participate, such as achievement systems, reward mechanisms, etc. Enhance players’ sense of belonging and community through the game community.

Game Prototype Interface Display

This study follows a prototype based on the DPE framework and systematically develops the visual and interactive design of the “Old Friends” game. We follow the WCAG 2.1 [32] accessibility standard and are explicitly committed to minimizing the visual and cognitive burden on older users throughout the design process.

The interface design intentionally used large, clear, and legible fonts (18 pt) and high contrast (4.5:1), consistently applied to all core visual components, including the game logo, card elements, login paper, main game, win or lose, and homeland interface. These choices directly supported the user experience goal emphasized in the design framework, focusing on clarity and usability. A key outcome of this process was the “Old Friends” logo (Figure 1). Its design exemplifies how

accessibility principles translate into tangible visual identity. We chose the Bauhaus 93 font for its high legibility and clear glyphs, and simultaneously, we used a simple black-on-white color scheme to ensure maximum readability in accordance with the “minimum visual load principle.” However, this functional clarity is complemented by carefully designed emotional elements. The name “Old Friends” and its warm, friendly presentation aim to resonate with the project’s core objectives, emotional support, and social interaction, as detailed in our “Storytelling” and “Experience” dimensions. We believe this semantic and visual consistency enhances the emotional connection between users and the game, thereby increasing user engagement and strengthening community awareness. Therefore, the logo is not merely an aesthetic element but a reflection of a design philosophy that perfectly blends technological usability with an empathetic, goal-oriented narrative design.

Figure 1. Logo of the “Old Friends” serious card game, compliant with WCAG (Web Content Accessibility Guidelines) 2.1 standards and embodying high-contrast visual identity principles.



Figure 2 shows the final card design, developed through iterative iterations while consistently prioritizing senior friendliness. A core design decision was to use the clear and legible Berlin Sans Demi font to represent the numbers 1-7 at an oversized 120-point font size to ensure quick recognition. This design, combined with a high-contrast black-on-white scheme, forms the functional foundation of the card. Subsequently, the designers consciously layered various aesthetic elements on top of this easily understood foundation. Light blue and medium

yellow circular decorations were introduced, adding visual interest and warmth without compromising clarity. The “Old Friends” logo, with its coordinated typography, visually and thematically unifies the entire component. This fusion of clear functionality and warm aesthetics was a thoughtful choice, designed to directly support the research goals of “emotional support” and “social interaction” by making the interface more appealing and personal for older users.

Figure 2. Game card interface of the “Old Friends” serious card game.



The login screen (Figure 3) is designed to immediately create a user-friendly atmosphere. To this end, we used a soft, light background, accented with soft yellow, orange, and blue circular graphic elements. This color scheme was deliberately chosen to create a warm and vibrant atmosphere without being overly stimulating to the senses, directly adhering to the “minimum visual load principle” and minimizing strong contrasts in unnecessary areas. Functional clarity is paramount. The central title, “Old Friends,” uses a high-contrast 60-point Bauhaus 93 font to ensure users can immediately recognize the brand. Similarly, the key operation buttons at the bottom of the screen

(Start, Register, Homeland, Archives) use a standard, clear, and legible 18-point Arial font. They are arranged with ample spacing on a white background, a carefully designed interaction method intended to prevent accidental clicks and guide users easily into the game. Beyond functionality, the illustration of older people playing card games in the background also carries significant narrative meaning. It was included to visually convey the core themes of “emotional support” and “social interaction” from the outset, preparing users for a positive social experience and fostering a direct emotional connection with the game’s objectives.

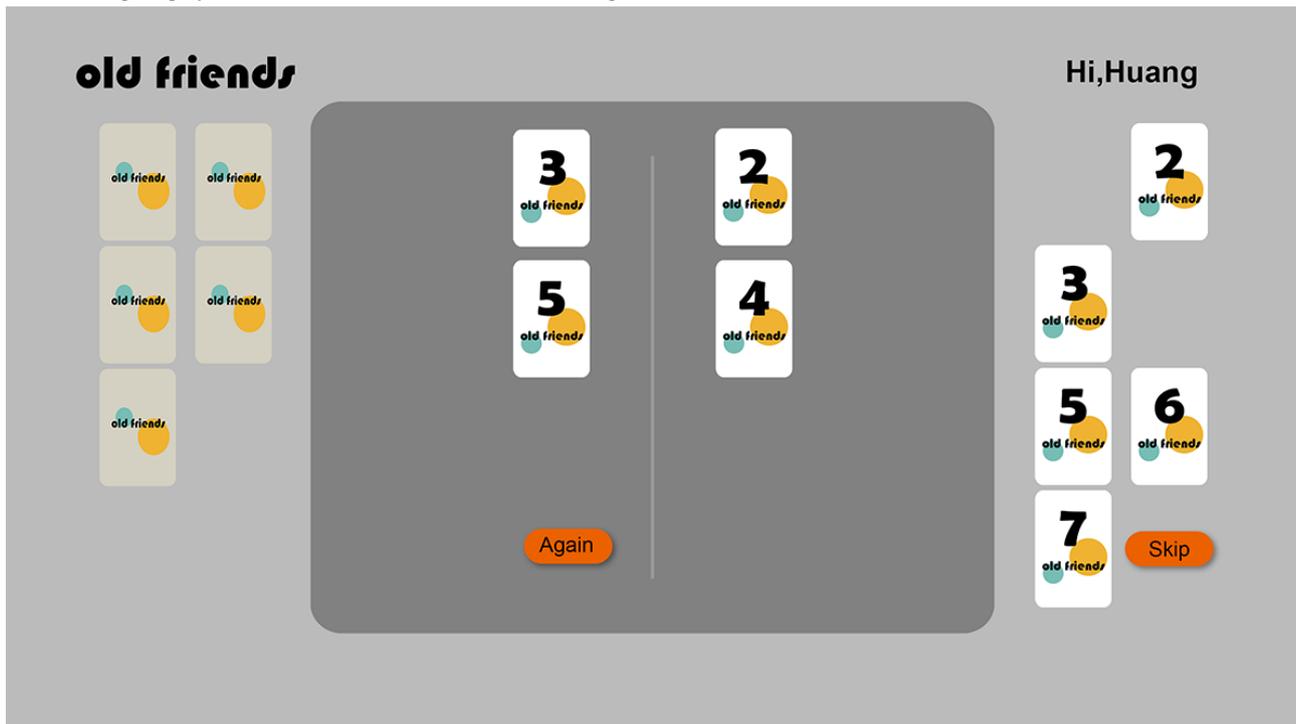
Figure 3. Login page interface of the “Old Friends” serious card game.



The design of the game’s main interface (Figure 4) continues the aesthetic and functional principles established on the login page, aiming to create a low-stimulation environment for extended gameplay. We chose a neutral gray background, forming a calm and professional base color scheme that effectively reduces visual fatigue, adhering to the “minimum visual load principle.” Against this background, key interactive components, such as white cards and orange function buttons, are presented with high contrast to ensure immediate visual recognition and readability. The interface layout is carefully planned to guide the user’s interaction flow, with the left panel displaying the opponent’s hand, while the right area is dedicated to displaying the user’s hand, creating a clear and intuitive game cycle. The “Again” and “Skip” buttons at the bottom provide

core operational controls, with their orange color ensuring high visibility. We clearly defined their functions, “Again” for haptic recognition of both players’ cards again, and “Skip” for skipping the current round, thus supporting the user’s cognitive rhythm and giving them clear operational autonomy. Emotional design remains indispensable at this stage. The consistently displayed “Old Friends” branding and personalized welcome messages, such as “Hello, Mr. Huang,” are not mere decorations. They are thoughtfully designed features to reinforce a sense of familiarity and social presence. This consistent emotional framework is crucial for aligning the core gameplay with the overall research objectives of “emotional support” and “social interaction.”

Figure 4. Main gameplay interface of the “Old Friends” serious card game.



The core design of the victory and defeat interfaces (Figures 5 and 6) is to provide emotionally intelligent feedback for older users. Both interfaces strictly adhere to the design language of the main interface, using a low-visual-load gray background and high-contrast text to ensure absolute clarity of information. The key design difference lies in the refined tuning of emotional semantics. The victory interface uses strongly affirmative language, such as “YOU ARE THE BEST!” to maximize the user’s sense of accomplishment and self-efficacy, while the defeat interface uses encouraging language, such as “The next one will be better!” to buffer frustration and guide the outcome toward a positive and constructive direction. This differentiated copywriting strategy is a conscious design choice to achieve

the goal of “emotional support.” At the interaction level, the “Again” button is placed in the exact same position on both interfaces, following the principle of consistency. This allows users to quickly start a new game without having to search for it again, regardless of the outcome, reducing cognitive load and operational hesitation. The brand logo and personalized greeting (“Hi, Huang”) at the top serve as a consistent emotional thread, continuously reinforcing the game’s social sense of belonging. Overall, these 2 interfaces are not only the end point of the game process but also key nodes that connect the beginning and the end, maintaining the user’s emotional investment and continued participation.

Figure 5. Victory outcome interface of the “Old Friends” serious card game.

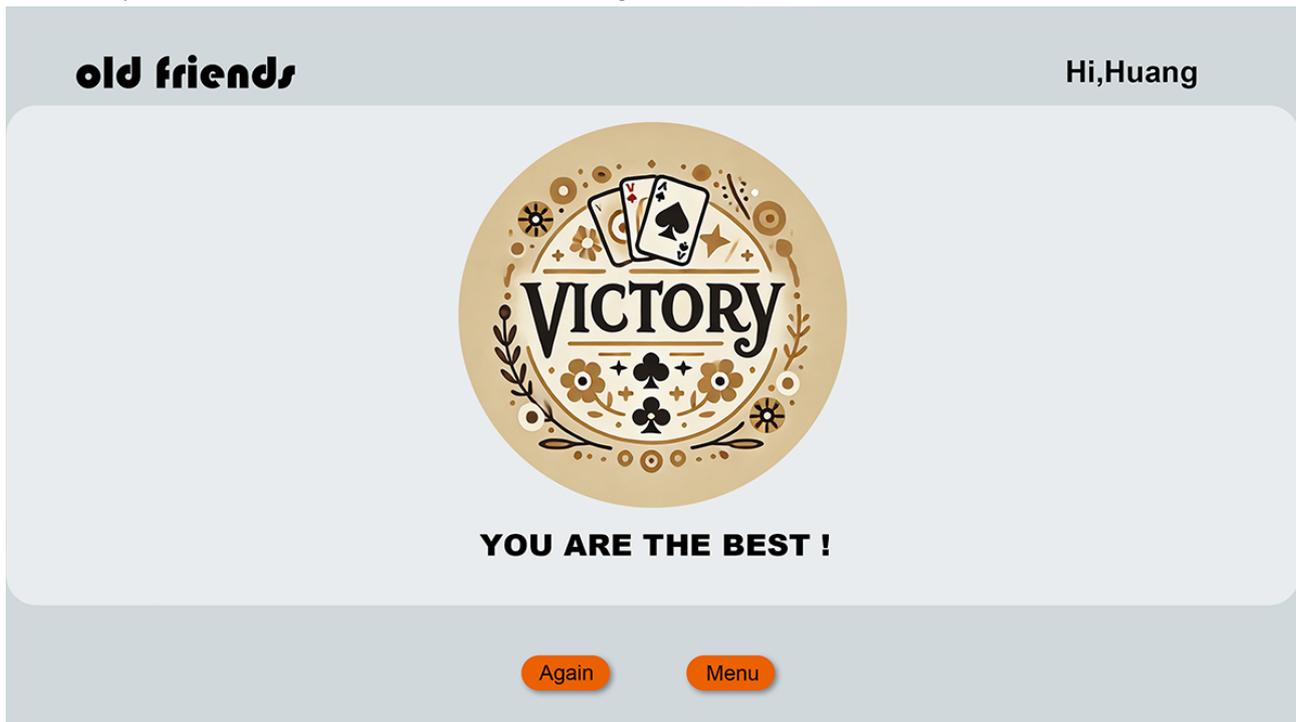
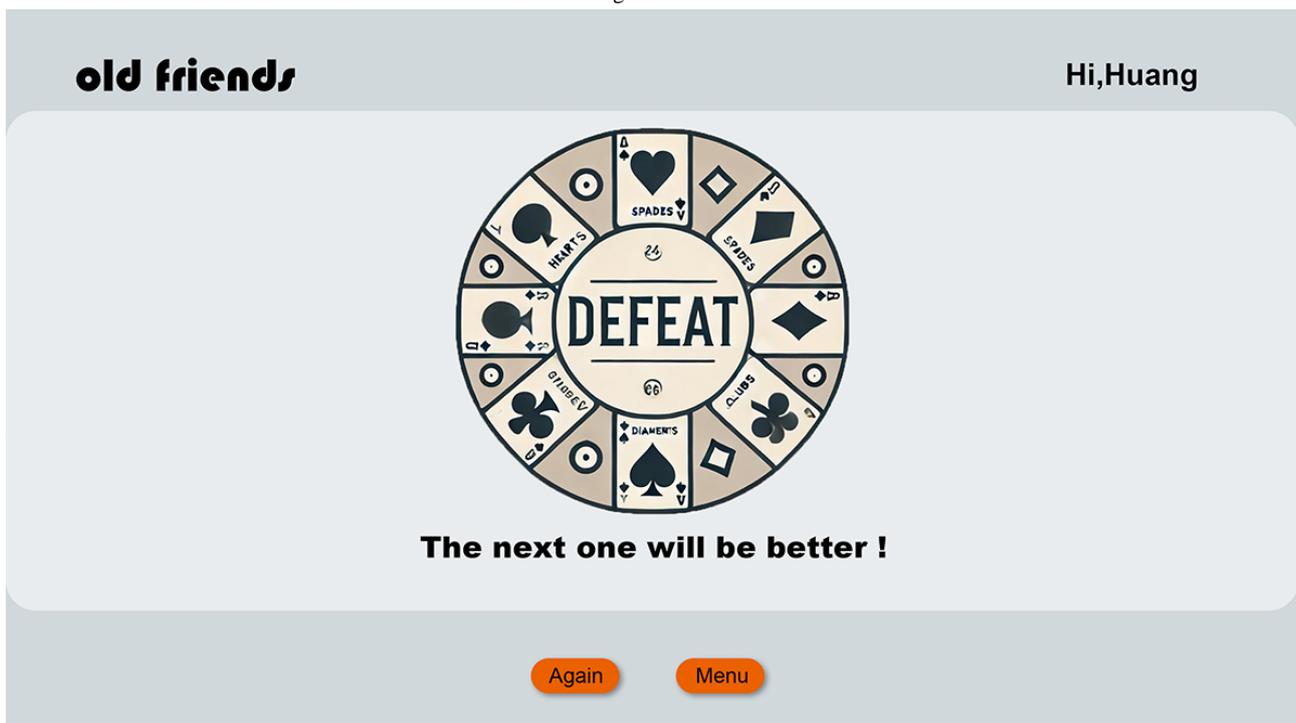


Figure 6. Failure outcome interface of the “Old Friends” serious card game.



The “Homeland” interface (Figure 7), serving as the core social and progress hub of the game, is designed to transform the abstract sense of community belonging into a concrete visual expression. The interface continues the game’s overall principle of low visual load, using soft backgrounds and high contrast numbers to ensure clear readability. Its core design feature is a series of circular blocks labeled with numbers. We chose circles because their smooth, rounded shape conveys a sense of inclusivity and harmony, perfectly aligning with the “Homeland” theme. The numbers within each block represent virtual coins

needed for community building. This quantified display makes the abstract goal of “community building” concrete and controllable, thereby enhancing the clarity of tasks and operational confidence for older users. Overall, this interface is not merely a collection of functions but also a visual metaphor, as breaking down the “community” into interactive and achievable units aims to strengthen older users’ sense of participation and belonging, directly supporting the study’s core objectives of “social interaction” and “emotional support.”

Figure 7. "Homeland" interface within the "Old Friends" serious card game.

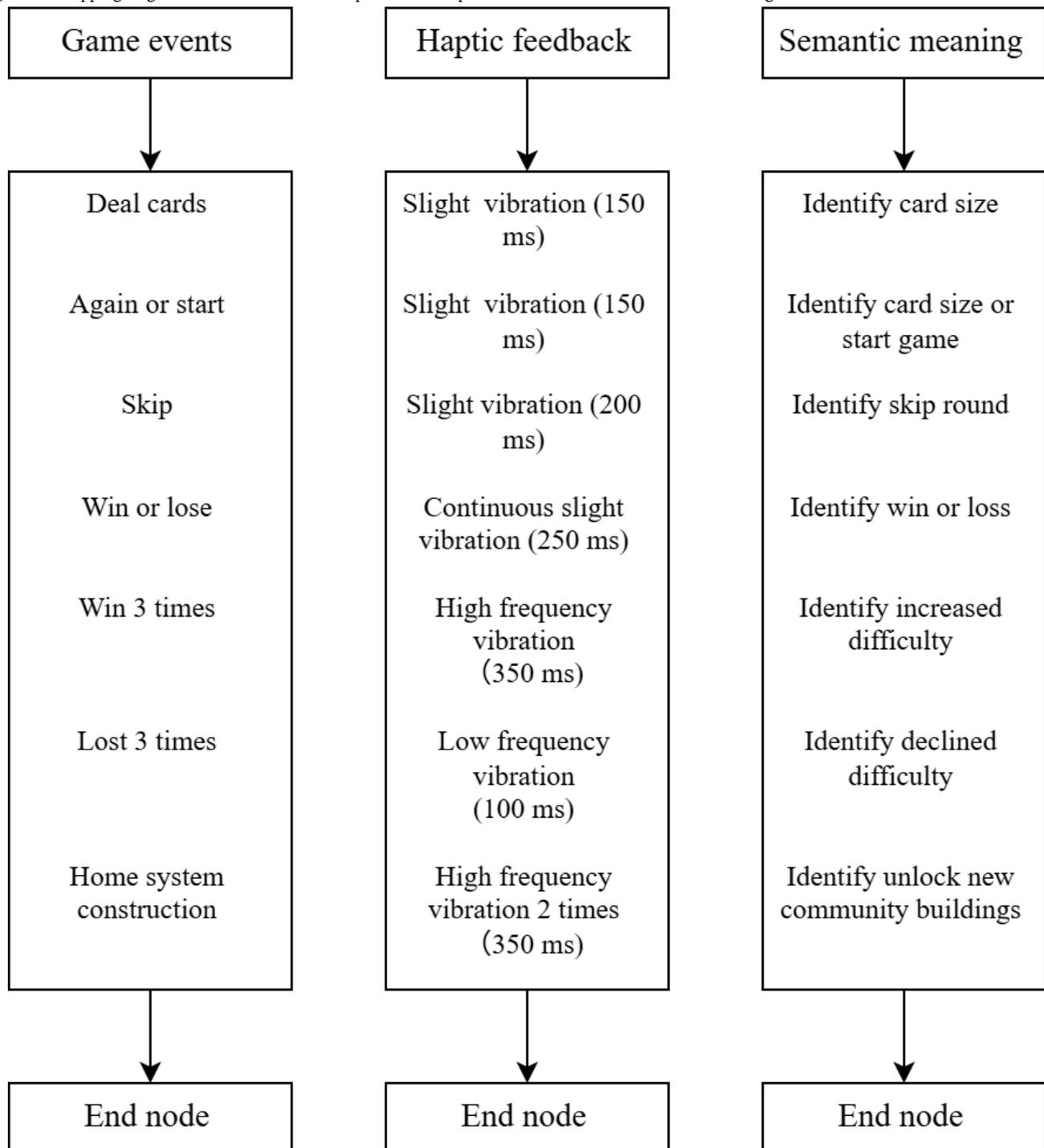


Haptic Interaction Flowchart

To create an accessible and semantically rich gaming experience for older users, we designed a structured haptic feedback system that maps specific vibration patterns to in-game events, as shown

in Figure 8. This design follows a core principle of using noninvasive haptic cues to convey game status information, thereby reducing reliance on vision and providing support for users with visual impairments.

Figure 8. Mapping of game events to structured haptic feedback patterns in the "Old Friends" serious card game.



The mapping relationship is carefully designed to ensure intuitiveness. Slight vibration (150 ms) is used for "Deal cards" and pressing the "Again" and "Start" buttons. Conversely, a continuous slight vibration (250 ms) is dedicated to important game outcomes (win or loss), providing a more significant and emotionally resonant cue. Furthermore, the system dynamically adjusts the game difficulty based on player performance, conveying this through differentiated vibrations, with a high-frequency vibration (350 ms) after 3 consecutive wins indicating an increased challenge, while a low-frequency vibration (100 ms) after 3 consecutive losses indicates a decreased difficulty. Unlocking new community buildings is

achieved with 2 high-frequency vibrations (350 ms each). This closed-loop design aims to maintain an appropriate level of challenge and alleviate player frustration.

In terms of technical implementation, this haptic solution uses an Android linear resonant motor, chosen for its ability to accurately and stably generate predetermined vibration parameters. Therefore, this structured approach ensures that haptic feedback is not merely decorative but serves as an indispensable and meaningful communication channel within the game's accessibility design framework.

Discussion

Overview of Findings in Relation to Objectives

This pilot study successfully achieved the 2 main objectives outlined in the introduction. First, it systematically applied the DPE framework to design a haptic-driven prototype of serious game theory for older adults. Second, it conducted a preliminary assessment of the prototype's usability, accessibility, and user experience in the target population. The results confirmed the feasibility and preliminary acceptability of the approach. Quantitative analysis showed that the prototype achieved an excellent SUS score. In the qualitative analysis, thematic analysis revealed three key user experience themes: (1) the intuitiveness of haptic feedback, (2) its effectiveness in reducing perceived eye strain, and (3) the appeal of simplified game mechanics. These results collectively provide fundamental validation that DPE-guided design with haptic feedback as a core component can create an engaging and visually low-load cognitive training tool for older adults, thus directly achieving the objectives of this study.

Comparisons to Existing Literature

This study addresses a specific gap in the existing literature. While previous studies have successfully integrated haptic technology into serious games, their main focus has often been on enhancing realism, immersion, or motor skill training [17,19]. In contrast, our app is inherently accessibility-driven. We repurpose a common hardware feature of smartphones (vibration motors) to address a key and often overlooked problem of the pervasive visual dependence in tools designed for older populations [14,15]. Furthermore, by adopting the DPE framework, we provide a systematic, theory-based design approach that goes beyond the ad hoc or technology-centric design approaches seen in some previous work [18,20], thus providing a replicable model for future research and development in the field of inclusive aging technologies.

Interpretations

Positive user feedback regarding the intuitiveness of haptic feedback, its reduction of eye strain, and the game's ease of use validated our design approach in several ways. First, users quickly grasped the ability to identify cards through vibration cues, demonstrating that the carefully designed haptic symbol system successfully served as an effective alternative to traditional information channels. This directly confirms the feasibility of "haptic feedback as a substitute for vision" as a practical design principle [17,19]. Second, user feedback on reduced eye strain indicates that the strategy of combining high-contrast visuals with haptic cues successfully alleviated the burden on the visual system [15,32]. Finally, user appreciation for the simplified card mechanics suggests that for our target user group, reducing the cognitive burden associated with complex rules is just as important as sensory adaptation [33], ensuring user engagement and minimizing frustration.

Implications

These findings are significant. Theoretically, they provide concrete empirical support for applying sensory substitution [34] and cognitive load theories [35] to the design of technology

for older adults. Our research shows that information can be strategically redistributed across different sensory channels to create a more accessible and less burdensome user experience for visually impaired older adults. Practically, the results validate the DPE framework as a feasible blueprint for building such designs. The "design" dimension ensures the learnability of haptic symbols, the "play" dimension uses these symbols to implement core mechanisms, such as difficulty adjustment, and the "experience" dimension uses them to provide emotional feedback. This provides developers with a clear path to create serious games that are not only cognitively stimulating but also inherently accessible, going beyond the practice of adding accessibility features after the fact.

Limitations

As a pilot study, this research has several limitations that must be explained to better understand the findings and guide future research. First, method limitations are significant, as the small sample size ($n=10$) and recruitment from a single old-age university solely through convenience sampling limit statistical power and the general applicability of the results. Second, reliance on self-reported measurements (eg, perceptions of eye fatigue during interviews) and the lack of objective physiological or cognitive indicators mean that while the reported benefits are encouraging, they remain preliminary and subjective. Third, the nature of the intervention itself also presents limitations, as a cross-sectional study design cannot assess long-term engagement, cognitive maintenance effects, or changes in the potential haptic learning curve over time. Furthermore, the technical parameters of the haptic feedback (eg, vibration pattern and duration) are based on the initial design choices and require systematic optimization.

To overcome these specific limitations, future research should prioritize the following directions: (1) conducting large-scale randomized controlled trials, including diverse populations with clinical characteristics, to verify cognitive effects using statistical methods and incorporate objective indicators; (2) using longitudinal study designs to assess users' sustained engagement, long-term adherence, and the trajectory of cognitive outcomes; (3) systematically testing haptic parameters to establish evidence-based intuitive and emotionally resonant guidelines for older adults; and (4) expanding the prototype by incorporating richer social interaction mechanisms and adaptive narratives to further enhance long-term motivation. Ultimately, transforming this framework into standardized design guidelines will be key to advancing the field of accessible cognitive training.

Conclusion

This pilot study validated a prototype of a haptic-driven serious game for older adults. The main innovation of this study lies in the systematic application of the DPE framework, putting "haptic as a substitute for vision" as a core design principle into practice. This approach fundamentally differs from existing studies, which largely use haptics to enhance overall immersion rather than addressing specific accessibility challenges posed by visual impairments. The main contribution of this study to the field is providing a reusable and theoretically sound design blueprint, offering a structured approach to creating efficient

and easy-to-use cognitive training tools. Its implications in the real world, the prototype demonstrates a practical and scalable model for deploying low-visual-load interventions in communities and care facilities. These findings have broader implications, as they point to a more inclusive digital health

future where therapeutic tools can be designed with sensory impairments in mind from the outset. By reducing barriers to participation, such approaches promise to promote social inclusion and support cognitive health among older adults, thereby meeting an important public health need.

Acknowledgments

We would like to thank the participants from the Dangtu Old Age University. The authors used DeepSeek and Google Translate to perform a grammar check on this article.

Funding

Funding was partly supported by a research grant from Universiti Kebangsaan Malaysia (grant GUP-2019-066).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: XH (lead), NMA (equal)

Data curation: XH (lead), NMA (equal)

Formal analysis: XH (lead), NMA (equal), YZ (equal)

Funding acquisition: NMA

Methodology: XH (lead), NMA (equal)

Resources: XH

Supervision: NMA (lead), SS (equal)

Validation: XH (lead), NMA (equal), SS (equal)

Visualization: XH

Writing—original draft: XH

Writing—review & editing: XH (lead), NMA (equal), LWL (equal)

Conflicts of Interest

None declared.

References

1. Ageing and health. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health> [accessed 2025-11-15]
2. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 1999;56(3):303-308. [doi: [10.1001/archneur.56.3.303](https://doi.org/10.1001/archneur.56.3.303)] [Medline: [10190820](https://pubmed.ncbi.nlm.nih.gov/10190820/)]
3. Larson EB, Yaffe K, Langa KM. New insights into the dementia epidemic. *N Engl J Med* 2013;369(24):2275-2277 [FREE Full text] [doi: [10.1056/NEJMp1311405](https://doi.org/10.1056/NEJMp1311405)] [Medline: [24283198](https://pubmed.ncbi.nlm.nih.gov/24283198/)]
4. Mouchaileh N, Hughes AJ. Pharmacological management of Parkinson's disease in older people. *Pharmacy Practice and Res* 2020;50(5):445-454. [doi: [10.1002/jppr.1683](https://doi.org/10.1002/jppr.1683)]
5. Del Ser T, Zea M, Valentí M, Olazarán J, López-Álvarez J, Rebollo-Vázquez A, et al. Effects of commonly prescribed drugs on cognition and mild cognitive impairment in healthy elderly people. *J Psychopharmacol* 2019;33(8):965-974. [doi: [10.1177/0269881119857206](https://doi.org/10.1177/0269881119857206)] [Medline: [31241413](https://pubmed.ncbi.nlm.nih.gov/31241413/)]
6. Petersen RC, Bennett D. Mild cognitive impairment: is it Alzheimer's disease or not? *J Alzheimers Dis* 2005;7(3):241-245. [doi: [10.3233/jad-2005-7307](https://doi.org/10.3233/jad-2005-7307)] [Medline: [16006668](https://pubmed.ncbi.nlm.nih.gov/16006668/)]
7. Yu J, Lam CL, Lee TMC. White matter microstructural abnormalities in amnesic mild cognitive impairment: a meta-analysis of whole-brain and ROI-based studies. *Neurosci Biobehav Rev* 2017;83:405-416 [FREE Full text] [doi: [10.1016/j.neubiorev.2017.10.026](https://doi.org/10.1016/j.neubiorev.2017.10.026)] [Medline: [29092777](https://pubmed.ncbi.nlm.nih.gov/29092777/)]
8. Dementia. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/dementia> [accessed 2025-11-04]
9. Umphred DA, Lazaro RT, Roller M, Burton G. *Neurological Rehabilitation*. Amsterdam, Netherlands: Elsevier Health Sciences; 2013.
10. Parkinson disease. Merck Manual Professional Edition. URL: <https://www.merckmanuals.com/professional/neurologic-disorders/movement-and-cerebellar-disorders/parkinson-disease> [accessed 2025-11-15]

11. González-González CS, Toledo-Delgado PA, Muñoz-Cruz V, Torres-Carrion PV. Serious games for rehabilitation: gestural interaction in personalized gamified exercises through a recommender system. *J Biomed Inform* 2019;97:103266 [FREE Full text] [doi: [10.1016/j.jbi.2019.103266](https://doi.org/10.1016/j.jbi.2019.103266)] [Medline: [31408732](https://pubmed.ncbi.nlm.nih.gov/31408732/)]
12. Abt CC. *Serious Games*. London, United Kingdom: Bloomsbury Academic; 1987.
13. Benveniste S, Jouvelot P, Péquignot R. The MINWii project: Renarcissization of patients suffering from Alzheimer's disease through video game-based music therapy. In: Yang HS, Malaka R, Hoshino J, Han JH, editors. *Entertainment Computing - ICEC*. Berlin, Heidelberg: Springer; 2010.
14. Huang X, Ali NM, Sahrani S. Evolution and future of serious game technology for older adults. *Information* 2024;15(7):385. [doi: [10.3390/info15070385](https://doi.org/10.3390/info15070385)]
15. Farage MA, Miller KW, Ajayi F, Hutchins D. Design principles to accommodate older adults. *Glob J Health Sci* 2012;4(2):2-25 [FREE Full text] [doi: [10.5539/gjhs.v4n2p2](https://doi.org/10.5539/gjhs.v4n2p2)] [Medline: [22980147](https://pubmed.ncbi.nlm.nih.gov/22980147/)]
16. Konstantinidis EI, Bamparopoulos G, Bamidis PD. Moving real exergaming engines on the web: the webFitForAll case study in an active and healthy ageing living lab environment. *IEEE J Biomed Health Inform* 2017;21(3):859-866. [doi: [10.1109/JBHI.2016.2559787](https://doi.org/10.1109/JBHI.2016.2559787)] [Medline: [28113566](https://pubmed.ncbi.nlm.nih.gov/28113566/)]
17. Deng S, Chang J, Zhang JJ. A survey of haptics in serious gaming. In: De Gloria A, editor. *Games and Learning Alliance*. Cham: Springer International Publishing; 2014:130-144.
18. LI Y, Yoo Y, Weill-Duflos A, Cooperstock J. Towards context-aware automatic haptic effect generation for home theatre environments. In: *VRST '21: Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*. 2021 Presented at: Association for Computing Machinery; December 8 - 10, 2021; Osaka, Japan p. 1-11. [doi: [10.1145/3489849.3489887](https://doi.org/10.1145/3489849.3489887)]
19. Pacheco-Barrios K, Ortega-Márquez J, Fregni F. Haptic technology: exploring its underexplored clinical applications-a systematic review. *Biomedicines* 2024;12(12):2802 [FREE Full text] [doi: [10.3390/biomedicines12122802](https://doi.org/10.3390/biomedicines12122802)] [Medline: [39767709](https://pubmed.ncbi.nlm.nih.gov/39767709/)]
20. Deng S, Kirkby JA, Chang J, Zhang JJ. Multimodality with eye tracking and haptics: a new horizon for serious games? *Int J Serious Games* 2014;1(4):17-34. [doi: [10.17083/ijsg.v1i4.24](https://doi.org/10.17083/ijsg.v1i4.24)]
21. Silva AJR, Restivo T, Gabriel J. A serious game with a thermal haptic mouse. *Int J Onl Eng* 2013;9(S8):74-76. [doi: [10.3991/ijoe.v9is8.3399](https://doi.org/10.3991/ijoe.v9is8.3399)]
22. Ferdig RE. *Handbook of Research on Effective Electronic Gaming in Education (3 Volumes)*. IGI Global Scientific Publishing. 2009. URL: <https://www.igi-global.com/book/handbook-research-effective-electronic-gaming/459> [accessed 2025-11-15]
23. Golzar J, Noor S, Tajik O. Convenience sampling. *Int J Educ Lang Stud* 2022;1(2):72-77. [doi: [10.4135/9781071812082.n116](https://doi.org/10.4135/9781071812082.n116)]
24. Su Z, Li Y, Xie Y, Huang Z, Cheng A, Zhou X, et al. Acute and long COVID-19 symptoms and associated factors in the omicron-dominant period: a nationwide survey via the online platform Wenjuanxing in China. *BMC Public Health* 2024 Aug 01;24(1):2086 [FREE Full text] [doi: [10.1186/s12889-024-19510-w](https://doi.org/10.1186/s12889-024-19510-w)] [Medline: [39090598](https://pubmed.ncbi.nlm.nih.gov/39090598/)]
25. Brooke J. *SUS: a 'Quick and Dirty' usability scale*. In: *Usability Evaluation In Industry*. Boca Raton, FL: CRC Press; 1996.
26. Adeoye - Olatunde OA, Olenik NL. Research and scholarly methods: semi - structured interviews. *J Am Coll Clin Pharm* 2021;4(10):1358-1367. [doi: [10.1002/jac5.1441](https://doi.org/10.1002/jac5.1441)]
27. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Hum Comput Interact* 2008;24(6):574-594. [doi: [10.1080/10447310802205776](https://doi.org/10.1080/10447310802205776)]
28. Sampaio A, Lewis JR. *Quantifying the User Experience: Practical Statistics for User Research*. Amsterdam Waltham, MA: Elsevier/Morgan Kaufmann; 2013.
29. Appelbaum M, Cooper H, Kline RB, Mayo-Wilson E, Nezu AM, Rao SM. Journal article reporting standards for quantitative research in psychology: the APA publications and communications board task force report. *Am Psychol* 2018;73(1):3-25. [doi: [10.1037/amp0000191](https://doi.org/10.1037/amp0000191)] [Medline: [29345484](https://pubmed.ncbi.nlm.nih.gov/29345484/)]
30. Willig C, Rogers WS. *The SAGE Handbook of Qualitative Research in Psychology*. Thousand Oaks, CA: SAGE; 2017.
31. Luo R, Wang J, Wang Y. Undergraduate students' perceptions of using videoconferencing for EFL learning: evidence from tencent meeting application. *Heliyon* 2023;9(12):e22993 [FREE Full text] [doi: [10.1016/j.heliyon.2023.e22993](https://doi.org/10.1016/j.heliyon.2023.e22993)] [Medline: [38144280](https://pubmed.ncbi.nlm.nih.gov/38144280/)]
32. *Web Content Accessibility Guidelines (WCAG) 2.1*. URL: <https://www.w3.org/TR/WCAG21/> [accessed 2025-11-04]
33. Bilau I, Koo B, Fu E, Chau W, Kwon E, Yang E. Visual accessibility through open shelving: impacts on cognitive load, motivation, physical activity, and user perception in older adults with mild cognitive impairment. *J Aging Environ* 2025 Nov 12:1-28. [doi: [10.1080/26892618.2025.2576213](https://doi.org/10.1080/26892618.2025.2576213)]
34. Bach-y-Rita P, Kercel SW. Sensory substitution and the human-machine interface. *Trends Cogn Sci* 2003;7(12):541-546. [doi: [10.1016/j.tics.2003.10.013](https://doi.org/10.1016/j.tics.2003.10.013)] [Medline: [14643370](https://pubmed.ncbi.nlm.nih.gov/14643370/)]
35. *Cognitive load theory*. In: *Psychology of Learning and Motivation*. Cambridge, MA: Academic Press; 2011:37-76.

Abbreviations

DPE: Design-Play-Experience

JARS: Journal Article Reporting Standards

SUS: System Usability Scale

Edited by S Brini; submitted 22.Oct.2025; peer-reviewed by W Chuangchai, Z Amin; comments to author 12.Nov.2025; accepted 04.Dec.2025; published 30.Jan.2026.

Please cite as:

Huang X, Mohamad Ali N, Sahrani S, Zhang Y

A Haptic-Driven Serious Game for Cognitive Stimulation and Visual Impairment Mitigation in Older Adults Based on the Design-Play-Experience Framework: Cross-Sectional Mixed Methods Pilot Study

JMIR Serious Games 2026;14:e86290

URL: <https://games.jmir.org/2026/1/e86290>

doi: [10.2196/86290](https://doi.org/10.2196/86290)

PMID:

©Xin Huang, Nazlena Mohamad Ali, Shafrida Sahrani, Yue Zhang. Originally published in JMIR Serious Games (<https://games.jmir.org>), 30.Jan.2026. 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.

Associations Between Parental Gaming Behaviors and Conversion From Internet Gaming Disorder Noncases to Cases Among Adolescents: Prospective Longitudinal Cohort Study

Joseph TF Lau^{1,2}, PhD; Yanqiu Yu³, PhD

¹Zhejiang Provincial Clinical Research Center for Mental Disorders, The Affiliated Wenzhou Kangning Hospital, Wenzhou Medical University, Wenzhou, China

²Public Mental Health Center, School of Mental Health, Wenzhou Medical University, Wenzhou, China

³School of Public Health, Fudan University, 130 Dong'an Road, Shanghai, China

Corresponding Author:

Yanqiu Yu, PhD

School of Public Health, Fudan University, 130 Dong'an Road, Shanghai, China

Abstract

Background: Parental factors are known determinants of internet gaming disorder (IGD) among adolescents. However, the associations between gaming-specific parental factors (eg, parental gaming frequency and parental invitations for cogaming) and IGD have been less investigated, and relevant longitudinal evidence is lacking to inform effective IGD interventions.

Objective: This study aimed to investigate (1) the prevalence of IGD conversion (from a noncase at baseline to a case at follow-up) and (2) the prospective associations between two parental gaming behaviors and IGD conversion as well as their mediation mechanisms via parental supportive attitude toward adolescents' gaming behaviors and behavioral intention of increasing gaming time.

Methods: A 12-month prospective longitudinal study, with the baseline survey in December 2018 (T1) and the follow-up survey in December 2019 (T2), was conducted among students from 6 convenience-selection junior middle schools in Chengdu and Guangzhou, China. All grade 7 and 8 students of these schools were invited for participation; students self-administered the structured questionnaire on paper and pencil, in classroom settings, and in the absence of schoolteachers. The 9-item DSM-5 Internet Gaming Disorder Symptoms Checklist was used to assess IGD (Cronbach $\alpha=0.73$). Those with IGD at T1 were excluded, and the final sample size was 2172 (mean age 12.56, SD 0.02, 95% CI 12.52-12.60 years; n=1102, 50.7%, 95% CI 48.7%-52.8% female). The prevalence of IGD conversion was 5.2% (113/2172; 95% CI 4.4%-6.1%).

Results: Adjusted for background factors and respective mediator and outcome scores at T1, two path analysis models showed satisfactory model fit indices (ie, root mean square error approximation=0.02 and 0.03, comparative fit index=0.97 for both models, and standardized root mean square residual=0.02 for both models). The prospective associations between perceived parental gaming frequency and perceived parental invitations for cogaming and IGD conversion were significantly and fully mediated by the 1-mediator indirect path via parental supportive attitude ($\beta=0.02$, 95% CI 0.01-0.04 and $\beta=0.02$, 95% CI 0.01-0.04) and the 2-mediator indirect path first via parental supportive attitude and then via behavioral intention of increasing gaming time ($\beta=0.006$, 95% CI 0.002-0.009 and $\beta=0.005$, 95% CI 0.001-0.009) but not the 1-mediator indirect path via the behavioral intention ($\beta=0.01$, 95% CI -0.01 to 0.03 and $\beta=0.01$, 95% CI -0.01 to 0.03), respectively.

Conclusions: This longitudinal study revealed the prospective associations between 2 parental gaming behaviors and adolescent IGD conversion, as well as the mediation mechanisms, addressing inconsistencies in previous cross-sectional studies and filling in the knowledge gap in longitudinal studies that overlooked parental gaming behaviors. It suggests that parental gaming behaviors may form important family environments shaping adolescents' perceptions and behaviors related to internet gaming. Family-based prevention and intervention programs on IGD conversion may hence take into account these findings.

(JMIR Serious Games 2026;14:e80061) doi:[10.2196/80061](https://doi.org/10.2196/80061)

KEYWORDS

internet gaming disorder; parental factors; social cognitive theory; theory of planned behavior; adolescent; longitudinal study

Introduction

Internet gaming disorder (IGD) was classified as “a condition for further study” in Section III of the *Diagnostic and Statistical*

Manual of Mental Disorders, Fifth Edition (DSM-5) [1] to encourage research into the etiology and treatment of pathological gaming. The clinical relevance was reinforced by the inclusion of gaming disorder in the *International*

Classification of Diseases, 11th Revision (ICD-11) [2]. Notably, while *ICD-11* focuses on functional impairments, the *DSM-5* operationalizes IGD through 9 cognitive, psychological, and behavioral criteria. This study is grounded in the *DSM-5* framework, focusing on the developmental mechanisms of IGD conversion (ie, the transition from an IGD noncase to a case over time). A meta-analysis reported a pooled global prevalence of IGD of 8.6% among adolescents, with the highest prevalence observed in China (11.7%) [3]. Adolescent IGD was prospectively associated with numerous negative consequences, including depression, anxiety, suicidal ideation, sleep problems, and conduct problems [4-11].

Parental factors are known determinants of adolescent IGD, as parents convey the earliest, strong, and most direct influences on adolescents [12]. This study conceptually categorizes parental factors into 2 domains: general (nonspecific) factors and gaming-specific factors. General factors encompass broad relational constructs, such as parental style, attachment, and parent-child communication [13,14]. In contrast, specific factors refer to parental behaviors directly targeting the child's gaming activities, such as parental supervision and monitoring [15]. A review highlighted 5 gaming-specific parental behaviors, including no (parental) intervention, cogaming, active mediation (eg, communicating with children without criticism), monitoring (eg, checking children's gaming without in-depth communication), and restriction (eg, setting rules) [15]. Notably, the literature has focused more on nonspecific parental factors, and there is a dearth of studies investigating modifiable parental gaming behaviors, of which parental gaming frequency and invitations for cogaming (with children) are important. To our knowledge, only 6 cross-sectional studies were located investigating the associations between parental gaming behaviors and adolescent IGD. Three studies reported positive associations among adolescents in mainland [16], Hong Kong [17], and Taiwan [18] China, while the other 3 studies reported nonsignificant results among adolescents in Malaysia [19] and Korea [20,21]. The findings were hence mixed, reflecting the limitations of cross-sectional designs. For instance, the nonsignificant findings might result from the inability to distinguish between parents who game with their children to foster connection (protector) and those whose gaming creates a family environment of excessive gaming (risk). In addition, cross-sectional designs cannot rule out reverse causality (ie, parents might increase cogaming to monitor a child with IGD). Furthermore, although several studies looked at prospective IGD conversion [22-28], none of them involved parental gaming behaviors as predictors. Thus, longitudinal evidence is required to understand the temporal precedence of these parental gaming behaviors in predicting IGD conversion and related mechanisms.

Parental supportive attitude toward the adolescent's gaming behavior was another parental gaming-specific factor of IGD [15], which may mediate the association between parental gaming behaviors and IGD conversion. According to the theory of cognitive dissonance, parents with gaming behaviors may change their perceptions toward accepting gaming behaviors of themselves and their children to minimize mental discomfort [29]. Hence, parental gaming behaviors may lead to stronger parental supportive attitude toward the adolescent's gaming

behavior (a positive association). In addition, parental supportive attitude might be positively associated with IGD conversion, as it forms a subjective norm supporting adolescent gaming. Subjective norm is a key construct of the theory of planned behavior (TPB) [30], postulating that behavioral attitude, subjective norm, and perceived behavioral control would determine behavioral intention, which would in turn determine the actual behavior (gaming behavior in this case). Parental supportive attitude may hence mediate the associations between parental gaming behaviors and IGD conversion.

Plausibly, behavioral intention of increasing gaming time would further mediate between parental supportive attitude and IGD conversion, as TPB implies that parental supportive attitude would increase adolescents' behavioral intention of increasing gaming time, which may in turn increase IGD. Behavioral intention of increasing gaming time was a significant predictor of adolescent IGD [30], and the reduction in gaming time has been commonly used as an intervention component and/or goal in IGD interventions [31,32]. However, to our knowledge, no study has investigated whether behavioral intention of increasing gaming time predicts IGD conversion. Parental gaming behaviors may also increase adolescents' behavioral intention of increasing gaming time, heightening the risk of IGD conversion. Thus, it is contended that behavioral intention of increasing gaming time would mediate the associations between unfavorable parental gaming behaviors and parental supportive attitude and IGD conversion.

The associations between parental gaming behaviors and parental supportive attitude and IGD conversion were further supported by the widely used social cognitive theory (SCT) [12]. First, the observational learning construct postulates that an individual performs a behavior by observing and imitating his or her role models [12]. Accordingly, adolescents may learn to game more by observing their parents' gaming behaviors, increasing the risk of IGD conversion. Second, according to the reciprocal determinism construct, health behaviors are determined by the interactions among environmental, personal, and behavioral factors [12]. Doubtlessly, parental supportive attitude is a salient feature of the family environment [12], which would affect IGD conversion.

Given the background, this prospective longitudinal study aimed to (1) investigate the prevalence of IGD conversion (ie, from non-IGD cases at baseline [T1] to IGD cases at follow-up [T2]) over a 12-month follow-up period among adolescents in China and (2) examine the mediation effects of parental supportive attitude toward the adolescent's gaming behavior and adolescents' behavioral intention of increasing gaming time on the associations between parental gaming behaviors (ie, perceived parental gaming frequency and perceived parental invitation for cogaming) and IGD conversion. It was hypothesized that (1) perceived parental gaming frequency (T1) would be prospectively and positively associated with parental supportive attitude (T2) that would be positively associated with IGD conversion (T2), (2) perceived parental gaming frequency (T1) would be prospectively and positively associated with adolescents' behavioral intention of increasing gaming time (T2) that would be positively associated with IGD conversion (T2), and (3) perceived parental gaming frequency (T1) would

be prospectively and positively associated with parental supportive attitude (T2) that would be positively associated with adolescents' behavioral intention of increasing gaming time (T2), which would be positively associated with IGD conversion (T2). Similar research hypotheses were generated with the variable of perceived parental gaming frequency replaced by perceived parental invitation for cogaming.

Methods

Participants and Procedures

A 12-month prospective longitudinal cohort study was conducted in Guangzhou and Chengdu, China, which had population sizes of 18.3 and 16.6 million, respectively. The baseline survey and follow-up survey were conducted in December 2018 (T1) and December 2019 (T2), respectively. Four junior middle schools from Guangzhou and 2 from Chengdu were selected by convenience sampling. All grade 7 and 8 (7 and 8 years of formal schooling) students of those schools were invited to join this study. Grade 9 students were not invited as they would leave the school before the end of the 12-month follow-up period. As we were investigating IGD conversion, those with IGD at baseline were removed from data analysis (an exclusion criterion). The structured questionnaire was self-administered anonymously using paper and pencil, in the classroom setting, and in the absence of the schoolteachers. Well-trained fieldworkers briefed the students about the objective, content, logistics, and voluntary nature of the study. The reporting of this study follows the Journal Article Reporting Standards for Studies Using Structural Equation Modeling [33].

Ethical Considerations

The ethics approval of this project was obtained from the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong (SBRE-18 - 430). This study used a parental opt-out procedure, and student participation was strictly voluntary and uncompensated. Due to the anonymous study design, written informed consent was waived; instead, the act of completing and returning the questionnaire indicated informed consent, which was explicitly stated on the cover page of the questionnaire. To facilitate longitudinal data matching while guaranteeing anonymity, a unique matching code was used combining participants' birthdays, the last 4 digits of their parents' mobile phone numbers, and the last 2 letters of their parents' given names. Consequently, no personally identifiable information was collected, and no identifiable information could be identified in both the datasets analyzed and this manuscript. No incentives were provided to participants, parents, and schoolteachers in this study.

Measurements

Background Factors Assessed at T1

The information included sex, age, city, whether participants had moved to the city, parental educational levels (junior middle school or below, senior middle school, or college or above), and perceived family financial situation (very good, good, moderate, poor, or very poor). These variables, except for age, were treated as categorical variables in data analysis.

Perceived Parental Gaming Behaviors Assessed at T1

Perceived parental gaming frequency was assessed by the item, "How often do your parents play internet games (1=never to 4=always; a continuous variable)?" Parental invitation for cogaming was assessed by the item, "How often do your parents invite you to play internet games together (1=never to 4=always; a continuous variable)?" Both items have been applied to Chinese adolescents in extant literature [16-18].

Adolescent IGD Conversion at T2

The 9-item *DSM-5* checklist was used to assess IGD [1]. It recorded the presence of 9 addictive symptoms, including preoccupation, withdrawal, tolerance, inability to control gaming, loss of interest in other activities, psychological and/or social problems, deception, avoidance, and significant loss due to gaming. IGD was defined by the endorsement of 5 or more items in the past 12 months (yes or no responses) [34]. The Chinese version of the *DSM-5* has been validated in adolescents with satisfactory psychometric properties [35]. The Cronbach α of the checklist was 0.73 at T2 in this study. This variable assessed IGD conversion from T1 to T2 and was treated as a binary variable.

Perceived Parental Supportive Attitude Toward the Adolescent's Gaming Behavior at T2

Parental supportive attitude toward the adolescent's gaming behavior was assessed using the item, "To what extent do your parents support your gaming behavior (1=strongly unsupportive to 5=strongly supportive; a continuous variable)?"

Behavioral Intention of Increasing Gaming Time at T2

Behavioral intention of increasing gaming time was assessed by the item, "Do you intend to increase or reduce your gaming time in the next year (5=intend to increase substantially; 4=intend to increase slightly/moderately; 3=no intention to change; 2=intend to reduce slightly/moderately; 1=intention to reduce substantially; a continuous variable)?"

Sample Size Planning

As path analysis was used as the key statistical method, the rule of thumb for its sample size planning requires a minimum of 10 cases per item or an observed variable [36]. Accordingly, this study requires a sample size of at least 140 participants, and the current sample size is adequate.

Statistical Analysis

Attrition analysis was conducted by examining between-group differences in the background variables and 2 parental gaming behaviors at T1 between follow-ups and dropouts. The Little missing completely at random (MCAR) test was conducted to evaluate the pattern of missing data due to dropouts; multiple imputation ($n=20$) was used to address missing data. Descriptive statistics and normality tests (via the Kolmogorov-Smirnov and Shapiro-Wilk tests) were performed for the key studied variables. Spearman correlation coefficients were derived for the correlations between two parental gaming behaviors (T1), parental supportive attitude (T2), behavioral intention of increasing gaming time (T2), and IGD conversion (T2). Path analysis was performed to test the hypothesized mediation

mechanisms, with the weighted least square mean and variance adjusted estimator. The models were adjusted for all studied background factors and respective baseline levels of mediators and outcomes (ie, parental supportive attitude at T1 for parental supportive attitude at T2). Satisfactory model fit indices for the path analysis models included comparative fit index ≥ 0.90 , root mean square error of approximation ≤ 0.08 , and standardized root mean square residual ≤ 0.08 . Path analysis was performed by using Mplus (version 7.0; Muthén & Muthén) and other analyses by SPSS (version 21.0; IBM Corp). Statistical significance was indicated as $P < .05$ and 95% CI not involving zero.

Results

Participant Characteristics

A total of 2436 students completed the baseline questionnaire; 264 (10.8%) with IGD at T1 were removed. Of the remaining 2172 participants, 1721 (79.2%) were matched for their baseline and follow-up questionnaires (Figure 1). Those lost to follow-up were more likely to be a Chengdu student, older, female, not having moved to the city, and having parental educational levels of junior middle school or below; their parents also played internet games and invited them to cogame more frequently (Table 1). The Little MCAR test was statistically significant ($\chi^2_{302}=599.5$; $P < .001$), suggesting that data were not missing completely at random.

Figure 1. Flowchart of the 2-wave longitudinal study on mediation mechanism between parental gaming behaviors and internet gaming disorder (IGD) conversion among adolescents from 2 Chinese cities from December 2018 to December 2019.

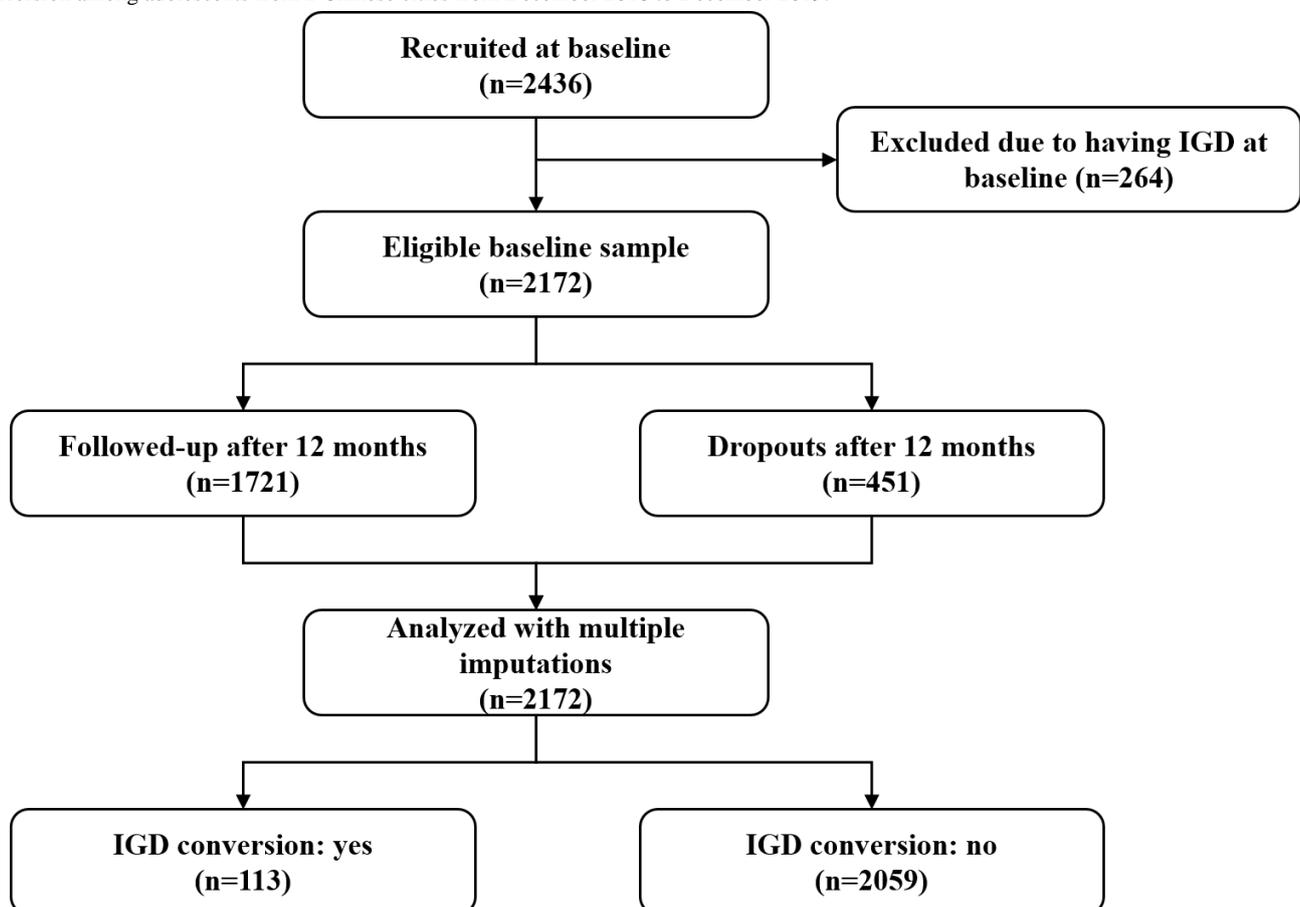


Table . Attrition analysis of the 2-wave longitudinal study on mediation mechanism between parental gaming behaviors and internet gaming disorder (IGD) conversion among 2172 adolescents from 2 Chinese cities from December 2018 to December 2019.

	Followed up (n=1721)	Lost to follow-up (n=451)	<i>P</i> value ^a
Study city, n (%; 95% CI)			<.001
Guangzhou	1100 (63.9; 61.6-66.2)	150 (33.3; 28.9-37.8)	
Chengdu	621 (36.1; 33.8-38.4)	301 (66.7; 62.2-71.1)	
Age (y) ^b , mean (SD; 95% CI)	12.54 (0.7; 12.51-12.58)	12.8 (0.7; 12.7-12.9)	<.001
Sex, n (%; 95% CI)			.002
Female	800 (46.5; 44.1-48.9)	251 (55.7; 51.1-60.1)	
Male	888 (51.6; 49.2-54.0)	195 (43.2; 38.7-47.9)	
Unknown	33 (1.9; 1.4-2.7)	5 (1.1; 0.4-2.7)	
Whether moving to the city, n (%; 95% CI)			<.001
No	288 (17.1; 15.0-18.6)	108 (24.8; 20.1-28.2)	
Yes	1393 (82.9; 79.0-84.6)	328 (75.2; 68.4-76.8)	
Father's educational level, n (%; 95% CI)			<.001
Junior middle school or below	664 (38.6; 36.3-40.9)	229 (50.8; 46.1-55.5)	
Senior middle school	454 (26.4; 24.3-28.5)	129 (28.6; 24.5-33.0)	
College or above	489 (28.4; 26.3-30.6)	58 (12.9; 9.9-16.3)	
Unknown	114 (6.6; 5.5-7.9)	35 (7.8; 5.5-10.6)	
Mother's educational level, n (%; 95% CI)			<.001
Junior middle school or below	678 (39.4; 37.1-41.7)	233 (51.7; 46.9-56.4)	
Senior middle school	475 (27.6; 25.5-29.8)	116 (25.7; 21.7-30.0)	
College or above	456 (26.5; 24.4-28.6)	63 (14; 10.9-17.5)	
Unknown	112 (6.5; 5.4-7.8)	39 (8.6; 6.2-11.6)	
Perceived family financial situation, n (%; 95% CI)			.29
Good or very good	463 (26.9; 24.8-29.1)	109 (24.2; 20.3-28.4)	
Moderate	1055 (61.3; 59.0-63.6)	286 (63.4; 58.8-67.9)	
Poor or very poor	168 (9.8; 8.4-11.3)	51 (11.3; 8.5-14.6)	
Unknown	35 (2; 1.4-2.8)	5 (1.1; 0.4-2.6)	
Perceived parental gaming frequency (range 1 to 4), mean (SD; 95% CI)	1.63 (0.8; 1.60-1.67)	1.8 (0.9; 1.7-2.0)	.003
Perceived parental invitation for cogaming (range 1 to 4), mean (SD; 95% CI)	1.18 (0.5; 1.15-1.20)	1.4 (0.6; 1.3-1.5)	<.001

^aChi-square tests and independent-sample *t* tests were used as applicable.

^bMean and SD are reported for continuous variables. Percentages are calculated within each follow-up group (followed up and lost to follow-up).

Of the 2172 participants based on the 20 imputed dataset, the mean age was 12.56 (SD 0.02, 95% CI 12.52-12.60) years. More than half were female (n=1102, 50.7%, 95% CI 48.7%-52.8%); about two-fifths of their fathers (n=953, 43.9%, 95% CI 41.8%-46.0%) and mothers (n=976, 44.9%, 95% CI 42.8%-47.0%) had received junior middle school or below education. Close to one-tenth self-reported a poor or very poor

family financial situation (n=223, 10.3%, 95% CI 9.0%-11.7%). The prevalence of IGD conversion over a 12-month follow-up period was 5.2% (n=113, 95% CI 4.4%-6.1%; Table 2). The mean scores were 1.66 (SD 0.02; 95% CI 1.63-1.69; range 1 - 4) for perceived parental gaming frequency, 1.21 (SD 0.01, 95% CI 1.19-1.23; range 1 - 4) for perceived parental invitation for cogaming, 2.01 (SD 0.02, 95% CI: 1.96-2.06; range 1 - 5) for

perceived parental supportive attitude, and 1.91 (SD 0.02, 95% CI 1.87-1.95; range 1 - 5) for behavioral intention of increasing gaming time.

Table . Participant characteristics of the 2-wave longitudinal study on mediation mechanism between parental gaming behaviors and internet gaming disorders (IGD) conversion among 2172 adolescents from 2 Chinese cities from December 2018 to December 2019. The frequencies and percentages represent pooled estimates after multiple imputation.

Background factors	Participants, n (%; 95% CI)
Study city	
Guangzhou	1250 (57.6; 55.5-59.6)
Chengdu	922 (42.4; 40.4-44.5)
Sex	
Male	1070 (49.3; 47.2-51.3)
Female	1102 (50.7; 48.7-52.8)
Whether moving to live in the studied city	
No	406 (18.7; 17.2-20.3)
Yes	1766 (81.3; 79.7-82.8)
Father's educational level	
Junior middle school or below	953 (43.9; 41.8-46.0)
Senior high school or equal	627 (28.9; 27.0-30.9)
College or above	592 (27.3; 25.4-29.2)
Mother's educational level	
Junior middle school or below	976 (44.9; 42.8-47.0)
Senior high school or equal	637 (29.3; 27.4-31.3)
College or above	559 (25.7; 23.9-27.7)
Perceived family financial situation	
Good or very good	585 (27; 25.1-28.9)
Moderate	1364 (62.8; 60.8-64.8)
Poor or very poor	223 (10.3; 9.0-11.7)
IGD conversion (T2 ^a)	
No	2059 (94.8; 93.9-95.6)
Yes	113 (5.2; 4.4-6.1)

^aT2: assessed at follow-up.

Correlation Analysis

Kolmogorov-Smirnov and Shapiro-Wilk tests showed that all studied key variables violated the normality assumption (all $P < .001$). Spearman correlation analysis showed that perceived parental gaming frequency, perceived parental invitation for cogaming, parental supportive attitude, and behavioral intention of increasing gaming time were significantly and positively

correlated with IGD conversion. Perceived parental gaming frequency and parental supportive attitude were positively correlated with behavioral intention of increasing gaming time, but the correlation between parental invitation for cogaming and behavioral intention was statistically nonsignificant. The 3 variables related to parental gaming behavior and attitudes were positively correlated with each other (Table 3).

Table . Spearman correlations among the key variables of the 2-wave longitudinal study on mediation mechanisms between parental gaming behaviors and internet gaming disorder (IGD) conversion among 2172 adolescents from 2 Chinese cities from December 2018 to December 2019.

	Parental gaming frequency (T1 ^a), ρ (95% CI)	Parental invitation for cogaming with children (T1), ρ (95% CI)	Parental supportive attitude toward adolescents' gaming behavior (T2 ^b), ρ (95% CI)	Behavioral intention of increasing gaming time (T2), ρ (95% CI)
Parental gaming frequency (T1)	1.00	— ^c	—	—
Parental invitation for cogaming with children (T1)	0.37 (0.34-0.41)	1.00	—	—
Parental supportive attitudes toward the adolescent's gaming behavior (T2)	0.13 (0.08-0.17)	0.07 (0.03-0.11)	1.00	—
Behavioral intention of increasing gaming time (T2)	0.07 (0.03-0.12)	0.04 (0.002-0.09)	0.13 (0.09-0.17)	1.00
IGD conversion (T2)	0.05 (0.01-0.09)	0.05 (0.01-0.09)	0.05 (0.01-0.10)	0.11 (0.07-0.15)

^aT1: assessed at baseline.

^bT2: assessed at follow-up.

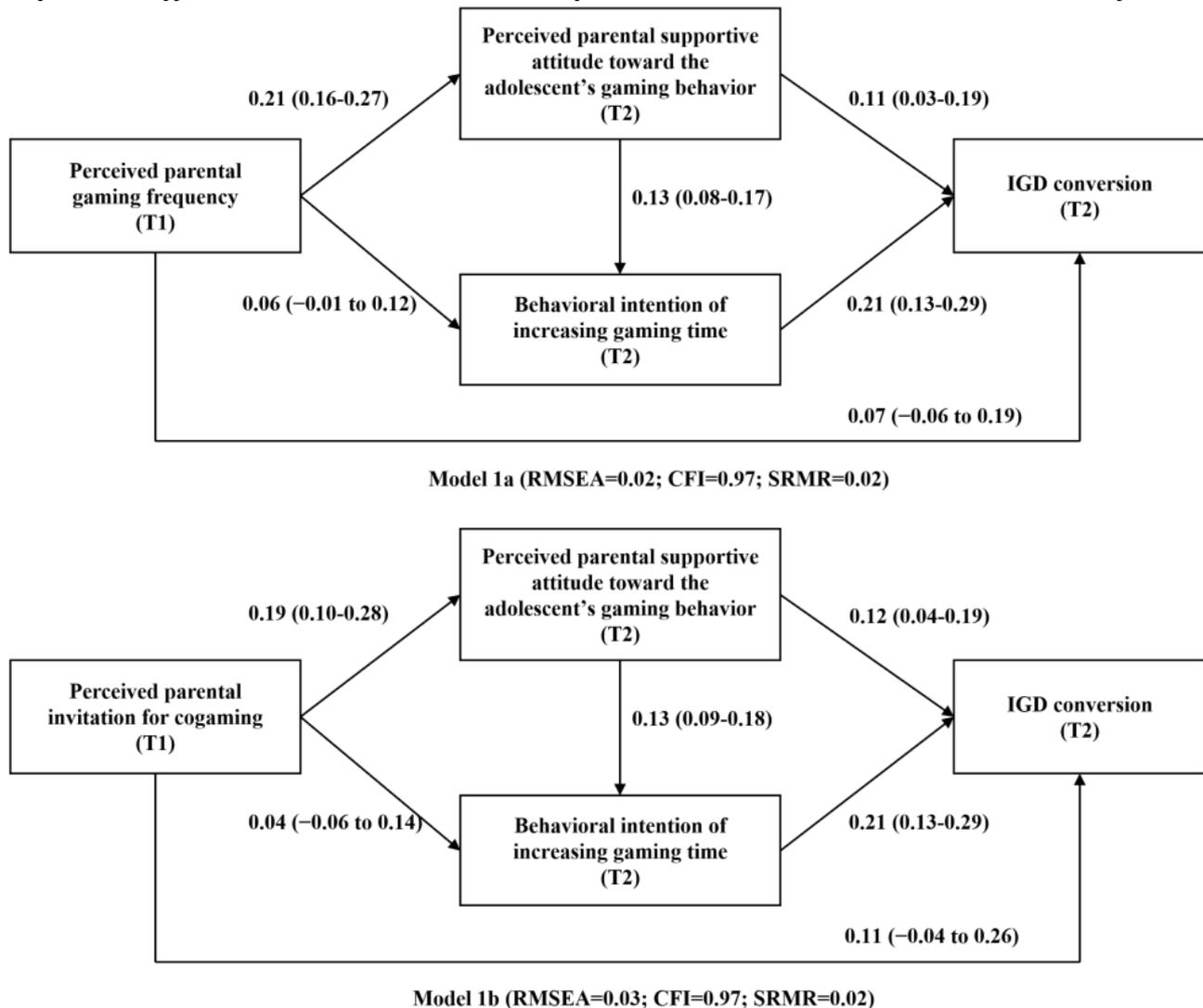
^cNot applicable.

Path Analysis

Figure 2A presents path analysis testing the mediation effects of parental supportive attitude and behavioral intention of increasing gaming time on the association between perceived parental gaming frequency and IGD conversion. The 1-mediator indirect path via parental supportive attitude ($\beta=0.02$, 95% CI 0.01-0.04) and the 2-mediator indirect path (first via parental supportive attitude then via behavioral intention; $\beta=0.006$, 95% CI 0.002-0.009) were statistically significant, whereas the 1-mediator indirect path via behavioral intention ($\beta=0.01$, 95% CI -0.01 to 0.03) was not. As the association between perceived parental gaming frequency and IGD conversion was statistically nonsignificant ($\beta=0.07$, 95% CI -0.06 to 0.19), full mediation via the 2 indirect paths was observed. Similar patterns were

observed for the mediation between perceived parental invitation for cogaming and IGD conversion (Figure 2). The 1-mediator indirect path via parental supportive attitude ($\beta=0.02$, 95% CI 0.01-0.04) and the 2-mediator indirect path (first via parental supportive attitude then via behavioral intention; $\beta=0.005$, 95% CI 0.001-0.009) were statistically significant, whereas the 1-mediator indirect path via behavioral intention of increasing gaming time ($\beta=0.01$, 95% CI -0.01 to 0.03) was not. As the association between perceived parental invitation for cogaming and IGD conversion was also statistically nonsignificant ($\beta=0.11$, 95% CI -0.04 to 0.26), full mediation via the 2 indirect paths was observed. Both models support the original hypothesized models, and the other relevant statistics are presented in Multimedia Appendix 1.

Figure 2. Path analysis models testing the mediation mechanism between parental gaming behaviors and internet gaming disorder (IGD) conversion in the 2-wave longitudinal study among 2172 adolescents from 2 Chinese cities from December 2018 to December 2019. Standardized coefficients were reported as β (95% CI). The models were adjusted for city, age, sex, whether participants had moved to the city, fathers' and mothers' educational level, perceived family financial situation, and respective baseline levels of the mediators and the outcome. CFI: comparative fit index; RMSEA: root mean square error of approximation; SRMR: standardized root mean square residual; T1: assessed at baseline; T2: assessed at follow-up.



Discussion

Principal Findings

This longitudinal study revealed the prospective associations between 2 parental gaming-specific factors of perceived parental gaming frequency and parental invitations for cogaming at T1 and IGD conversion at T2 among Chinese adolescents. Furthermore, the 2 individual path analysis models revealed that these prospective associations were fully mediated by a 2-mediator indirect path (first via parental supportive attitude toward the adolescent's gaming behavior and then via adolescents' behavioral intention of increasing gaming time) and the 1-mediator indirect path via parental supportive attitude, but not the other 1-mediator indirect path via behavioral intention. The results have implications for IGD interventions involving parents.

This study documented the prevalence of IGD conversion of 5.2% (113/2172) among previously non-IGD adolescents in China, which was slightly lower than 11.7% [22] and 1.7% [37]

among participants aged 14 years or older in Western countries. Our results identify 2 parental gaming behaviors positively associated with IGD conversion among Chinese adolescents, highlighting the role of environmental triggers [12]. Specifically, higher frequencies of parental gaming and invitations for cogaming may serve as environmental endorsements of adolescent gaming and facilitate their IGD conversion. Aside from the observational learning perspective of SCT, these parental gaming behaviors can be cues to action, prompting intensive gaming that may increase adolescent IGD conversion [38]. Hence, health education should target not only adolescents but also their parents. Behavioral intention of increasing gaming time was also positively associated with IGD conversion, supporting TPB that behavioral intentions predict corresponding health behaviors [30]. A behavioral intention is often not translated into the actual behavior [39], but this "intention-behavior gap" could be reduced by implementation planning interventions [40]. For instance, the if-then intervention ("if" facing temptation and "then" implementing a mentally

rehearsed preventive response) [40] has demonstrated strong effectiveness in achieving behavioral change goals [41].

In addition, the significant indirect paths shed insights into how to reduce IGD conversion. Both parental gaming behaviors were prospectively associated with parental supportive attitude that was positively associated with behavioral intention of increasing gaming time, which was associated with a higher risk of IGD conversion. The observed full mediation effects suggest that the mediation mechanism via parental supportive attitude alone and first via parental supportive attitude and then via behavioral intention of increasing gaming time may fully explain the associations between parental gaming behaviors at T1 and IGD conversion at T2. Notably, although both indirect paths reached statistical significance, the magnitudes of the involved associations were small. It is important to acknowledge that the statistical significance of these paths may be due to the large sample size of this study [42]. These small effect sizes reflect the multidimensional etiology of IGD conversion [43], wherein parental gaming behaviors and attitudes and personal cognition (behavioral intention) account for only a minute portion of the overall variance. Accordingly, while these variables (eg, parental gaming behaviors) are widespread and modifiable [12], they are better conceptualized as small, interacting components within a broader biopsychosocial network of risk factors, rather than stand-alone targets. This suggests that public health interventions should not rely on targeting single environmental factors but instead adopt holistic, multicomponent approaches that integrate family-level determinants alongside other individual and systemic factors. An additional remark is that the 1-mediator indirect path from parental gaming behaviors to IGD conversion via behavioral intention alone was statistically nonsignificant. This suggests that parental gaming behaviors were prospectively and directly associated with parental supportive attitude, but not behavioral intention, and that the association between parental supportive attitude and IGD conversion was mediated via behavioral intention of increasing gaming time. As TPB postulates that the association between attitude and behavior would be mediated via behavioral intention [30], this finding supports the TPB.

Overall, the earlier findings offer practical implications that prevention of IGD conversion should not neglect both parental and adolescent gaming behaviors and attitudes. Inevitably, parental gaming behaviors (eg, cogaming with children) and parental attitude are a double-edged sword, as they may both increase the risk of adolescent IGD [16-18] and improve parent-child relationships [44]. A delicate balance needs to be reached. Parents may use their gaming behaviors to set good examples in time management, nonexcessive gaming, and self-control for their children. Although excessive cogaming may be problematic, cogaming is a good opportunity for parent-child communication on healthy gaming [15]. Parents should be informed that parental supervision and monitoring are protective factors against adolescent IGD [14] and that parental supportive attitude toward the adolescent's gaming behavior needs to be carefully conveyed to the children to avoid the wrong impression of absent parental supervision and monitoring. Hence, effective family-based IGD interventions

on health education, parenting skills, and improving family relationships [45,46] may consider integrating these findings.

This study has several limitations. First, the attrition rate was 22.6%, and there were significant differences between the retained participants and dropouts. Specifically, those dropouts reported higher frequencies of parental gaming and invitations for cogaming. Furthermore, the MCAR test confirmed that the data were not missing completely at random. To reduce this selection bias and improve the validity of the estimates, multiple imputation was used in data analysis to account for these missing data patterns [47]. Second, convenience sampling may introduce sampling bias [48]. As there might be geographic variations, the generalizability of these findings to other regions and countries should be made with caution. Third, as the questionnaire was self-administered, social desirability bias might be present [48], and several measures were taken to mitigate this risk. The survey was anonymous, and no schoolteachers were present in the classroom during administration; data collection was managed entirely by external research staff to encourage honest reporting. Another notable limitation is the single-item measures to assess parental supportive attitude and behavioral intention of increasing gaming time. While single-item measures can reduce participant burden and are considered valid for concrete constructs [49], they preclude the estimation of internal consistency reliability. As path analysis assumes that observed variables are measured without error [50], the presence of potential random measurement errors in these single items may underestimate path coefficients [51]. Future studies should verify these findings using validated multi-item scales. Furthermore, IGD was assessed by the 9-item *DSM-5* IGD checklist instead of the *ICD-11* gaming disorder criteria, as the former was the prevailing diagnostic standard for large-scale epidemiological studies at the time of baseline data collection (December 2018). Although the clinical validity of the *DSM-5* IGD checklist has been validated in Chinese populations [34], this self-report screening tool may overestimate IGD prevalence compared with a clinical gold standard administered by a psychiatrist [52]. Furthermore, this prospective study used a 2-wave design, and both mediators and IGD conversion were assessed concurrently at T2. Although the baseline variables were adjusted, this design precludes the establishment of a complete temporal relationship (ie, predictor to mediator to outcome). As such, the specific causal relationships between the mediators and IGD conversion cannot be statistically confirmed and should be interpreted as associations within a prospective framework rather than causal mechanisms. Finally, this study did not include data on gaming time or frequency, which are known predictors of IGD conversion [43]; this limits the ability to assess whether the observed associations exist independently of gaming time. Furthermore, this study did not assess peer factors, which are important determinants of adolescent gaming [17]. This limits the ability to compare the relative strength of parental versus peer influences. Future research should integrate peer variables to provide a more comprehensive ecological model of IGD conversion.

Conclusions

In conclusion, this longitudinal study is novel to identify the prospective associations between specific parental gaming behaviors and the risk of IGD conversion among Chinese adolescents. Grounded in SCT and TPB, these findings elucidate the mediation mechanisms involving family environment and individual cognitive processes. Specifically, the identification of parental supportive attitude toward the adolescent's gaming behavior and adolescents' behavioral intention of increasing gaming time as key mediators reveals a nuanced parental behavior–adolescent cognition pathway, suggesting that parental modeling not only provides a behavioral template but also shapes the internal cognitive appraisals and behavioral intentions that precede IGD conversion.

These results bear theoretical and practical implications. Theoretically, they extend the application of SCT and TPB to the field of behavioral addiction by highlighting how distal environmental factors (parental behavior) are internalized through proximal cognitive factors. Practically, the findings underscore the necessity of shifting the focus of IGD prevention from solely adolescent-centered approaches to family-based interventions. Prevention strategies should involve educating parents on their own gaming habits and fostering healthier cognitive appraisals of gaming within the household. Moreover, mitigating the risk of adolescent IGD conversion requires a holistic approach that targets both modifiable parental influences and the subsequent cognitive distortions in adolescents.

Acknowledgments

The authors would like to thank the participants for their contribution to this study.

Funding

This study was funded by Shanghai Clinical Research Fund (Youth Project; 20244Y0052). The funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

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

Authors' Contributions

Conceptualization: YY, JTFL

Data curation: YY

Formal analysis: YY

Investigation: YY

Methodology: YY, JTFL

Resources: JTFL

Software: YY

Supervision: JTFL

Validation: YY

Writing—original draft: YY, JTFL

Writing—review and editing: YY, JTFL

Conflicts of Interest

None declared.

Multimedia Appendix 1

Path analysis testing the mediation mechanism between parental gaming behaviors and IGD conversion in the 2-wave longitudinal study among 2172 adolescents from 2 Chinese cities from December 2018 to December 2019.

[[DOCX File, 23 KB - games_v14i1e80061_app1.docx](#)]

References

1. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Association; 2013. [doi: [10.1176/appi.books.9780890425596](https://doi.org/10.1176/appi.books.9780890425596)]
2. Gaming disorder. World Health Organization. URL: <https://www.who.int/standards/classifications/frequently-asked-questions/gaming-disorder> [accessed 2026-04-21]
3. Satapathy P, Khatib MN, Balaraman AK, et al. Burden of gaming disorder among adolescents: a systemic review and meta-analysis. *Public Health Pract (Oxf)* 2024;9:100565. [doi: [10.1016/j.puhip.2024.100565](https://doi.org/10.1016/j.puhip.2024.100565)] [Medline: [40115446](https://pubmed.ncbi.nlm.nih.gov/40115446/)]

4. Peng P, Chen ZM, Ren SL, et al. Internet gaming disorder and depression mediated by impaired resilience and sleep distress: a three-wave longitudinal study among Chinese adolescents. *Epidemiol Psychiatr Sci* 2025 Feb 19;34:e11. [doi: [10.1017/S2045796025000046](https://doi.org/10.1017/S2045796025000046)] [Medline: [39965936](https://pubmed.ncbi.nlm.nih.gov/39965936/)]
5. Peng P, Jin J, Chen Z, et al. Impaired sleep quality mediates the relationship between internet gaming disorder and conduct problems among adolescents: a three-wave longitudinal study. *Child Adolesc Psychiatry Ment Health* 2025 Mar 21;19(1):26. [doi: [10.1186/s13034-025-00889-2](https://doi.org/10.1186/s13034-025-00889-2)] [Medline: [40119352](https://pubmed.ncbi.nlm.nih.gov/40119352/)]
6. Peng P, Chen Z, Ren S, et al. Internet gaming disorder predicts the incidence, persistence, and worsening of suicidal ideation: a population-based cohort study of 96,158 Chinese adolescents. *J Affect Disord* 2025 Jun 15;379:186-193. [doi: [10.1016/j.jad.2025.03.029](https://doi.org/10.1016/j.jad.2025.03.029)] [Medline: [40081581](https://pubmed.ncbi.nlm.nih.gov/40081581/)]
7. Peng P, Chen Z, Ren S, et al. Trajectory of internet gaming disorder among Chinese adolescents: course, predictors, and long-term mental health outcomes. *J Behav Addict* 2025 Jun;14(2):846-860. [doi: [10.1556/2006.2025.00054](https://doi.org/10.1556/2006.2025.00054)] [Medline: [40549959](https://pubmed.ncbi.nlm.nih.gov/40549959/)]
8. Jeong H, Yim HW, Lee SY, et al. Reciprocal relationship between depression and internet gaming disorder in children: a 12-month follow-up of the iCURE study using cross-lagged path analysis. *J Behav Addict* 2019 Dec 1;8(4):725-732. [doi: [10.1556/2006.8.2019.74](https://doi.org/10.1556/2006.8.2019.74)] [Medline: [32359239](https://pubmed.ncbi.nlm.nih.gov/32359239/)]
9. Mihara S, Higuchi S. Cross-sectional and longitudinal epidemiological studies of internet gaming disorder: a systematic review of the literature. *Psychiatry Clin Neurosci* 2017 Jul;71(7):425-444. [doi: [10.1111/pcn.12532](https://doi.org/10.1111/pcn.12532)] [Medline: [28436212](https://pubmed.ncbi.nlm.nih.gov/28436212/)]
10. Gioia F, Colella GM, Boursier V. Evidence on problematic online gaming and social anxiety over the past ten years: a systematic literature review. *Curr Addict Rep* 2022 Mar;9:32-47. [doi: [10.1007/s40429-021-00406-3](https://doi.org/10.1007/s40429-021-00406-3)]
11. Li S, Wu Z, Zhang Y, Xu M, Wang X, Ma X. Internet gaming disorder and aggression: a meta-analysis of teenagers and young adults. *Front Public Health* 2023;11:1111889. [doi: [10.3389/fpubh.2023.1111889](https://doi.org/10.3389/fpubh.2023.1111889)]
12. Nabavi RT, Bijandi MS. A literature review on Bandura's social learning theory & social cognitive learning theory. University of Science and Culture. 2012. URL: https://www.researchgate.net/publication/267750204_Bandura's_Social_Learning_Theory_Social_Cognitive_Learning_Theory [accessed 2026-04-21]
13. Nielsen P, Favez N, Rigter H. Parental and family factors associated with problematic gaming and problematic internet use in adolescents: a systematic literature review. *Curr Addict Rep* 2020 Sep;7:365-386. [doi: [10.1007/s40429-020-00320-0](https://doi.org/10.1007/s40429-020-00320-0)]
14. Schneider LA, King DL, Delfabbro PH. Family factors in adolescent problematic internet gaming: a systematic review. *J Behav Addict* 2017 Sep 1;6(3):321-333. [doi: [10.1556/2006.6.2017.035](https://doi.org/10.1556/2006.6.2017.035)] [Medline: [28762279](https://pubmed.ncbi.nlm.nih.gov/28762279/)]
15. Nielsen P, Favez N, Liddle H, Rigter H. Linking parental mediation practices to adolescents' problematic online screen use: a systematic literature review. *J Behav Addict* 2019 Dec 1;8(4):649-663. [doi: [10.1556/2006.8.2019.61](https://doi.org/10.1556/2006.8.2019.61)] [Medline: [31786936](https://pubmed.ncbi.nlm.nih.gov/31786936/)]
16. Yu Y, Mo PK, Zhang J, Li J, Lau JT. Impulsivity, self-control, interpersonal influences, and maladaptive cognitions as factors of internet gaming disorder among adolescents in China: cross-sectional mediation study. *J Med Internet Res* 2021 Oct 27;23(10):e26810. [doi: [10.2196/26810](https://doi.org/10.2196/26810)] [Medline: [34704960](https://pubmed.ncbi.nlm.nih.gov/34704960/)]
17. Wang X, Li Q, Wong SY, Yang X. Whether and how parental and friends' internet gaming behaviors may affect adolescent internet gaming disorder symptoms: the cognitive and behavioral mechanisms. *Int J Ment Health Addict* 2023;22(6):3922-3939. [doi: [10.1007/s11469-023-01091-1](https://doi.org/10.1007/s11469-023-01091-1)]
18. Wu JYW, Ko HC, Wong TY, Wu LA, Oei TP. Positive outcome expectancy mediates the relationship between peer influence and internet gaming addiction among adolescents in Taiwan. *Cyberpsychol Behav Soc Netw* 2016 Jan;19(1):49-55. [doi: [10.1089/cyber.2015.0345](https://doi.org/10.1089/cyber.2015.0345)] [Medline: [26716791](https://pubmed.ncbi.nlm.nih.gov/26716791/)]
19. Benrazavi R, Teimouri M, Griffiths MD. Utility of parental mediation model on youth's problematic online gaming. *Int J Ment Health Addict* 2015 Dec;13:712-727. [doi: [10.1007/s11469-015-9561-2](https://doi.org/10.1007/s11469-015-9561-2)]
20. Jeong EJ, Kim DH. Social activities, self-efficacy, game attitudes, and game addiction. *Cyberpsychol Behav Soc Netw* 2011 Apr;14(4):213-221. [doi: [10.1089/cyber.2009.0289](https://doi.org/10.1089/cyber.2009.0289)] [Medline: [21067285](https://pubmed.ncbi.nlm.nih.gov/21067285/)]
21. Lee C, Kim O. Predictors of online game addiction among Korean adolescents. *Addict Res Theory* 2017;25(1):58-66. [doi: [10.1080/16066359.2016.1198474](https://doi.org/10.1080/16066359.2016.1198474)]
22. Wartberg L, Kriston L, Zieglmeier M, Lincoln T, Kammerl R. A longitudinal study on psychosocial causes and consequences of internet gaming disorder in adolescence. *Psychol Med* 2019 Jan;49(2):287-294. [doi: [10.1017/S003329171800082X](https://doi.org/10.1017/S003329171800082X)] [Medline: [29622057](https://pubmed.ncbi.nlm.nih.gov/29622057/)]
23. Liu Y, Gong R, Yu Y, et al. Longitudinal predictors for incidence of internet gaming disorder among adolescents: the roles of time spent on gaming and depressive symptoms. *J Adolesc* 2021 Oct;92:1-9. [doi: [10.1016/j.adolescence.2021.06.008](https://doi.org/10.1016/j.adolescence.2021.06.008)] [Medline: [34246122](https://pubmed.ncbi.nlm.nih.gov/34246122/)]
24. Jeong H, Yim HW, Lee SY, Lee HK, Potenza MN, Lee H. Factors associated with severity, incidence or persistence of internet gaming disorder in children and adolescents: a 2-year longitudinal study. *Addiction* 2021 Jul;116(7):1828-1838. [doi: [10.1111/add.15366](https://doi.org/10.1111/add.15366)] [Medline: [33283397](https://pubmed.ncbi.nlm.nih.gov/33283397/)]
25. She R, Zhang Y, Yang X. Parental factors associated with internet gaming disorder among first-year high school students: longitudinal study. *JMIR Serious Games* 2022 Nov 8;10(4):e33806. [doi: [10.2196/33806](https://doi.org/10.2196/33806)] [Medline: [36346660](https://pubmed.ncbi.nlm.nih.gov/36346660/)]

26. Marrero RJ, Fumero A, Voltés D, González M, Peñate W. Individual and interpersonal factors associated with the incidence, persistence, and remission of internet gaming disorders symptoms in an adolescents sample. *Int J Environ Res Public Health* 2021 Nov 5;18(21):11638. [doi: [10.3390/ijerph182111638](https://doi.org/10.3390/ijerph182111638)] [Medline: [34770150](https://pubmed.ncbi.nlm.nih.gov/34770150/)]
27. Jeong H, Yim HW, Lee SY, Lee HK, Potenza MN, Shin Y. Preschool exposure to online games and internet gaming disorder in adolescents: a cohort study. *Front Pediatr* 2021;9:760348. [doi: [10.3389/fped.2021.760348](https://doi.org/10.3389/fped.2021.760348)] [Medline: [34888270](https://pubmed.ncbi.nlm.nih.gov/34888270/)]
28. Schettler LM, Thomasius R, Paschke K. Emotional dysregulation predicts problematic gaming in children and youths: a cross-sectional and longitudinal approach. *Eur Child Adolesc Psychiatry* 2024 Feb;33(2):605-616. [doi: [10.1007/s00787-023-02184-x](https://doi.org/10.1007/s00787-023-02184-x)] [Medline: [36932230](https://pubmed.ncbi.nlm.nih.gov/36932230/)]
29. Harmon-Jones E, Mills J. An introduction to cognitive dissonance theory and an overview of current perspectives on the theory. In: Harmon-Jones E, editor. *Cognitive Dissonance: Reexamining a Pivotal Theory in Psychology*, 2nd edition: American Psychological Association; 2019:3-24. [doi: [10.1037/0000135-000](https://doi.org/10.1037/0000135-000)]
30. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process* 1991 Dec;50(2):179-211. [doi: [10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)]
31. Zajac K, Ginley MK, Chang R. Treatments of internet gaming disorder: a systematic review of the evidence. *Expert Rev Neurother* 2020 Jan;20(1):85-93. [doi: [10.1080/14737175.2020.1671824](https://doi.org/10.1080/14737175.2020.1671824)] [Medline: [31544539](https://pubmed.ncbi.nlm.nih.gov/31544539/)]
32. Brailovskaia J, Meier-Faust J, Schillack H, Margraf J. A two-week gaming abstinence reduces internet gaming disorder and improves mental health: An experimental longitudinal intervention study. *Comput Human Behav* 2022 Sep;134:107334. [doi: [10.1016/j.chb.2022.107334](https://doi.org/10.1016/j.chb.2022.107334)]
33. JARS–Quant: table 7: reporting standards for studies using structural equation modeling. American Psychological Association. URL: <https://apastyle.apa.org/jars/quant-table-7.pdf> [accessed 2026-04-21]
34. Ko CH, Yen JY, Chen SH, Wang PW, Chen CS, Yen CF. Evaluation of the diagnostic criteria of internet gaming disorder in the DSM-5 among young adults in Taiwan. *J Psychiatr Res* 2014 Jun;53:103-110. [doi: [10.1016/j.jpsychires.2014.02.008](https://doi.org/10.1016/j.jpsychires.2014.02.008)] [Medline: [24581573](https://pubmed.ncbi.nlm.nih.gov/24581573/)]
35. Yang X, Jiang X, Wu AMS, et al. Validation of the internet gaming disorder symptoms checklist based on the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders in Chinese adolescents. *Child Psychiatry Hum Dev* 2023 Feb;54(1):26-33. [doi: [10.1007/s10578-021-01213-7](https://doi.org/10.1007/s10578-021-01213-7)] [Medline: [34338962](https://pubmed.ncbi.nlm.nih.gov/34338962/)]
36. Nunnally JC. An overview of psychological measurement. In: Wolman BB, editor. *Clinical Diagnosis of Mental Disorders*: Springer; 1978:97-146. [doi: [10.1007/978-1-4684-2490-4_4](https://doi.org/10.1007/978-1-4684-2490-4_4)]
37. Scharkow M, Festl R, Quandt T. Longitudinal patterns of problematic computer game use among adolescents and adults--a 2-year panel study. *Addiction* 2014 Nov;109(11):1910-1917. [doi: [10.1111/add.12662](https://doi.org/10.1111/add.12662)] [Medline: [24938480](https://pubmed.ncbi.nlm.nih.gov/24938480/)]
38. Janz NK, Becker MH. The health belief model: a decade later. *Health Educ Q* 1984;11(1):1-47. [doi: [10.1177/109019818401100101](https://doi.org/10.1177/109019818401100101)] [Medline: [6392204](https://pubmed.ncbi.nlm.nih.gov/6392204/)]
39. Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. *Psychol Bull* 2006 Mar;132(2):249-268. [doi: [10.1037/0033-2909.132.2.249](https://doi.org/10.1037/0033-2909.132.2.249)] [Medline: [16536643](https://pubmed.ncbi.nlm.nih.gov/16536643/)]
40. Rhodes RE, Grant S, de Bruijn GJ. Planning and implementation intention interventions. In: Hagger MS, Cameron LD, Hamilton K, Hankonen N, Lintunen T, editors. *The Handbook of Behavior Change*: Cambridge University Press; 2020:572-585. [doi: [10.1017/9781108677318.039](https://doi.org/10.1017/9781108677318.039)]
41. Gollwitzer PM, Sheeran P. Implementation intentions and goal achievement: a meta - analysis of effects and processes. *Adv Exp Soc Psychol* 2006;38:69-119. [doi: [10.1016/S0065-2601\(06\)38002-1](https://doi.org/10.1016/S0065-2601(06)38002-1)]
42. Götz FM, Gosling SD, Rentfrow PJ. Small effects: the indispensable foundation for a cumulative psychological science. *Perspect Psychol Sci* 2022 Jan;17(1):205-215. [doi: [10.1177/1745691620984483](https://doi.org/10.1177/1745691620984483)] [Medline: [34213378](https://pubmed.ncbi.nlm.nih.gov/34213378/)]
43. Zhuang X, Zhang Y, Tang X, Ng TK, Lin J, Yang X. Longitudinal modifiable risk and protective factors of internet gaming disorder: a systematic review and meta-analysis. *J Behav Addict* 2023;12(2):375-392. [doi: [10.1556/2006.2023.00017](https://doi.org/10.1556/2006.2023.00017)] [Medline: [37224007](https://pubmed.ncbi.nlm.nih.gov/37224007/)]
44. Musick G, Freeman G, McNeese NJ. Gaming as family time: digital game co-play in modern parent-child relationships. *Proc ACM Hum Comput Interact* 2021;5:1-25. [doi: [10.1145/3474678](https://doi.org/10.1145/3474678)]
45. Pornnoppadol C, Ratta-apha W, Chanpen S, et al. A comparative study of psychosocial interventions for internet gaming disorder among adolescents aged 13–17 years. *Int J Ment Health Addiction* 2020 Aug;18(4):932-948. [doi: [10.1007/s11469-018-9995-4](https://doi.org/10.1007/s11469-018-9995-4)]
46. González-Bueso V, Santamaría JJ, Fernández D, et al. Internet gaming disorder in adolescents: personality, psychopathology and evaluation of a psychological intervention combined with parent psychoeducation. *Front Psychol* 2018;9:787. [doi: [10.3389/fpsyg.2018.00787](https://doi.org/10.3389/fpsyg.2018.00787)] [Medline: [29892241](https://pubmed.ncbi.nlm.nih.gov/29892241/)]
47. Enders CK. *Applied Missing Data Analysis*, 2nd edition: Guilford Publications; 2022.
48. Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ. *Modern Epidemiology*: Wolters Kluwer; 2021.
49. Fisher GG, Matthews RA, Gibbons AM. Developing and investigating the use of single-item measures in organizational research. *J Occup Health Psychol* 2016 Jan;21(1):3-23. [doi: [10.1037/a0039139](https://doi.org/10.1037/a0039139)] [Medline: [25894198](https://pubmed.ncbi.nlm.nih.gov/25894198/)]
50. Kline RB. *Principles and Practice of Structural Equation Modeling*, 4th edition: The Guilford Press; 2016.
51. Allen MS, Iliescu D, Greiff S. Single item measures in psychological science. *Eur J Psychol Assess* 2022 Jan;38(1):1-5. [doi: [10.1027/1015-5759/a000699](https://doi.org/10.1027/1015-5759/a000699)]

52. Jo YS, Bhang SY, Choi JS, Lee HK, Lee SY, Kweon YS. Clinical characteristics of diagnosis for internet gaming disorder: comparison of DSM-5 IGD and ICD-11 GD diagnosis. *J Clin Med* 2019 Jun 28;8(7):945. [doi: [10.3390/jcm8070945](https://doi.org/10.3390/jcm8070945)] [Medline: [31261841](https://pubmed.ncbi.nlm.nih.gov/31261841/)]

Abbreviations:

DSM-5: *Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)*

ICD-11: *International Classification of Diseases, 11th Revision*

IGD: internet gaming disorder

MCAR: missing completely at random

SCT: social cognitive theory

TPB: theory of planned behavior

Edited by S Brini; submitted 03.Jul.2025; peer-reviewed by L Gao, ME Keshky, P Peng; revised version received 25.Mar.2026; accepted 31.Mar.2026; published 30.Apr.2026.

Please cite as:

Lau JTF, Yu Y

Associations Between Parental Gaming Behaviors and Conversion From Internet Gaming Disorder Noncases to Cases Among Adolescents: Prospective Longitudinal Cohort Study

JMIR Serious Games 2026;14:e80061

URL: <https://games.jmir.org/2026/1/e80061>

doi: [10.2196/80061](https://doi.org/10.2196/80061)

© Joseph TF Lau, Yanqiu Yu. Originally published in JMIR Serious Games (<https://games.jmir.org>), 30.Apr.2026. 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.

Original Paper

Conceptualization and Measurement of Withdrawal Symptoms in Gaming Disorder and Development and Psychometric Validation of the Gaming Withdrawal Symptoms Questionnaire: Cross-Sectional Study

Sylwia Starzec¹, MSc; Sławomir Śpiewak², PhD; Jolanta Starosta², PhD; Paweł Strojny², PhD

¹Doctoral School in the Social Sciences, Jagiellonian University, Kraków, Poland

²Institute of Applied Psychology, Faculty of Management and Social Communication, Jagiellonian University, Kraków, Poland

Corresponding Author:

Sylwia Starzec, MSc

Doctoral School in the Social Sciences

Jagiellonian University

Rynek Główny 34

Kraków, 31-010

Poland

Phone: 48 788310271

Email: sylwia.starzec@doctoral.uj.edu.pl

Abstract

Background: Gaming disorder (GD) is an emerging issue that leads to significant impairment, yet existing tools for measuring withdrawal symptoms in GD are limited and often fail to capture its multidimensional nature. Most current measures rely on single-item assessments or adapted tools from substance use disorders, overlooking cognitive, behavioral, and physiological components. A comprehensive, multidimensional questionnaire is needed to more accurately assess withdrawal in GD, aiding in early detection and intervention.

Objective: The objective of this study was to develop and psychometrically validate a comprehensive measurement tool, the Gaming Withdrawal Symptoms Questionnaire (GWSQ), capturing the multidimensional nature of withdrawal symptoms in GD, including affective, cognitive, behavioral, and physiological components.

Methods: A multistage psychometric approach was used, starting with item generation from a scoping literature review. Exploratory factor analysis and confirmatory factor analysis were conducted to refine the questionnaire. Reliability and validity were assessed using 2 cross-sectional studies. Data were collected anonymously via an online survey platform. Participants were recruited from gaming-related platforms and social media (eg, Discord, Reddit, and Facebook) and restricted to actively engaged adult gamers who passed attention check questions to ensure data quality.

Results: Study 1 involved 480 adults (mean age 23, SD 4.96 years; n=327, 68.1% male). Study 2 included 565 adults (mean age 25, SD 5.55 years; n=245, 43% male). Exploratory factor analysis revealed a 3-factor model of withdrawal symptoms: (1) motivational and cognitive symptoms, (2) affective symptoms, and (3) physical symptoms, explaining 54% of the variance. Confirmatory factor analysis confirmed adequate model fit ($\chi^2_{227}=887.8$; $P<.001$; comparative fit index=0.91; Tucker - Lewis index=0.90; root-mean-square error of approximation=0.072). The GWSQ demonstrated high internal consistency, with Cronbach α ranging from 0.89 (motivational and cognitive symptoms) to 0.90 (affective symptoms and physical symptoms). Correlations with related constructs (Internet Gaming Disorder Scale-Short Form [IGDS9-SF], Patient Health Questionnaire-9 items, and Generalized Anxiety Disorder-7 items) confirmed convergent validity with moderate associations (eg, IGDS9-SF: $r=0.48$, 95% CI 0.32-0.61) and discriminant validity, and normative data (sten scores) were established for the general population.

Conclusions: The GWSQ is the first validated multidimensional tool specifically designed to assess withdrawal symptoms in GD, representing a conceptual and methodological innovation. It addresses critical gaps in GD diagnosis and research by capturing a broader spectrum of symptoms beyond affective distress. The questionnaire's tripartite structure provides a framework for advancing theoretical models and informing etiological studies of GD. The GWSQ offers a robust measure for clinical research and enables differentiated assessment of symptom clusters. In real-world contexts, it can serve as a reliable patient-reported

outcome tool in forthcoming clinical trials of GD interventions, enabling precise monitoring of treatment effects. Nevertheless, given the nonclinical sample, further cultural validation and studies involving clinical populations are required.

(*JMIR Serious Games* 2026;14:e81407) doi:[10.2196/81407](https://doi.org/10.2196/81407)

KEYWORDS

behavioral addictions; DSM-5; gaming disorder; IGD-11; internet gaming disorder; psychological assessment; psychological testing; psychometrics; videogame addiction; withdrawal symptoms

Introduction

Background

Playing video games is an increasingly popular form of leisure activity with more than 3.3 billion gamers worldwide [1]. The popularity of games is growing with technological developments and easier access to immersive virtual worlds [2]. Although for most people gaming is a nonproblematic recreational activity, for a minority of users it has become poorly controlled and has negative psychosocial consequences [3-5]. According to different sources, disordered gaming impacts 3%-17% of the population, depending on the geographic region [6-9].

In 2013, for the first time, internet gaming disorder (IGD) was recognized as a potential new disorder and included as a “condition for further study” in the *DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition)* [10]. The positioning of IGD in Section III of the *DSM-5* was intended to encourage researchers to conduct further studies to determine whether the disorder should be included in the next generation of the *Diagnostic and Statistical Manual of Mental Disorders*. The criteria used to diagnose IGD include preoccupation, withdrawal symptoms, tolerance, loss of control, giving up other activities, continuation of gaming despite negative consequences, deception, using games to escape, and losing relationships/career/education because of gaming. More recently, in 2019, gaming disorder (GD) was recognized as a mental disorder and listed as a “disorder due to addictive behaviors” in the *International Classification of Diseases, Eleventh Edition (ICD-11)* [11]. Notably, the World Health Organization has taken a more conservative approach and proposed GD criteria characterized by 3 obligatory features (loss of control, increasing priority given to games, and continuation of gaming despite negative consequences) associated with clinically significant impairment. Despite the scientific debate on diagnostic criteria for gaming-related disorders and intensive research in the field, the most recent version of the *DSM-5 (DSM-5 TR [Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision])* does not include an updated definition of IGD [12].

The presence of the withdrawal symptoms criterion is the main difference between the diagnosis of GD proposed in *DSM-5* and *ICD-11* [13]. Withdrawal symptoms, according to the definition proposed by the American Psychological Association (APA) in the *DSM-5*, refer to symptoms such as anxiety, irritability, and sadness that occur when gaming is not possible [10]. This implies that withdrawal is not necessarily physiological in nature. Since the definition extends only to the emotions a gamer may experience, this criterion is one of the most criticized aspects of the entire GD diagnosis [13-16]. It has been argued that withdrawal in the context of GD lacks a

neurobiological basis, as there is no psychoactive substance interaction as in substance addictions [16]. Furthermore, it was also debated whether the withdrawal criterion is capable of distinguishing between passionate engagement in gaming and pathological gaming [14,17,18]. Nevertheless, recent studies have shown that withdrawal plays a central role in GD symptomatology and is a valid criterion for distinguishing players with GD from healthy gamers [19-23]. These studies are in line with the traditional division of addiction symptoms into core and peripheral ones, where withdrawal is classified as a core symptom that distinguishes addicted gamers from highly engaged gamers [24,25].

Current Perspectives on Withdrawal Symptoms in GD

Since withdrawal plays a crucial role in GD symptomatology, it requires an in-depth understanding, including precise characterization and methods for assessment. Based on the APA's 2013 definition, which describes withdrawal symptoms as sadness, anxiety, and irritability, this characterization appears insufficiently specific and lacks comprehensive detail. First, the withdrawal criterion should incorporate a defined time frame that specifies when symptoms are expected to emerge. Petry et al [15] emphasize that withdrawal symptoms should be carefully differentiated from emotional reactions that result from an external interruption or prevention of gaming. Genuine withdrawal manifests itself as distress lasting several hours to days following a cessation of gaming, distinguishing it from the immediate emotional responses triggered by external factors, such as a person forcibly stopping the game [14,15,26]. To distinguish negative emotions related to withdrawal from those caused by other factors, the assessment of withdrawal symptoms should be specified in such a way that symptoms disappear once the gamer plays again [14].

Continuing further, withdrawal may be more complex and multidimensional than initially described. Starzec et al [27], in their scoping review on the conceptual definition of withdrawal, which analyzed 3701 publications from 2018 to 2024, confirmed the presence not only of the affective component, but also of the cognitive, behavioral, physiological, and neurological dimensions. Based on three sources: (1) the specific definition provided by the authors, (2) the symptoms identified in the measurement tools through withdrawal check items, and (3) descriptions of participants' experiences of withdrawal, the review authors significantly expanded the list of abstinence symptoms in GD. In the affective component, they list symptoms such as anxiety, feeling depressed, sadness, feelings of stress, fear, restlessness, anger, impatience, nervousness, feelings of unhappiness, emptiness, irritability, frustration, and boredom. The physiological component includes symptoms such as craving, decreased hunger, decreased eating, stomach problems,

sweating, increased sleeping, headaches, and muscle pain. Furthermore, the cognitive component of withdrawal symptoms was extended to include difficulty concentrating, preoccupation, and intense thoughts of gaming. Behavioral symptoms (eg, inability to do other things and inability to relax) were also identified, as well as neurological symptoms (eg, functional connectivity was observed between the striatum and thalamus) [27]. These findings highlight the multifaceted nature of withdrawal in GD, underscoring the need for comprehensive assessment methods that capture its full complexity.

Furthermore, the methods of measuring withdrawal have also been examined in detail. A scoping review of the role of withdrawal in contemporary GD research provided further insight into how these symptoms are operationalized [27]. One of the most commonly used methods for assessing withdrawal involves self-report instruments, but other approaches, such as clinical interviews and physiological and neurophysiological measures, have also been applied [27,28]. While self-report instruments are widely used for assessing withdrawal symptoms, they have several limitations. First of all, in GD research, withdrawal is examined as a single questionnaire item across the IGD screening tools. This is the case with the Internet Gaming Disorder Scale-Short Form (IGDS9-SF), where withdrawal symptoms are assessed based on the question: “Do you feel more irritability, anxiety or even sadness when you try to either reduce or stop your gaming activity?” [29]. A similar approach to the withdrawal assessment was used in the Internet Addiction Test [30], Compulsive Internet Use Scale [31], 10-Item Internet Gaming Disorder Test [32], Chen Internet Addiction Scale-Gaming Version [33], Internet Gaming Disorder Criteria Checklist [15], and Screening Test for Problematic Gaming [34]. Additionally, several withdrawal screening tools have been developed based on withdrawal symptoms in substance dependence; however, these symptoms may differ from those found in behavioral addictions. Examples of this include the Internet Gaming Withdrawal Scale [35], which is a modified version of the Penn Alcohol Craving Scale [36]. Likewise, the Abstinence Symptoms Checklist, adapted to the gaming context by Giordano et al [37], was originally designed for cocaine withdrawal symptoms. To the best of our knowledge, to date, there is no withdrawal assessment tool that takes into account the specificity of GD and is directly designed for behavioral addictions, which may differ from substance use disorders. Moreover, there is no standardized tool that captures the multidimensionality of withdrawal, not only from an affective dimension, but also from a cognitive, behavioral, and physiological perspective.

The development of a more refined and comprehensive withdrawal questionnaire is essential for advancing both research and clinical practice. Current research highlights that withdrawal symptoms, including cravings and mood disturbances, are not only key indicators of GD but also predictors of more severe cases [38]. A specialized questionnaire would be crucial in accurately differentiating between normal and pathological gaming behavior, making it vital for early detection and monitoring the progression of the disorder. Moreover, such a tool would be instrumental in capturing the multidimensionality of withdrawal, considering its affective, cognitive, behavioral,

and physiological dimensions, which are often overlooked in existing assessments. This would allow for a more nuanced understanding of withdrawal, helping clinicians develop tailored treatment and prevention strategies. Furthermore, research supports the claim that GD is a major risk factor for the development of other mental diseases, such as depression and anxiety disorders [19,39,40]. This indicates that a comprehensive evaluation of withdrawal symptoms (eg, mood swings, sleep disturbances, and concentration difficulties) could help detect emerging mood disorders at an early stage. Identifying these symptoms early would allow for the implementation of preventive interventions, potentially averting the escalation of both gaming-related issues and cooccurring mental health problems.

Overview of the Work

To achieve the objectives of this research, we implemented a structured, multistage approach. The development process began with a reanalysis of data from a scoping literature review, aimed at identifying measurement methods and withdrawal symptoms associated with GD. These findings guided the construction of the Gaming Withdrawal Symptoms Questionnaire (GWSQ). Next, during the validation process, we conducted 2 empirical studies. Study 1 involved an exploratory factor analysis (EFA) to identify the underlying structure of the questionnaire, along with an assessment of internal consistency. Study 2 used a confirmatory factor analysis (CFA) to validate this structure, alongside evaluations of convergent and divergent validity to ensure the robustness of the GWSQ as a psychometric tool. Finally, the normalization phase focused on establishing interpretation guidelines, including the evaluation of sten standards, to facilitate the practical application of the GWSQ in research and clinical settings.

Aim of the Study

In order to advance research and further the scientific understanding of gaming-related disorders, we aim to develop the first comprehensive withdrawal inventory (covering all symptoms identified in the literature) to evaluate the core criteria for GD [27]. This is especially important due to the absence of precise diagnostic criteria for withdrawal and limited comparability between the IGD and the GD criteria. This is a crucial step in GD research in order to provide empirical data related to the diagnostic properties and efficacy of withdrawal criteria. Consequently, it will influence not only GD research but also clinical practice.

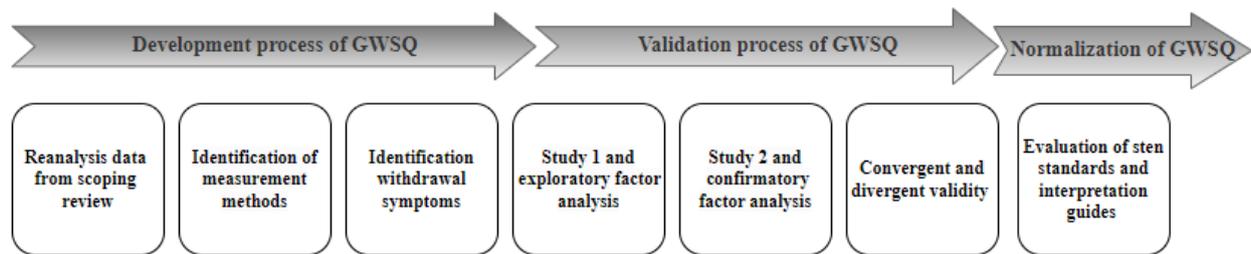
Methods

Development of GWSQ

The development and validation of the GWSQ, which involved 2 cross-sectional phases (Study 1: EFA, Study 2: CFA), followed recommendations for transparent and complete reporting of instrument development studies (Figure 1). Specifically, the research adhered to established psychometric standards, including the principles outlined in the COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments) guidelines [41] and the STROBE (Strengthening the Reporting of Observational Studies in

Epidemiology) statement [42] for cross-sectional studies. The completed checklists for these reporting guidelines have been provided in [Multimedia Appendix 1](#).

Figure 1. Methodological framework for the Gaming Withdrawal Symptoms Questionnaire (GWSQ): development, validation, and normalization. This figure summarizes the 2-study design used for the conceptualization and exploratory factor analysis (study 1; n=480) and the psychometric validation via confirmatory factor analysis (study 2; n=565) of GWSQ. The studies were conducted among a predominantly Polish-speaking sample of gamers.



The GWSQ was developed and validated in the English language. The primary phase of instrument creation focused on establishing content validity and generating a comprehensive item pool.

1. Content domain definition: the initial item pool was generated through a scoping literature review aimed at systematically identifying all available empirical measures and conceptual definitions of gaming withdrawal symptoms. This systematic review established the content domain of the GWSQ, ensuring it was comprehensive and grounded in existing research and clinical practice. A detailed description of the review methodology, selection process, and results regarding the operational and conceptual

definition of gaming withdrawal symptoms is presented in a previous publication by the authors [27].

2. Initial item pool: the synthesis of the scoping review findings resulted in a pool of 47 distinct withdrawal symptoms. These symptoms were formulated into preliminary questionnaire items, with each item using a 5-point Likert scale (1=strongly disagree to 5=strongly agree). The initial pool of items was designed to reflect 4 preliminary theoretical dimensions: affective, cognitive, behavioral, and physical symptoms.

[Table 1](#) presents the specific symptoms identified and used for item generation.

Table 1. Identified gaming withdrawal symptoms based on a scoping review of existing measurement tools used to assess withdrawal in the context of gaming disorder. The table groups symptoms into preliminary domains (affective, cognitive, behavioral, and physical). “n” indicates the number of tools in the scoping review [27] that contained the specific symptom. This step informed the item pool generation for the Gaming Withdrawal Symptoms Questionnaire development.

Symptom	Value, n
Affective symptoms	
Irritability	13
Anxiety	11
Sadness	6
Stress	2
Fear	1
Feeling bad	1
Restlessness	8
Feeling moody	2
Anger	3
Feeling depressed	8
Impatience	2
Feeling upset	1
Feeling uneasy	1
Frustration	3
Nervousness	1
Feeling lethargic	1
Dysphoria	2
Unsatisfaction	1
Boredom	2
Distress	1
Emptiness	1
Cognitive symptoms	
Thoughts about gaming	2
Difficulty concentrating	1
Preoccupation with gaming	2
Behavioral symptoms	
Inability to experience pleasure	2
Lack of pleasure	1
Disinclination for activity	2
Inability to do other things	1
Inability to identify activities to do	1
Attempts to find games to play	1
Lack of motivation	1
Inability to relax	1
Physical symptoms	
Urge	2
Craving	2
Decreased hunger	1
Decreased eating	1

Symptom	Value, n
Increased eating	1
Headache	1
Stomach problems	1
Sweating	1
Insomnia	1
Increased sleeping	1
Increased dreaming	1
Muscle pain	1
Chills, tremors/shakes, and twitching	1
Backaches or other physical discomfort	1
Fatigue	1

Given that the 47 items were systematically derived and operationalized directly from the empirical and clinical literature via the published scoping review, the authors used large-scale factor analysis for subsequent item refinement. The EFA performed in Study 1 thus served as the initial robust empirical method for item purification and scale construction. This statistical approach was used to reduce the initial pool of 47 items, verify the underlying latent factor structure, and ensure the construct validity and internal consistency of the emerging GWSQ subscales. The subsequent CFA in Study 2 provided final validation of the structure.

Study 1: Exploratory Validation

Participants and Procedure

Participants were recruited through advertisements on gaming-related social media groups using convenience and snowball-sampling methodologies. The posts included information about the study, that is, an invitation to participate, a description of the topic of the study, an estimated time to complete the survey, an assurance of anonymity, an opportunity to withdraw at any time, assurance of confidentiality, and an explanation of the purpose of data use. Data collection was conducted through an online survey hosted on Qualtrics [43], which was performed between January and February 2024. Participation was entirely voluntary, and no financial compensation was offered to participants.

Before starting the questionnaire, participants provided their informed consent to continue and take part in the study. The eligibility criteria required participants to be 18 years of age or older and actively involved in video gaming. Participants became eligible to take part in the study upon responding to the following screening question: “Have you played any video games in the past 12 months (yes/no)?” This criterion was chosen in accordance with previous studies [13,44,45] that demonstrated its effectiveness in identifying individuals with gaming experience. Only participants who provided a confirming answer were allowed to participate in the study. A total of 21 participants were excluded based on this inclusion criterion. To qualify for analysis, participants had to pass at least 1 attention check. The attention check questions were designed to ensure that participants were reading the questions

carefully and engaging with the task. Specifically, participants were asked to select a predetermined answer. Finally, the results of 101 participants were excluded on the basis of a negative answer to this question. A total of 602 people took part in the study. Due to the unsuccessful screening question (n=21) and the failure of the attention check (n=101), 480 results were included in the analysis. The minimum required sample size for the EFA was determined based on the established psychometric guideline of a minimum 10:1 participant-to-item ratio (47 items), necessitating a sample of 470 participants [46,47].

Measures

Sociodemographic Information

Sociodemographic data included participants’ gender, age, and nationality. If participants entered their age as <18 years, they were informed that they could not participate further in the study. Participants aged <18 years were excluded from the study due to potential concerns about the influence of external factors, such as parental restrictions on gaming behavior.

GWSQ Tool

The GWSQ is an assessment tool designed to measure the intensity of withdrawal symptoms. The GWSQ items were developed in the English language to ensure the scale’s content was universal and applicable to international studies. The GWSQ includes 47 items reflecting all symptoms identified during the systematic review process. The items are organized conceptually into 4 components: affective (eg, sadness and irritability), behavioral (eg, attempts to find games to play and inability to relax), cognitive (eg, concentration problems and thoughts about gaming), and physiological (eg, headaches and sweating). Respondents were instructed to indicate the intensity of those feelings or states they experienced in situations when they desired to play games but were unable to do so. All items are rated on a 5-point Likert scale: 1 (“not at all”), 2 (“slightly intense”), 3 (“moderately intense”), 4 (“intense”), and 5 (“extremely intense”). The psychometric properties of the final scale, including its internal consistency reliability (Cronbach α), are detailed in the Results section.

Gaming Involvement Scale

The Gaming Involvement Scale (GIS) is a self-report tool to measure involvement in video game-related activities across six different indications: (1) playing video games, (2) watching game-related video, (3) thinking about video games, (4) reading game-related content, (5) talking about video games, and (6) considering game-related purchases. The English version of the scale was administered to all participants. Participants were asked to estimate the amount of time (in minutes) they dedicated to these activities on an average workday and an average weekend day. The scale was developed in previous studies by the team [48]. The internal consistency was good (Cronbach $\alpha=0.80$).

IGDS9-SF Measurement

The IGDS9-SF [29] is a psychometric measurement adapted from the 9 IGD criteria according to the *DSM-5* classification [10]. The English version of the tool was used. The tool consists of 9 items that are rated on a 5-point Likert scale: 1 (“never”), 2 (“rarely”), 3 (“sometimes”), 4 (“often”), and 5 (“very often”). Total scores can range from 9 to 45, with higher scores indicative of higher levels of disordered gaming. An example of an item is: “Do you feel more irritability, anxiety or even sadness when you try to either reduce or stop your gaming activity?” The internal consistency was good (Cronbach $\alpha=0.84$).

Statistical Approach

All statistical analyses were performed using SPSS Statistics software (version 29; IBM Corp). Prior to data analysis, the completeness of the GWSQ items was assessed in Study 1. No missing data were observed, which precluded the need for applying the missing completely at random test or handling missing-data techniques such as multiple imputation. Then, the statistical analysis of the collected data included descriptive statistics of the main sample’s characteristics. Subsequently, to examine the underlying structure of the GWSQ, we used EFA. Given that the data were not approximately normally distributed, we used the principal-axis method as the extraction method. To determine the number of factors to fit the data, we used scree-plot inspection combined with the Kaiser rule. Due to the high likelihood of between-item correlations, oblimin rotation was used for extraction. Items with a loading of 0.4 or higher were included in the factor, and if items were loading on 2 factors, they were included in the factor with the highest loading.

Study 2: Confirmatory Validation

Participants and Procedure

The inclusion criteria for participation in Study 2 were the same as in the previous one (being 18 years of age or older and playing video games). If the respondent did not meet the age and gaming criteria, they were not allowed to participate in the study. The instructions indicated that the participation was anonymous, voluntary, and that the data were collected for research purposes only. Respondents to Study 2 were sought via gaming-related social media (Facebook, Reddit, Discord, and Steam Community), and all participants had not participated in Study 1. Data collection was performed using an online survey hosted on the Qualtrics platform [43] and took place

between November and December 2024. The survey included sociodemographic questions and a set of questionnaires. In order to control the quality of the data, 2 attention check questions were used in the study. From the initial sample of 590, we excluded participants who did not play games ($n=4$) or failed the double attention check ($n=21$). Subsequently, a total of 565 participants were included in the final analysis.

Measures

Sociodemographic Information

The study collected demographic data, which included gender, age, and nationality of participants.

GWSQ Tool

The GWSQ consists of 23 items. This version is based on the results of an EFA conducted in Study 1. The English version of the scale was used. Responses were given on a 5-point Likert scale, ranging from 1 (“not at all”) to 5 (“extremely intense”). Examples of items include: “Anxiety,” “Irritability,” and “Thoughts about gaming.”

IGDS9-SF Tool

The IGDS9-SF [49] is a short, standardized psychometric tool for the assessment of IGD. The scale contains 9 criteria that form the basis for a possible diagnosis according to APA’s guidelines [10]. The participants answered 9 questions about their gaming behavior and experiences over the past 12 months. Cronbach α coefficient for the questionnaire was 0.74.

Patient Health Questionnaire-9

The Patient Health Questionnaire-9 items (PHQ-9) [50] is a useful tool for screening for depression. It was used to measure the level of depressive symptoms of study participants. This scale is based on the criteria for the disorder, as contained in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* [51]. The answers of the questionnaire are rated on a 4-point Likert scale, ranging from 0 (“not at all”) to 3 (“nearly every day”). An example of an item is: “Little interest or pleasure in doing things.” In this study, the internal consistency was high (Cronbach α coefficient for the questionnaire was 0.87).

Generalized Anxiety Disorder-7 items

The Generalized Anxiety Disorder-7 items (GAD-7) [52] is a brief self-report to identify the presence of anxiety symptoms in participants. All participants completed the English version of the scale. The scale uses a 4-point Likert scale, ranging from 0 (“not at all”) to 3 (“nearly every day”). An example of an item is: “Feeling nervous, anxious, or on edge.” The internal consistency of the questionnaire in this study was high (Cronbach $\alpha=0.89$).

GIS Questionnaire

The GIS [48] is a questionnaire designed to assess the extent of involvement in gaming activities. Participants received the English version of the scale. The participant was asked to answer a question regarding how they spent their time on particular video game-playing activities. The respondent was required to determine the number of minutes spent on 6 activities during both working and weekend days. These activities included

playing video games, contemplating playing video games, reading or watching tutorials, reviews, theory, or additional lore related to video games, watching video game streams and gameplay videos (including e-sports games), talking or writing about video games, and purchasing additional gaming devices or gadgets. In this study, the scale demonstrated acceptable internal consistency (Cronbach $\alpha=0.79$).

Gaming Motivation Inventory

The Gaming Motivation Inventory (GMI) [53] was included in the study to control the level of motivation to play video games among participants. The English version of the questionnaire was given to all participants. The selected subscales were “Recreation” and “Introjected regulation.” The “introjected regulation” subscale includes items such as “because I must play to feel good about myself” and “because otherwise I would feel bad about myself,” reflecting extrinsic motivation driven by internal pressures or guilt. This subscale was hypothesized to have strong convergence with withdrawal symptoms, as both are indicative of problematic gaming behavior. On the other hand, the “Recreation” subscale, which includes items such as “because it is fun” and “to relax,” represents intrinsic motivation tied to the enjoyment and leisure aspects of gaming. This subscale was expected to show weaker or no correlation with withdrawal symptoms, serving as a measure of divergent validity. A total of 6 GMI items were used in the survey. The questionnaire used a 7-point Likert scale ranging from 1 (“It does not correspond at all”) to 7 (“It corresponds exactly”). An example of an item from the “Introjected regulation” subscale is “because I must play to feel good about myself.” The respondent was given the opportunity to select one of the options from 1 “It does not correspond at all” to 7 “It corresponds exactly.” The internal consistency of the scale was good (Cronbach $\alpha=0.81$).

Satisfaction With Life Scale

The Satisfaction With Life Scale (SWLS) [54] focuses on happiness in hedonistic terms (occurrence of positive emotions, nonoccurrence of negative emotions, and satisfaction with the situation). Participants received the English version of the scale. The scale contains 5 items. The 7-point Likert scale ranges from 1 (“strongly disagree”) to 7 (“strongly agree”). An example of an item is “In most ways my life is close to my ideal.” Once all scores have been aggregated, the resulting total score provides an indication of overall satisfaction with life. The score ranges from 5 to 35 points. In this study, the internal consistency was high (Cronbach α coefficient for the questionnaire was 0.87).

Psychological Well-Being Scale

The Psychological Well-Being Scale [55] measures eudaimonic happiness, focusing on self-actualization, self-acceptance,

personal development, and mastery over the environment. All participants responded to the English version of the scale. The scale consists of 18 items. A 7-point Likert scale was used in the questionnaire, ranging from 1 (“strongly disagree”) to 7 (“strongly agree”). An example of an item is “I like most parts of my personality.” In this study, the internal consistency was high (Cronbach α coefficient for the questionnaire was 0.85).

Statistical Approach

Similar to Study 1, an assessment of data completeness for all GWSQ items in Study 2 revealed no missing data. Consequently, neither the Little missing completely at random test nor any missing-data handling techniques were required for subsequent analyses. Statistical analyses were performed using SPSS Statistics software (version 29) and the *lavaan* package in R software (version R-4.4.2; R Foundation for Statistical Computing). First, descriptive statistics of the main sample’s characteristics were examined using SPSS. Subsequently, CFA was conducted using the *lavaan* package in R software. Maximum likelihood estimation was used. The hypothesized model was a 3-factor model as described above. The model was assessed using the chi-square with respective *P* values ($P>.01$), root-mean-square error of approximation (<0.06), comparative fit index (>0.95), Tucker - Lewis index (>0.95), and standardized root-mean-square residual (<0.08) [56].

Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki. The approval was granted by the Research Ethics Committee at the Institute of Applied Psychology of the Jagiellonian University (opinion number 102/2021). The approval date was July 19, 2021. Informed consent was obtained from all individual participants included in the study. Privacy and confidentiality were rigorously protected, as the study was anonymous. Participation was voluntary, and participants received no financial compensation. No images or supplementary materials included in this manuscript allow for the identification of individual participants.

Results

Study 1: Exploratory Validation

Descriptive Statistics

Descriptive statistics for the sample demographics, gaming engagement, and severity of GD symptoms are presented in Table 2. The total sample for Study 1 consisted of 480 participants who were predominantly male ($n=327$, 68.1%) with a mean age of 23 (SD 4.96) years. The sample was predominantly Polish ($n=473$, 98%), which necessitates caution regarding generalizability.

Table 2. Demographic and gaming-related information of the sample (Study 1).

Characteristic	Value (n=480)
Gender, n (%)	
Male	327 (68.1)
Female	138 (28.7)
Nonbinary	12 (2.5)
Refused to state	3 (0.6)
Age (years), mean (SD; range)	23 (4.96; 18-53)
Nationality, n (%)	
Poland	473 (98)
United Kingdom	4 (0.8)
Germany	1 (0.2)
Australia	1 (0.2)
Poland-Germany	1 (0.2)
Gaming engagement (hours), mean (SD)	
Playing video games	13.6 (16.88)
Watching video game streams and gameplay videos	8.27 (9.17)
Reading or watching game-related content	7.11 (11.63)
Talking or writing about video games	7.15 (10.13)
Considering buying additional content or collectibles related to games	2.65 (6.55)
Severity of GD symptoms, mean (SD)	
IGDS9-SF ^a	17.27 (6.02)

^aIGDS9-SF: Internet Gaming Disorder Scale-Short Form.

Analysis of gaming engagement showed that participants spent an average of 23 (SD 16.39) hours per week actively playing video games. Beyond primary gameplay, thinking about video games was the most time-consuming secondary activity, while the least amount of time was dedicated to considering the purchase of additional content.

The severity of GD symptoms, as measured by the IGDS9-SF, demonstrated a mean score of 17.27 (SD 6.02). The median score of 16 (IQR 13-20) and a mode of 11 suggest a positive skew in the distribution of IGD symptoms within this cohort.

EFA Method

An initial analysis was performed to verify the justification for subjecting the collected material to factor analysis. Kaiser-Meyer-Olkin measure of sampling adequacy was 0.94. The Bartlett sphericity test result was as follows: $\chi^2_{1128}=12468.1$; $P<.001$, which indicates the validity of

performing a factor analysis due to the heterogeneous structure of the questions.

Following initial exploration, the principal-axis method was conducted on the full item set. Based on the Kaiser criterion ($l>1$), 9 factors were initially extracted, explaining 62% of the common variance. Items with factor loadings below 0.40 were systematically excluded from the analyses. A total of 24 items did not meet the factor loading threshold ($\geq.40$) and were removed from the scale. After the removal of these items, 23 items remained for further analysis. Consequently, the initial 9-factor solution was reduced to a 3-factor model.

To confirm the dimensionality of the retained items, the EFA was rerun using the principal-axis method with an oblimin rotation (due to expected correlations between factors). The final 3-factor solution was extracted, which explained 54% of the common variance. The complete factor loading matrix for this structure is shown in [Table 3](#).

Table 3. Factor structure of the Gaming Withdrawal Symptoms Questionnaire derived from exploratory factor analysis (EFA; study 1). The EFA used principal-axis method with oblimin rotation and Kaiser normalization on a sample of 480 gamers (n=473, 98% Polish-speaking participants). Data were collected online between January and February 2024.

	Factor 1 ^a	Factor 2 ^b	Factor 3 ^c
Inability to do other things	0.84 ^d	-0.01	0.03
Inability to identify activities to do	0.84 ^d	-0.04	-0.07
Disinclination for activity	0.78 ^d	-0.05	0.05
Inability to relax	0.77 ^d	0.07	-0.06
Difficulty concentrating	0.69 ^d	0.09	0.07
Lack of motivation	0.68 ^d	0.09	0.18
Preoccupation with gaming	0.58 ^d	0.14	0.09
Inability to experience pleasure	0.54 ^d	0.05	0.26
Frustration	0.03	0.84 ^d	-0.08
Anger	-0.10	0.82 ^d	0.13
Irritability	-0.08	0.76 ^d	0.05
Impatience	0.10	0.69 ^d	-0.15
Feeling upset	0.16	0.69 ^d	-0.09
Nervousness	0.05	0.66 ^d	0.27
Feeling bad	0.34	0.46 ^d	0.00
Sweating	-0.17	0.14	0.79 ^d
Muscle pain	0.09	-0.12	0.70 ^d
Chills, tremors/shaking, and twitching	-0.05	0.09	0.69 ^d
Stomach problems	0.04	-0.05	0.69 ^d
Headaches	0.07	-0.04	0.68 ^d
Backaches or other physical discomfort	0.06	0.02	0.65 ^d
Decreased eating	0.20	-0.04	0.46 ^d
Increased dreaming	0.20	-0.03	0.45 ^d

^aFactor 1: motivational and cognitive withdrawal symptoms.

^bFactor 2: affective withdrawal symptoms.

^cFactor 3: physical withdrawal symptoms.

^dValues indicate the highest factor loading for each item (primary factor loading).

Consistent with the use of the oblimin oblique rotation, the 3 extracted factors were correlated. The interfactor correlations ranged from 0.19 to 0.49. Specifically, the correlation between the motivational and cognitive factor (F1) and the affective factor (F2) was $\phi=0.39$; between F1 and the physical factor (F3) was $\phi=0.49$; and between F2 and F3 was $\phi=0.19$. These interfactor correlations support the existence of a higher-order latent construct (gaming withdrawal symptoms) while confirming that the 3 dimensions are distinct, as no correlation exceeded the threshold of $\phi=0.85$, which typically suggests redundant factors.

In connection with the analysis of items included in individual subscales, the following names were chosen for the extracted components:

- Factor 1 (Motivational and Cognitive Withdrawal Symptoms): the subscale includes the following 8 items connected with an individual's involvement in various activities: "Inability to do other things," "Inability to identify activities to do," "Disinclination for activities," "Inability to relax," "Lack of motivation," "Difficulty concentrating," "Preoccupation with gaming," and "Inability to experience pleasure." The factor loading values for the given items range from 0.54 to 0.84. The subscale

demonstrated high internal consistency with Cronbach $\alpha=0.90$.

- Factor 2 (Affective Withdrawal Symptoms): the factor contains the following 7 items related to heightened emotional states: “Frustration,” “Anger,” “Irritability,” “Nervousness,” “Impatience,” “Feeling upset,” and “Feeling bad.” The factor loading values for the given items range from 0.46 to 0.84. The internal consistency of the subscale was high (Cronbach $\alpha=0.86$).
- Factor 3 (Physical Withdrawal Symptoms): the subscale includes the following 8 items related to experienced physiological symptoms: “Sweating,” “Chills, tremors/shaking, and twitching,” “Headaches,” “Muscle pain,” “Stomach problems,” “Backaches or other physical discomfort,” “Decreased eating,” “Increased dreaming.” The factor loading values for the given items range from 0.45 to 0.79. Factor 3 demonstrated good reliability, assessed using Cronbach α (Cronbach $\alpha=0.80$).

The initial EFA confirmed the validity of subjecting the collected data to a factorial structure examination. Although an initial 9-factor solution was extracted following the Kaiser criterion, a refinement process led to the identification of a 3-factor model that explained 54% of the total variance. This model includes Motivational and Cognitive Withdrawal Symptoms, Affective Withdrawal Symptoms, and Physical

Withdrawal Symptoms, all demonstrating good internal consistency (Cronbach α ranging from 0.80 to 0.90). The structure highlights the complexity of gaming withdrawal, extending beyond emotional distress to include cognitive difficulties and somatic complaints. To further validate the identified 3-factor model, a CFA was conducted in the next stage of the study. Additionally, convergent and divergent validity and the evaluation of scale standards were performed. To perform such analyses, Study 2 applied additional measurement tools: the PHQ-9 [50], the GAD-7 [52], the SWLS [54], the Psychological Well-Being Scale [55], and the GMI [53].

Study 2: Confirmatory Validation

Descriptive Statistics

Descriptive statistics characterizing the sample’s demographics and gaming engagement patterns for the total sample (N=565) are presented in Table 4. The sample exhibited a concentration of Polish participants, accounting for the vast majority of respondents. Other nationalities were present in very low numbers. Regarding gender distribution, the proportion of female participants was slightly higher than that of males. A small number of individuals identified as nonbinary or declined to specify their gender. Participant ages ranged from 18 to 45 years.

Table 4. Demographic and gaming-related information of the sample (study 2).

Demographic	Value (n=565)
Gender, n (%)	
Male	245 (43)
Female	294 (52)
Nonbinary	13 (2)
Refused to state	10 (1.77)
Age (years), mean (SD; range)	25 (5.55; 18-45)
Nationality, n (%)	
Poland	545 (96.4)
United Kingdom	4 (0.7)
Nederland	4 (0.7)
United States	2 (0.35)
Canada	1 (0.17)
China	1 (0.17)
France	1 (0.17)
India	1 (0.17)
Italy	1 (0.17)
Malaysia	1 (0.17)
Scotland	1 (0.17)
Singapore	1 (0.17)
Thailand	1 (0.17)
Vietnam	1 (0.17)
Gaming engagement (hours), mean (SD)	
Playing video games	9.25 (7.5)
Watching video game streams and gameplay videos	3.65 (6.5)
Reading or watching game-related content	1.86 (3.08)
Talking or writing about video games	1.73 (3.83)
Considering buying additional content and collectibles related to games	0.43 (1.3)
Severity of GD symptoms, mean (SD)	
IGDS9-SF ^a	18.64 (5.79)

^aIGDS9-SF stands for the Internet Gaming Disorder Scale-Short Form.

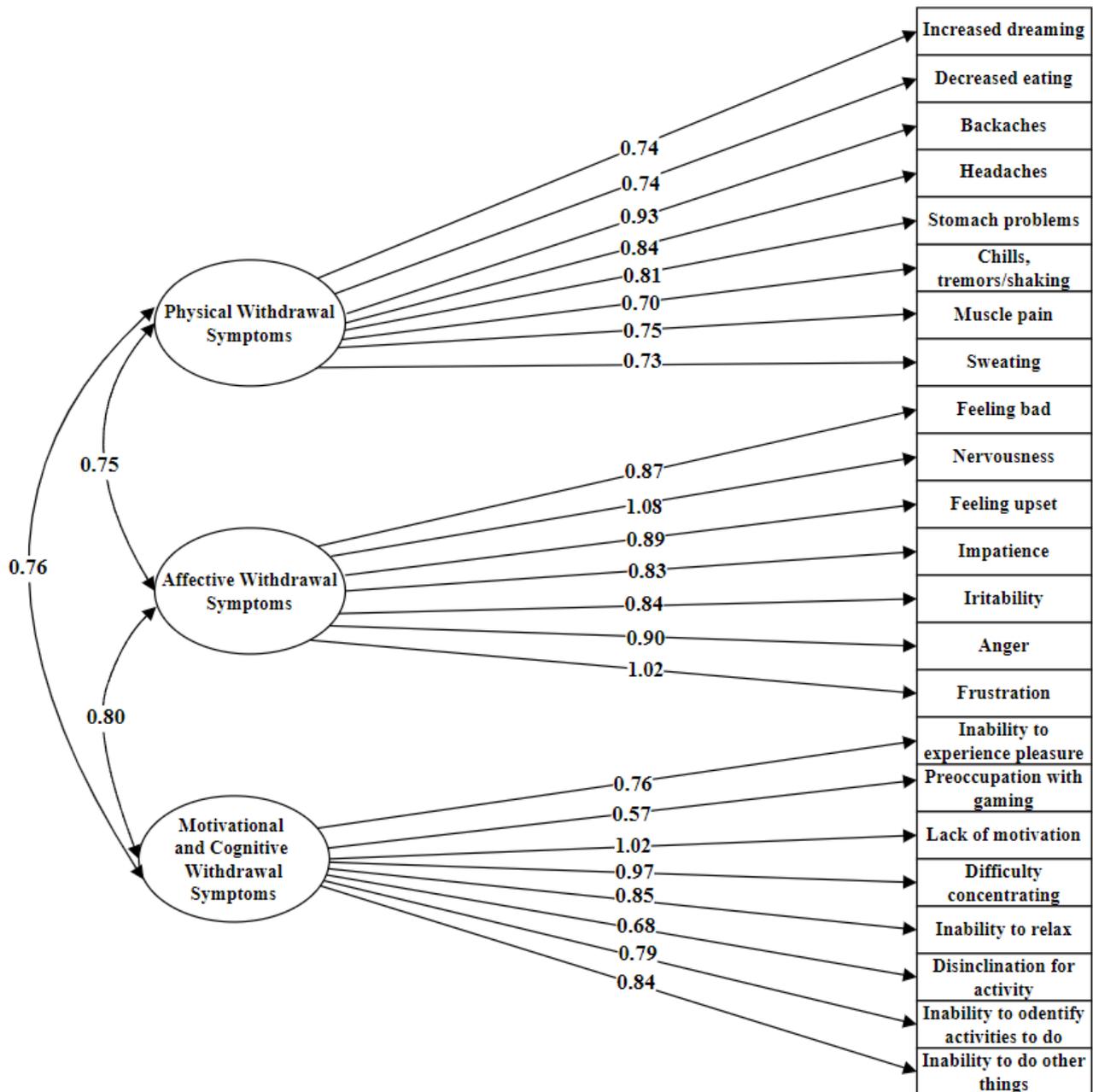
Analysis of self-reported weekly gaming-related behaviors based on the GIS indicated that participants spent the most time actively playing video games. The next highest time commitment was devoted to watching video game streams and gameplay videos. Other related activities, such as reading game-related content or talking about video games, constituted significantly less weekly time commitment. The least amount of time was spent on considering the purchase of additional game content.

Descriptive statistics were also computed for the IGDS9-SF to characterize self-reported symptom severity.

CFA Results

The hypothesized model demonstrated an acceptable, though not perfect, fit to the data. CFA conducted on an independent sample of 565 gamers (n=545, 96.4% Polish-speaking participants) showed the following fit indices: $\chi^2_{227}=887.8$; $P<.001$; comparative fit index=0.91; Tucker-Lewis index=0.90; root-mean-square error of approximation=0.072 (90% CI 0.067-0.077); standardized root-mean-square residual=0.049. The model was estimated using maximum likelihood estimation. All factor loadings were statistically significant ($P<.001$). The standardized factor loadings are presented in [Figure 2](#).

Figure 2. Three-factor model of the Gaming Withdrawal Symptoms Questionnaire (GWSQ). The figure illustrates the final confirmatory factor analysis structure (motivational and cognitive, affective, and physical) of GWSQ. The model was tested on the study 2 sample (n=565 predominantly Polish-speaking participants). This figure displays the 3-factor structure of GWSQ, including standardized factor loadings and interfactor correlations.



Convergent and Divergent Validity

To assess the convergent and discriminant validity of the GWSQ, Pearson correlations were conducted with selected

measures assessing related and distinct constructs. The results are presented in [Tables 5](#) and [6](#).

Table 5. Correlations between the Gaming Withdrawal Symptoms Questionnaire (GWSQ) and measures of convergent validity. Three decimal places were reported for enhanced precision, particularly to better differentiate between closely related values.

Variable	GWSQ, <i>r</i> (95% CI)
IGDS9-SF ^a	0.476 ^b (0.321-0.606)
PHQ-9 ^c	0.479 ^b (0.321-0.612)
GAD-7 ^d	0.415 ^b (0.246-0.560)
GMI ^e (introjected regulation)	0.324 ^b (0.137-0.488)

^aIGDS9-SF: Internet Gaming Disorder Scale-Short Form.

^b $P < .01$ (2-tailed).

^cPHQ-9: Patient Health Questionnaire-9 items.

^dGAD-7: Generalized Anxiety Disorder-7 items.

^eGMI: Gaming Motivation Inventory.

Table 6. Correlations between the Gaming Withdrawal Symptoms Questionnaire (GWSQ) and measures of discriminant validity.

Variable	GWSQ, <i>r</i> (95% CI)
SWLS ^a	0.28 ^b (0.09 to 0.45)
RYFF-18 ^c	-0.36 ^b (-0.52 to -0.17)
GMI ^d (Recreation)	-0.009 (-0.21 to 0.19)

^aSWLS: Satisfaction With Life Scale.

^b $P < .01$ (2-tailed).

^cRYFF-18: Ryff Scales of Psychological Well-Being-18 items.

^dGMI: Gaming Motivation Inventory.

As shown, the GWSQ demonstrated significant positive correlations with all measures. The strongest correlations were observed with the IGDS9-SF ($r=0.476$, 95% CI 0.321-0.606; $P<.001$) and the PHQ-9 ($r=0.479$, 95% CI 0.321-0.612; $P<.001$). The correlation with the GAD-7 was also significant, though somewhat weaker ($r=0.415$, 95% CI 0.246-0.560; $P<.001$). These correlations were precisely estimated, as evidenced by the consistently narrow 95% CIs, which increase confidence that these observed effect sizes accurately reflect the true population association. A similar pattern was found with the Introjected regulation subscale of the GMI ($r=0.324$, 95% CI 0.137-0.488; $P<.001$).

To assess the discriminant validity of the GWSQ, correlations were examined between the GWSQ and measures of constructs theoretically unrelated to gaming withdrawal symptoms. Specifically, correlations were calculated with the SWLS, the Ryff Psychological Well-Being Scale-18 items (RYFF-18), and

the Recreation subscale of the GMI. Table 6 presents the Pearson correlation coefficients between the GWSQ and these measures.

Discriminant validity analysis demonstrated weaker, but still significant correlations between GWSQ and SWLS ($r=0.28$, 95% CI 0.09-0.45; $P=.004$), as well as a moderate negative correlation with RYFF-18 ($r=-0.36$, 95% CI -0.52 to -0.17; $P<.001$). Notably, there was no significant correlation between GWSQ and the GMI Recreation subscale ($r=-0.009$, 95% CI -0.21 to 0.19; $P=.93$), suggesting that the GWSQ does not capture general gaming motivation to play for recreational purposes.

Then, the reliability of all 3 subscales of the GWSQ was assessed. Table 7 presents Cronbach α , composite reliability (CR), and average variance extracted (AVE) statistics. The diagonal shows the square root of the AVE for each construct. The lower triangle of the last 3 columns of the table presents the correlation coefficients.

Table 7. Results of reliability analysis of Gaming Withdrawal Symptoms Questionnaire (GWSQ) and all subscales.

Subscale	CR ^a	Cronbach α^b	AVE ^c	Motivational and Cognitive Withdrawal Symptoms, r	Affective Withdrawal Symptoms, r	Physical Withdrawal Symptoms, r
Motivational and Cognitive Withdrawal Symptoms	0.89	0.89	0.51	0.80	N/A ^d	N/A
Affective Withdrawal Symptoms	0.90	0.90	0.56	0.80 ^e	0.96	N/A
Physical Withdrawal Symptoms	0.91	0.90	0.55	0.76 ^e	0.75 ^e	0.88
Total GWSQ	0.95	0.95	N/A	N/A	N/A	N/A

^aCR: composite reliability.

^bDesired coefficient value ≥ 0.75 .

^cAVE: average variance extracted (desired coefficient value ≥ 0.50).

^dN/A: not applicable.

^e $P < .001$.

The results of the reliability analysis and the explained variances for the total GWSQ and all scales indicate good internal consistency and validity. CR and Cronbach α values for each scale exceed the desired thresholds of 0.80, demonstrating strong reliability. Additionally, the AVE values were all > 0.50 , supporting construct validity. The square root of the AVE for each construct was compared with the correlation coefficients between the constructs. The square roots of the AVE values on the diagonal were greater than the corresponding interconstruct correlations in the lower triangle of the matrix, confirming the validity of the measurement.

Evaluation of Sten Standards and Interpretation Guidelines

To facilitate the interpretation of the scores obtained from the GWSQ, a standardization process was conducted. The raw scores were converted into sten scores to standardize individual performance relative to a reference group. To ensure accurate standardization, the reference group was selected to be representative of the general population, taking into account factors such as age, gender, and gaming experience. The raw scores were grouped into deciles, with sten scores ranging from 1 to 10, where lower scores (1-4) represent minimal withdrawal symptoms, average scores (5-6) indicate moderate withdrawal symptoms, and higher scores (7-10) suggest pronounced withdrawal symptoms [57]. The calculated stens for the GWSQ are presented in Table 8.

Table 8. Sten norms for the Gaming Withdrawal Symptoms Questionnaire and its subscales. Sten norms were derived from the large validation sample of 565 gamers (predominantly Polish-speaking participants). These norms are based on a general population of gamers rather than a clinical sample diagnosed with gaming disorder and should be interpreted accordingly when used in a clinical context. Sten scores are standard 10-point scores with a mean of 5.5 (SD 2).

Sten	Total score	Motivational Cognitive Withdrawal Symptoms	Affective Withdrawal Symptoms	Physical Withdrawal Symptoms
1	23-29	8-9	7-9	8-9
2	30-45	10-14	10-12	10-13
3	46-50	15-16	13-15	14-15
4	51-58	17-19	16-18	16-17
5	59-67	20-23	19-21	18-20
6	68-77	24-27	22-25	21-24
7	78-92	28-32	26-29	25-30
8	93-97	33-35	30-33	31-34
9	98-104	36-38	34	35-37
10	105-115	39-40	35	38-40

Interpretation Guidelines

- Low scores (sten 1-4): individuals scoring within this range exhibit minimal withdrawal symptoms related to gaming, suggesting that gaming cessation does not significantly impact their emotional, cognitive, or physical functioning.
- Average scores (sten 5-6): this range indicates moderate withdrawal symptoms. Individuals in this category may experience some discomfort when reducing or ceasing gaming, but these effects are typically not severe.
- High scores (sten 7-10): individuals in this range report pronounced withdrawal symptoms, including strong motivational urges to continue gaming, heightened

emotional distress, and notable physical symptoms. These scores may indicate potential difficulties in regulating gaming behavior and warrant further psychological evaluation.

The same interpretive guidelines can be used to interpret scores on the 3 subscales (Motivational and Cognitive Withdrawal Symptoms, Affective Withdrawal Symptoms, and Physical

Withdrawal Symptoms), providing a more qualitative assessment of results.

Results of GWSQ

Table 9 presents the descriptive statistics for the total score and the 3 subscales of the GWSQ. The dataset includes responses from 565 participants.

Table 9. Descriptive statistics of the Gaming Withdrawal Symptoms Questionnaire (GWSQ).

	n	Minimum-maximum	Mean (SD)
Motivational and Cognitive Withdrawal Symptoms	565	8-40	19.46 (6.85)
Affective Withdrawal Symptoms	565	7-35	19.56 (6.77)
Physical Withdrawal Symptoms	565	8-40	17.53 (6.55)
Total GWSQ score	565	23-115	56.57 (18.11)

The Motivational and Cognitive Withdrawal Symptoms subscale yielded scores ranging from 8 to 40 (mean 19.47, SD 6.86). The Affective Withdrawal Symptoms subscale had a similar range, with scores between 7 and 35 (mean 19.56, SD 6.78). The Physical Withdrawal Symptoms subscale exhibited scores ranging from 8 to 48 (mean 17.54, SD 6.55).

The overall GWSQ score varied between 23 and 115, with a mean score of 56.57 (SD 18.12). These results suggest variability in the severity of withdrawal symptoms among respondents.

Discussion

Principal Findings

This study aimed to develop and validate a comprehensive withdrawal inventory of GD, addressing the lack of standardized diagnostic criteria and measurement tools. We applied a multistage approach combining the development, validation, and normalization of the GWSQ through 2 studies.

Summarizing the contribution, we developed the GWSQ based on collected data from the scoping literature review [27]. Initially, the questionnaire consisted of 47 items encompassing the full spectrum of withdrawal symptoms (affective, cognitive, behavioral, and physical components). Psychometric validation, conducted in 2 separate empirical studies, identified a 3-factor structure of the GWSQ. The CFA results confirmed the extracted structure, which consists of the following subscales: Motivational and Cognitive Withdrawal Symptoms, Affective Withdrawal Symptoms, Physical Withdrawal Symptoms ($\chi^2_{227}=887.8$; $P<.001$; comparative fit index=0.91; Tucker - Lewis index=0.90; root-mean-square error of approximation=0.072, 90% CI 0.067-0.077; standardized root-mean-square residual=0.049). The results of the reliability analysis support the validity of the construct (Cronbach $\alpha=0.95$; CR=0.95). Subsequently, the findings of the convergent and discriminant validity analysis confirm the relevance of the created tool. Therefore, the GWSQ is the first comprehensive and validated tool to measure withdrawal symptoms in GD.

The 3-factor structure of withdrawal symptoms identified in the GWSQ, comprising Affective Withdrawal Symptoms,

Motivational and Cognitive Withdrawal Symptoms, and Physical Withdrawal Symptoms, aligns with existing conceptualizations of behavioral addictions and withdrawal theories. The multidimensional approach to withdrawal symptomatology in GD is consistent with the “components model of addiction,” where withdrawal is defined as the spectrum of psychological states and physiological effects [58,59]. These findings support the multidimensional nature of withdrawal presented in previous studies [27,28,35,37]. Furthermore, compared to prior research on withdrawal symptoms in GD, the GWSQ allows a more comprehensive identification of symptoms and their structure. Unlike the unidimensional approach, which primarily focuses on the affective component of withdrawal in existing IGD screening tools (eg, sadness and irritability; IGDS9-SF [29], 20-Item Internet Gaming Disorder Test [49], and Internet Addiction Test [30]), the GWSQ integrates affective, motivational, cognitive, and physiological components, broadening the understanding of withdrawal experiences. The expanded classification of symptoms in this study suggests that previous assessment tools may have provided an incomplete representation of the full spectrum of withdrawal experiences in GD.

This study contributes to the ongoing debate and criticism regarding the presence of withdrawal symptoms in behavioral addictions [13-16]. Compared to substance use disorders, withdrawal in GD has a more psychological nature, encompassing affective symptoms (eg, irritability, frustration, anger, and impatience), cognitive symptoms (eg, difficulty concentrating and preoccupation with gaming), and behavioral and motivational symptoms (eg, lack of motivation and inability to identify activities to do). Nevertheless, our research also points to the presence of physiological symptoms, including headaches, muscle pain, and sweating. However, the physiological symptoms observed in gaming withdrawal appear to be milder than those associated with substance withdrawal. This may be linked to the psychosomatic nature of emotions and their multidimensional structure, which includes cognitive, behavioral, and physiological components [60,61].

Strengths of the Research and Its Implications

This study offers several methodological and conceptual advancements in the assessment of withdrawal symptoms in

GD. One of its key strengths is the comprehensive approach to symptom identification, which was based on a systematic review of the literature. This ensured that the conceptualization of withdrawal in GD was grounded in empirical evidence and reflected the multidimensional nature of the phenomenon.

Regarding the methodological aspect of the research, although both surveys were conducted online, their strength lies in the direct individual contact established with respondents. Specifically, we recruited participants directly from gaming-related social media platforms such as Facebook, Reddit, Discord, and Steam Community, ensuring that those who participated were active gamers. This allowed for a more focused and engaged sample, minimizing the risks associated with the use of commercial panels, such as nontargeted or less motivated participants. Moreover, the high quality of the data collected is reflected in several aspects: the high reliability of the measurement tools, the consistency of results across the 2 surveys, and the indicators of high participant attention during the evaluation. Moreover, while most of the respondents were from Poland, the data included participants from various nationalities and geographic backgrounds.

Another notable strength is the rigorous validation process, which included both EFA and CFA conducted on independent samples. This 2-step approach ensured the robustness of the identified factor structure and provided strong psychometric evidence supporting the validity of the GWSQ. The CFA results confirmed an adequate fit of the 3-factor model, reinforcing the theoretical basis of the proposed classification. In addition, the study introduces norm-based scoring (sten scores), which enhances the practical applicability of the tool. By providing normative data, the GWSQ allows for a more precise interpretation of individual scores, facilitating both clinical and research applications. This is particularly valuable for distinguishing between different levels of withdrawal severity, aiding in the early identification of GD, as well as serving as a marker for more severe cases of the disorder [38].

Additionally, the high internal consistency (Cronbach α and CR) and satisfactory construct validity of the GWSQ further attest to the reliability and accuracy of the developed measure. The observed positive correlations with established instruments assessing related constructs (eg, IGDS9-SF, PHQ-9, and GAD-7) and weaker or nonsignificant correlations with theoretically unrelated constructs (eg, SWLS and RYFF-18) support the tool's convergent and discriminant validity.

From a practical perspective, the GWSQ can be a valuable addition to existing GD screening tools, such as the IGDS9-SF [29], by enhancing epidemiological research, identifying risk factors, and tracking the progression of GD [38]. In clinical practice, it can help differentiate between problematic gaming and high engagement, enabling more accurate diagnoses and tailored intervention plans based on the severity of withdrawal symptoms [19-22]. Additionally, the GWSQ can monitor symptom reduction throughout treatment and assist in the early detection of individuals at risk of developing GD and cooccurring mental health conditions, such as depression and anxiety [19,39,40]. Moreover, as a potentially change-sensitive tool, the GWSQ holds immense promise as a patient-reported

outcome measure in formal clinical trials for future GD therapeutics, serving as a standardized end point for evaluating treatment efficacy. By providing a standardized measure of withdrawal symptoms, the GWSQ can inform the development of targeted prevention and treatment strategies, ultimately improving intervention outcomes [62].

Limitations and Future Directions

Despite its contributions, this study has certain limitations that should be considered when interpreting the findings and assessing their generalizability. First, the sample in both validation studies was predominantly Polish, which may limit the applicability of the results to other cultural and gaming populations. Although there were participants from other nations, their numbers were limited. Given that withdrawal experiences can be influenced by sociocultural factors which can impact social obligations, interpersonal gaming dynamics and habits, as well as norms surrounding digital leisure, emotional expression (particularly the emotional and physical expression of distress), power dynamics, and gaming motivations, this limitation is particularly important, as the English version of the GWSQ was administered to a predominantly nonnative English-speaking sample. Future studies should assess the psychometric properties of the GWSQ in diverse cultural contexts, especially among native English-speaking populations, to ensure its cross-cultural applicability and validity. Additionally, this study included a diverse gaming sample, with substantial variation in time spent on gaming and related activities. While this diversity increases ecological validity, it may also introduce interpretative challenges. Differences in gaming engagement levels could influence withdrawal symptom severity. Future research should consider more homogeneous subsamples or analyses by subgroup (eg, casual vs highly engaged players) to better capture the relationship between gaming behavior and withdrawal symptoms.

Moreover, the sten norms presented in Table 9 are based on the reference group that was representative of the general population by age, gender, and gaming habits. They provide a standardized interpretation of GWSQ scores, enabling comparisons between individuals. However, future work should focus on the applicability of these norms to clinical or high-risk groups, such as individuals with GD. Further studies are required to assess the applicability of these norms in clinical and treatment-seeking populations. Furthermore, replicating studies with more specific samples of players engaged in particular games or game mechanics could provide more concrete insights into their impact on withdrawal symptoms. Studies show that reward-based, competitive, and multiplayer genres, such as multiplayer online battle arena, first-person shooter, and massively multiplayer online role-playing game, are usually related to higher withdrawal symptoms than casual games [63,64]. Specific mechanics such as reward loops, ranking systems, “just one more turn/quest” behavior, unclear progression, or cliffhangers may trigger craving in situations of negative disengagement [65]. Future research should also investigate how specific withdrawal symptoms vary across different video game genres.

Subsequently, the study relied on self-report measures, which, although standard in psychometric research, are subject to recall bias and social-desirability effects [66]. Self-reported withdrawal symptoms may not fully capture the complexity of the phenomenon. As the study asked participants to reflect on a period in which they desired but were unable to play video games, it may not fully reflect the real-time dynamics of withdrawal. Future research could benefit from incorporating multimethod approaches and diverse study designs, including behavioral or physiological assessments and experimental methods, where participants attempt to withdraw from gaming under controlled conditions. This approach could provide more accurate insights into the withdrawal process. Furthermore, it should be acknowledged that IGDS9-SF, which was used as a primary validation tool, measures IGD—a related but distinct construct from GD. Therefore, correlations observed in the study with the GWSQ should be interpreted with this distinction in mind.

Building upon these considerations, additional directions for future research should aim to further explore the role of gaming withdrawal symptoms. First, longitudinal studies are needed to assess the temporal stability of withdrawal symptoms and the predictive validity of the GWSQ. The cross-sectional nature of this study does not allow conclusions about the persistence or fluctuation of symptoms over time. Examining how withdrawal symptoms evolve in relation to gaming habits, psychological states, and external influences will be essential for understanding their long-term impact. Additionally, future research should explore the underlying neurobiological and psychological mechanisms of gaming withdrawal. Neuroimaging studies, such as functional magnetic resonance imaging or electroencephalography, could provide insight into the brain activity associated with withdrawal symptoms, while

neurophysiological assessments (eg, autonomic nervous system measures) could help differentiate the psychological and biological components of withdrawal. Experimental designs manipulating gaming access could help establish causal relationships. Finally, qualitative studies exploring the lived experiences of individuals undergoing withdrawal could reveal psychological processes and coping strategies not captured through standard questionnaires.

Conclusions

This study introduces the first validated instrument specifically designed to assess withdrawal symptoms in GD, addressing limitations of existing measures that primarily focus on affective distress or general symptom severity. It provides empirical support for the multidimensional nature of gaming withdrawal by validating the GWSQ as a reliable and valid assessment tool. The findings highlight the importance of considering withdrawal as a core diagnostic feature of GD, demonstrating its affective, cognitive, motivational, and physiological components. The GWSQ's strong psychometric properties, confirmed through EFA and CFA, reinforce its use for both research and clinical applications.

From a practical standpoint, the GWSQ offers a standardized method for identifying withdrawal symptoms, aiding early detection, diagnosis, and treatment monitoring. Its application extends to epidemiological research, thereby enabling a better understanding of GD risk factors and progression. Furthermore, the tool has the potential to differentiate problematic gaming from high engagement and support clinicians in tailoring therapeutic approaches. Future research should focus on cross-cultural validation, longitudinal assessment of withdrawal symptoms, and the integration of biopsychosocial correlates to further enhance the use of the GWSQ.

Acknowledgments

The authors would like to thank Aleksandra Zajas for her valuable support with the statistical analysis. We also thank Izabela Pieńkowska for her assistance during the data collection process. Their contribution was significant to this research.

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the GAIDeT taxonomy [67], the following tasks were delegated to GenAI tools under full human supervision: proofreading and editing, quality assessment, and publication support. The GenAI tool used was ChatGPT (OpenAI) and Gemini (Google). Responsibility for the final manuscript lies entirely with the authors.

Data Availability

In line with the best practices, the preregistration for the studies is available for download on the Open Science Framework [68]. The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Funding

This research is part of the SONATA BIS-11 research project “The Transformation Process From Gaming Involvement to Gaming Disorder: Delineating Social and Motivational Antecedents From Consequences,” number 2021/42/E/HS6/00068, funded by the National Science Center (NCN) and performed at the Faculty of Management and Social Communication, Jagiellonian University, Kraków, Poland. The funder (NCN) had no involvement in the study design, data collection, analysis, interpretation of results, or the writing of the manuscript. The publication has been supported by a grant from the Faculty of Management and Social Communication under the Strategic Programme Excellence Initiative at Jagiellonian University.

Authors' Contributions

Conceptualization: SS, SŚ, PS

Methodology: SS, SŚ, PS

Computation: SS

Investigation: SS

Formal Analysis: SS

Data Visualization: SS

Writing—Original Draft: SS

Writing—Review & Editing: SŚ, JS, PS

Project Administration: SS

Resources: SS

Funding Acquisition: PS

Supervision: PS

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire and STROBE checklist.

[[DOCX File, 28 KB - games_v14i1e81407_app1.docx](#)]

References

1. Jovanovic B. Gamer demographics: facts and stats about the most popular hobby in the world. DataProt. URL: <https://blog.dataprot.net/gamer-demographics> [accessed 2023-04-11]
2. Castro-Calvo J, King DL, Stein DJ, Brand M, Carmi L, Chamberlain SR, et al. Expert appraisal of criteria for assessing gaming disorder: an international Delphi study. *Addiction* 2021;116(9):2463-2475 [FREE Full text] [doi: [10.1111/add.15411](https://doi.org/10.1111/add.15411)] [Medline: [33449441](https://pubmed.ncbi.nlm.nih.gov/33449441/)]
3. Billieux J, Flayelle M, Rumpf HJ, Stein DJ. High involvement versus pathological involvement in video games: a crucial distinction for ensuring the validity and utility of gaming disorder. *Curr Addict Rep* 2019;6(3):323-330. [doi: [10.1007/s40429-019-00259-x](https://doi.org/10.1007/s40429-019-00259-x)]
4. Jo YS, Bhang SY, Choi JS, Lee HK, Lee SY, Kweon Y. Clinical characteristics of diagnosis for internet gaming disorder: comparison of DSM-5 IGD and ICD-11 GD diagnosis. *J Clin Med* 2019;8(7):945 [FREE Full text] [doi: [10.3390/jcm8070945](https://doi.org/10.3390/jcm8070945)] [Medline: [31261841](https://pubmed.ncbi.nlm.nih.gov/31261841/)]
5. Männikkö N, Ruotsalainen H, Miettunen J, Pontes HM, Kääriäinen M. Problematic gaming behaviour and health-related outcomes: a systematic review and meta-analysis. *J Health Psychol* 2020;25(1):67-81. [doi: [10.1177/1359105317740414](https://doi.org/10.1177/1359105317740414)] [Medline: [29192524](https://pubmed.ncbi.nlm.nih.gov/29192524/)]
6. Long J, Liu T, Liu Y, Hao W, Maurage P, Billieux J. Prevalence and correlates of problematic online gaming: a systematic review of the evidence published in Chinese. *Curr Addict Rep* 2018;5(3):359-371. [doi: [10.1007/s40429-018-0219-6](https://doi.org/10.1007/s40429-018-0219-6)]
7. Lopez-Fernandez O, Honrubia-Serrano ML, Baguley T, Griffiths MD. Pathological video game playing in Spanish and British adolescents: towards the exploration of internet gaming disorder symptomatology. *Comput Hum Behav* 2014;41:304-312. [doi: [10.1016/j.chb.2014.10.011](https://doi.org/10.1016/j.chb.2014.10.011)]
8. Schivinski B, Brzozowska-Woś M, Buchanan EM, Griffiths MD, Pontes HM. Psychometric assessment of the internet gaming disorder diagnostic criteria: an item response theory study. *Addict Behav Rep* 2018;8:176-184 [FREE Full text] [doi: [10.1016/j.abrep.2018.06.004](https://doi.org/10.1016/j.abrep.2018.06.004)] [Medline: [30505924](https://pubmed.ncbi.nlm.nih.gov/30505924/)]
9. 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;106(1):205-212. [doi: [10.1111/j.1360-0443.2010.03104.x](https://doi.org/10.1111/j.1360-0443.2010.03104.x)] [Medline: [20840209](https://pubmed.ncbi.nlm.nih.gov/20840209/)]
10. Diagnostic and Statistical Manual of Mental Disorders 5th ed. Arlington, VA, US: American Psychiatric Publishing, Inc; 2013.
11. ICD-11: gaming disorder. World Health Organization. 2024. URL: <https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1448597234> [accessed 2026-03-10]
12. First MB, Yousif LH, Clarke DE, Wang PS, Gogtay N, Appelbaum PS. DSM-5-TR: overview of what's new and what's changed. *World Psychiatry* 2022;21(2):218-219 [FREE Full text] [doi: [10.1002/wps.20989](https://doi.org/10.1002/wps.20989)] [Medline: [35524596](https://pubmed.ncbi.nlm.nih.gov/35524596/)]
13. Pontes HM, Schivinski B, Sindermann C, Li M, Becker B, Zhou M, et al. Measurement and conceptualization of gaming disorder according to the World Health Organization framework: the development of the gaming disorder test. *Int J Ment Health Addict* 2019;19(2):508-528. [doi: [10.1007/s11469-019-00088-z](https://doi.org/10.1007/s11469-019-00088-z)]

14. Griffiths MD, van Rooij AJ, Kardefelt-Winther D, Starcevic V, Király O, Pallesen S, et al. Working towards an international consensus on criteria for assessing internet gaming disorder: a critical commentary on Petry et al. (2014). *Addiction* 2016;111(1):167-175 [FREE Full text] [doi: [10.1111/add.13057](https://doi.org/10.1111/add.13057)] [Medline: [26669530](https://pubmed.ncbi.nlm.nih.gov/26669530/)]
15. Petry NM, Rehbein F, Gentile DA, Lemmens JS, Rumpf H, Mößle T, et al. An international consensus for assessing internet gaming disorder using the new DSM-5 approach. *Addiction* 2014;109(9):1399-1406. [doi: [10.1111/add.12457](https://doi.org/10.1111/add.12457)] [Medline: [24456155](https://pubmed.ncbi.nlm.nih.gov/24456155/)]
16. Starcevic V. Tolerance and withdrawal symptoms may not be helpful to enhance understanding of behavioural addictions. *Addiction* 2016;111(7):1307-1308. [doi: [10.1111/add.13381](https://doi.org/10.1111/add.13381)] [Medline: [27095413](https://pubmed.ncbi.nlm.nih.gov/27095413/)]
17. Deleuze J, Christiaens M, Nuyens F, Billieux J. Shoot at first sight! First person shooter players display reduced reaction time and compromised inhibitory control in comparison to other video game players. *Comput Hum Behav* 2017;72:570-576. [doi: [10.1016/j.chb.2017.02.027](https://doi.org/10.1016/j.chb.2017.02.027)]
18. Kuss DJ, Griffiths MD, Pontes HM. Chaos and confusion in DSM-5 diagnosis of internet gaming disorder: issues, concerns, and recommendations for clarity in the field. *J Behav Addict* 2017;6(2):103-109 [FREE Full text] [doi: [10.1556/2006.5.2016.062](https://doi.org/10.1556/2006.5.2016.062)] [Medline: [27599673](https://pubmed.ncbi.nlm.nih.gov/27599673/)]
19. Chew PKH, Au JQM. Internet gaming disorder in the DSM-5: engagement vs addiction. *J Technol Behav Sci* 2023;9(4):652-660. [doi: [10.1007/s41347-023-00364-9](https://doi.org/10.1007/s41347-023-00364-9)]
20. Infanti A, Valls-Serrano C, Perales J, Vögele C, Billieux J. Gaming passion contributes to the definition and identification of problematic gaming. *Addict Behav* 2023;147:107805 [FREE Full text] [doi: [10.1016/j.addbeh.2023.107805](https://doi.org/10.1016/j.addbeh.2023.107805)] [Medline: [37523871](https://pubmed.ncbi.nlm.nih.gov/37523871/)]
21. Li Y, Mu W, Xie X, Kwok SY. Network analysis of internet gaming disorder, problematic social media use, problematic smartphone use, psychological distress, and meaning in life among adolescents. *Digit Health* 2023;9 [FREE Full text] [doi: [10.1177/20552076231158036](https://doi.org/10.1177/20552076231158036)]
22. Tang MC, Ebrahimi OV, Cheng C. Symptomatology, risk, and protective factors of gaming disorder: a network analytical approach. *Comput Hum Behav* 2023;148:107899. [doi: [10.1016/j.chb.2023.107899](https://doi.org/10.1016/j.chb.2023.107899)]
23. Yen J, Lin P, Wu H, Ko C. The withdrawal-related affective, gaming urge, and anhedonia symptoms of internet gaming disorder during abstinence. *J Behav Addict* 2022 [FREE Full text] [doi: [10.1556/2006.2022.00008](https://doi.org/10.1556/2006.2022.00008)]
24. Brown RIF. A theoretical model of the behavioural addictions-applied to offending. In: Hodge JE, McMurran M, Hollins CR, editors. *Addicted Crime*. Chichester, UK: John Wiley; 1997:13-65.
25. Charlton JP, Danforth IDW. Distinguishing addiction and high engagement in the context of online game playing. *Comput Hum Behav* 2007;23(3):1531-1548. [doi: [10.1016/j.chb.2005.07.002](https://doi.org/10.1016/j.chb.2005.07.002)]
26. Kardefelt-Winther D. A critical account of DSM-5 criteria for internet gaming disorder. *Addict Res Theory* 2014;23(2):93-98. [doi: [10.3109/16066359.2014.935350](https://doi.org/10.3109/16066359.2014.935350)]
27. Starzec S, Starosta J, Zajas A, Kiszka P, Śpiewak S, Strojny P. Scoping review of withdrawal's role in contemporary gaming disorder research: conceptualizations and operationalizations. *Clin Psychol Rev* 2024;114:102478 [FREE Full text] [doi: [10.1016/j.cpr.2024.102478](https://doi.org/10.1016/j.cpr.2024.102478)] [Medline: [39236625](https://pubmed.ncbi.nlm.nih.gov/39236625/)]
28. Kaptis D, King DL, Delfabbro PH, Gradisar M. Withdrawal symptoms in internet gaming disorder: a systematic review. *Clin Psychol Rev* 2016;43:58-66. [doi: [10.1016/j.cpr.2015.11.006](https://doi.org/10.1016/j.cpr.2015.11.006)] [Medline: [26704173](https://pubmed.ncbi.nlm.nih.gov/26704173/)]
29. Pontes HM, Griffiths MD. Measuring DSM-5 internet gaming disorder: development and validation of a short psychometric scale. *Comput Hum Behav* 2015;45:137-143. [doi: [10.1016/j.chb.2014.12.006](https://doi.org/10.1016/j.chb.2014.12.006)]
30. Young K. *Caught in The Net: How to Recognize the Signs of Internet Addiction and a Winning Strategy for Recovery*. New Jersey: John Wiley & Sons; 1998.
31. Meerkerk G, Van Den Eijnden RJM, Vermulst AA, Garretsen HFL. The Compulsive Internet Use Scale (CIUS): some psychometric properties. *Cyberpsychol Behav* 2009;12(1):1-6. [doi: [10.1089/cpb.2008.0181](https://doi.org/10.1089/cpb.2008.0181)] [Medline: [19072079](https://pubmed.ncbi.nlm.nih.gov/19072079/)]
32. Király O, Slezcka P, Pontes HM, Urbán R, Griffiths MD, Demetrovics Z. Validation of the ten-item Internet Gaming Disorder Test (IGDT-10) and evaluation of the nine DSM-5 internet gaming disorder criteria. *Addict Behav* 2017;64:253-260. [doi: [10.1016/j.addbeh.2015.11.005](https://doi.org/10.1016/j.addbeh.2015.11.005)] [Medline: [26632194](https://pubmed.ncbi.nlm.nih.gov/26632194/)]
33. Chen SH, Weng LJ, Su YJ, Wu HM, Yang PF. Development of a Chinese Internet Addiction Scale and its psychometric study. *Chin J Psychol* 2003 [FREE Full text]
34. Holstein BE, Pedersen TP, Bendtsen P, Madsen KR, Meilstrup CR, Nielsen L, et al. Perceived problems with computer gaming and internet use among adolescents: measurement tool for non-clinical survey studies. *BMC Public Health* 2014;14:361 [FREE Full text] [doi: [10.1186/1471-2458-14-361](https://doi.org/10.1186/1471-2458-14-361)] [Medline: [24731270](https://pubmed.ncbi.nlm.nih.gov/24731270/)]
35. Kaptis D, King DL, Delfabbro PH, Gradisar M. Trajectories of abstinence-induced internet gaming withdrawal symptoms: a prospective pilot study. *Addict Behav Rep* 2016;4:24-30 [FREE Full text] [doi: [10.1016/j.abrep.2016.06.002](https://doi.org/10.1016/j.abrep.2016.06.002)] [Medline: [29511720](https://pubmed.ncbi.nlm.nih.gov/29511720/)]
36. Flannery BA, Volpicelli JR, Pettinati HM. Psychometric properties of the Penn Alcohol Craving Scale. *Alcohol Clin Exp Res* 2006;23(8):1289-1295. [doi: [10.1111/j.1530-0277.1999.tb04349.x](https://doi.org/10.1111/j.1530-0277.1999.tb04349.x)]
37. Giordano AL, Prosek EA, Bain C, Malacara A, Turner J, Schunemann K. Withdrawal symptoms among American collegiate internet gamers. *J Ment Health Couns* 2020;42(1):63-77. [doi: [10.17744/mehc.42.1.05](https://doi.org/10.17744/mehc.42.1.05)]

38. Lee S, Lee HK, Jeong H, Yim HW, Bhang S, Jo S, et al. The hierarchical implications of internet gaming disorder criteria: which indicate more severe pathology? *Psychiatry Investig* 2017;14(3):249-259 [FREE Full text] [doi: [10.4306/pi.2017.14.3.249](https://doi.org/10.4306/pi.2017.14.3.249)] [Medline: [28539943](https://pubmed.ncbi.nlm.nih.gov/28539943/)]
39. Jeong H, Yim HW, Lee SY, Lee HK, Potenza MN, Jo SJ, et al. Reciprocal relationship between depression and Internet gaming disorder in children: a 12-month follow-up of the iCURE study using cross-lagged path analysis. *J Behav Addict* 2019;8(4):725-732 [FREE Full text] [doi: [10.1556/2006.8.2019.74](https://doi.org/10.1556/2006.8.2019.74)] [Medline: [32359239](https://pubmed.ncbi.nlm.nih.gov/32359239/)]
40. Zhang MX, Kam CCS, Wu AM. The reciprocity between psychological need frustration and adolescent problematic smartphone use. *J Appl Dev Psychol* 2024;91:101634. [doi: [10.1016/j.appdev.2024.101634](https://doi.org/10.1016/j.appdev.2024.101634)]
41. Gagnier JJ, Lai J, Morkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res* 2021;30(8):2197-2218. [doi: [10.1007/s11136-021-02822-4](https://doi.org/10.1007/s11136-021-02822-4)] [Medline: [33818733](https://pubmed.ncbi.nlm.nih.gov/33818733/)]
42. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Int J Surg* 2014;12(12):1495-1499 [FREE Full text] [doi: [10.1016/j.ijvsu.2014.07.013](https://doi.org/10.1016/j.ijvsu.2014.07.013)] [Medline: [25046131](https://pubmed.ncbi.nlm.nih.gov/25046131/)]
43. Qualtrics. URL: <https://qualtrics.com> [accessed 2026-03-12]
44. Alhamoud M, Alkhalifah A, Althunyan A, Mustafa T, Alqahtani H, Awad F. Internet gaming disorder: its prevalence and associated gaming behavior, anxiety, and depression among high school male students, Dammam, Saudi Arabia. *J Fam Community Med* 2022;29(2):93-101 [FREE Full text] [doi: [10.4103/jfcm.jfcm_48_22](https://doi.org/10.4103/jfcm.jfcm_48_22)] [Medline: [35754755](https://pubmed.ncbi.nlm.nih.gov/35754755/)]
45. González-Cabrera J, Basterra-González A, Ortega-Barón J, Caba-Machado V, Díaz-López A, Pontes HM, et al. Loot box purchases and their relationship with internet gaming disorder and online gambling disorder in adolescents: a prospective study. *Comput Hum Behav* 2023;143:107685. [doi: [10.1016/j.chb.2023.107685](https://doi.org/10.1016/j.chb.2023.107685)]
46. Hair JF. Multivariate data analysis. DigitalCommons. 2009. URL: <https://digitalcommons.kennesaw.edu/facpubs/2925/> [accessed 2025-10-24]
47. Ho R. Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS. Boca Raton: Chapman and Hall/CRC; 2006.
48. Strojny P, Kiszka P, Starosta J, Szyszka PD, Starzec S, Winiarska A. It's not just about how long you play. Indirect gaming involvement and genre preferences in predicting gaming disorder risk: evidence from preregistered studies. *Front Psychiatry* 2023 [FREE Full text] [doi: [10.3389/fpsy.2023.1230774](https://doi.org/10.3389/fpsy.2023.1230774)]
49. Pontes HM, Király O, Demetrovics Z, Griffiths MD. The conceptualisation and measurement of DSM-5 internet gaming disorder: the development of the IGD-20 test. *PLoS One* 2014;9(10):e110137 [FREE Full text] [doi: [10.1371/journal.pone.0110137](https://doi.org/10.1371/journal.pone.0110137)] [Medline: [25313515](https://pubmed.ncbi.nlm.nih.gov/25313515/)]
50. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
51. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV. Washington, DC: American Psychiatric Association; 1994.
52. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
53. Király O, Billieux J, King DL, Urbán R, Koncz P, Polgár E, et al. A comprehensive model to understand and assess the motivational background of video game use: the Gaming Motivation Inventory (GMI). *J Behav Addict* 2022;11(3):796-819 [FREE Full text] [doi: [10.1556/2006.2022.00048](https://doi.org/10.1556/2006.2022.00048)] [Medline: [35939353](https://pubmed.ncbi.nlm.nih.gov/35939353/)]
54. Diener E, Emmons RA, Larsen RJ, Griffin S. The Satisfaction With Life Scale. *J Pers Assess* 1985;49(1):71-75. [doi: [10.1207/s15327752jpa4901_13](https://doi.org/10.1207/s15327752jpa4901_13)] [Medline: [16367493](https://pubmed.ncbi.nlm.nih.gov/16367493/)]
55. Ryff CD. Psychological well-being revisited: advances in the science and practice of eudaimonia. *Psychother Psychosom* 2014;83(1):10-28 [FREE Full text] [doi: [10.1159/000353263](https://doi.org/10.1159/000353263)] [Medline: [24281296](https://pubmed.ncbi.nlm.nih.gov/24281296/)]
56. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J* 1999;6(1):1-55. [doi: [10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)]
57. Brzeziński J. Metodologia badań psychologicznych: Wydanie nowe. Wydawnictwo Naukowe PWN. 2023. URL: <https://scholar.google.com/scholar?cluster=12695081681016493748&hl=en&oi=scholar> [accessed 2025-02-14]
58. Griffiths M. A 'components' model of addiction within a biopsychosocial framework. *J Subst Use* 2009;10(4):191-197. [doi: [10.1080/14659890500114359](https://doi.org/10.1080/14659890500114359)]
59. Griffiths MD. The role of context in online gaming excess and addiction: some case study evidence. *Int J Ment Health Addict* 2009;8(1):119-125. [doi: [10.1007/s11469-009-9229-x](https://doi.org/10.1007/s11469-009-9229-x)]
60. Ekman PE, Davidson RJ. The Nature of Emotion: Fundamental Questions. Oxford: Oxford University Press; 1994.
61. LeDoux J. Mózgowe interakcje poznawczo-emocjonalne. In: Ekman P, Davidson RJ, editors. *Natura Emocji*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne; 1998:190-197.
62. Wang Z. Diagnosis of internet gaming disorder: issues, impact, and recommendations. *EHSS* 2023;8:2247-2252. [doi: [10.54097/ehss.v8i.4684](https://doi.org/10.54097/ehss.v8i.4684)]
63. Jo H. How do different game genres impact adolescent players' problematic game use? *Entertain Comput* 2025;52:100897. [doi: [10.1016/j.entcom.2024.100897](https://doi.org/10.1016/j.entcom.2024.100897)]

64. Bao Y, Zhou H, Geng F, Hu Y. The relation between game disorder and interruption during game is mediated by game craving. *Front Psychol* 2025;16 [FREE Full text] [doi: [10.3389/fpsyg.2025.1579016](https://doi.org/10.3389/fpsyg.2025.1579016)]
65. Alexandrovsky D, Gerling K, Opp MS, Hahn CB, Birk MV, Alsheail M. Disengagement from games: characterizing the experience and process of exiting play sessions. *Proc ACM Hum-Comput Interact* 2024;8(CHI PLAY):1-27. [doi: [10.1145/3677066](https://doi.org/10.1145/3677066)]
66. Freichel R, Christensen E, Mrkonja L, de Jong PJ, Cousijn J, Franken I, et al. Attentional biases and their association with substance-use-related problems and addictive behaviors: the utility of a gamified value-modulated attentional capture task. *Addict Behav Rep* 2024;19:100534 [FREE Full text] [doi: [10.1016/j.abrep.2024.100534](https://doi.org/10.1016/j.abrep.2024.100534)] [Medline: [38404750](https://pubmed.ncbi.nlm.nih.gov/38404750/)]
67. Suchikova Y, Tsybuliak N, Teixeira da Silva JA, Nazarovets S. GAIDeT (Generative AI Delegation Taxonomy): a taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. *Account Res* 2026;33(3):2544331. [doi: [10.1080/08989621.2025.2544331](https://doi.org/10.1080/08989621.2025.2544331)] [Medline: [40781729](https://pubmed.ncbi.nlm.nih.gov/40781729/)]
68. Open Science Framework. URL: <https://osf.io/evyu9> [accessed 2026-03-13]

Abbreviations

APA: American Psychological Association
AVE: average variance extracted
CFA: confirmatory factor analysis
COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments
CR: composite reliability
DSM-5 TR: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
EFA: exploratory factor analysis
F1: motivational and cognitive factor
F2: affective factor
F3: physical factor
GAD-7: Generalized Anxiety Disorder-7 items
GD: gaming disorder
GIS: Gaming Involvement Scale
GMI: Gaming Motivation Inventory
GWSQ: Gaming Withdrawal Symptoms Questionnaire
ICD-11: International Classification of Diseases, Eleventh Edition
IGD: internet gaming disorder
IGDS9-SF: Internet Gaming Disorder Scale-Short Form
PHQ-9: Patient Health Questionnaire-9 items
RYFF-18: Ryff Psychological Well-Being Scale-18 items
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
SWLS: Satisfaction With Life Scale

Edited by S Brini; submitted 30.Jul.2025; peer-reviewed by P Larionow, B Imbert; comments to author 17.Oct.2025; revised version received 14.Jan.2026; accepted 15.Jan.2026; published 31.Mar.2026.

Please cite as:

Starzec S, Śpiewak S, Starosta J, Strojny P
Conceptualization and Measurement of Withdrawal Symptoms in Gaming Disorder and Development and Psychometric Validation of the Gaming Withdrawal Symptoms Questionnaire: Cross-Sectional Study
JMIR Serious Games 2026;14:e81407
URL: <https://games.jmir.org/2026/1/e81407>
doi: [10.2196/81407](https://doi.org/10.2196/81407)
PMID:

©Sylwia Starzec, Sławomir Śpiewak, Jolanta Starosta, Paweł Strojny. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 31.Mar.2026. 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.

Original Paper

Reciprocal Within-Person Dynamics Between Internet Gaming Disorder Symptoms, Physical Activity, and Loneliness Among Chinese Adolescent Gamers: Three-Wave Prospective Cohort Study

Chang Hu^{1,2}, PhD; Xuan Meng³, MS; Xi Wang⁴, MS; Jilin Huang⁵, PhD

¹Professional Tennis College, Wuhan City Polytechnic, Wuhan, Hubei, China

²School of Physical Education, Jiangxi Normal University, Nanchang, Jiangxi, China

³Nanxishan Hospital of Guangxi Zhuang Autonomous Region, Guilin, Guangxi, China

⁴General Education Department, Wuhan College, Wuhan, Hubei, China

⁵Faculty of General Education, Hubei College of Chinese Medicine, Jingzhou, Hubei, China

Corresponding Author:

Jilin Huang, PhD

Faculty of General Education

Hubei College of Chinese Medicine

No 4, Renmin Road

Jingzhou, Hubei, 434020

China

Phone: 86 15926516619

Email: H15926516619@163.com

Abstract

Background: Problematic gaming and internet gaming disorder (IGD) symptoms are prevalent in adolescence, yet the longitudinal interplay between physical activity (PA), loneliness, and IGD symptoms, as well as potential sexual differences, remains unclear.

Objective: This study aimed to examine reciprocal within-person associations between PA, loneliness, and IGD symptoms among Chinese adolescent gamers and test sexual differences in these associations.

Methods: We conducted a three-wave prospective cohort study among 1332 Chinese adolescents selected using convenience sampling from five middle schools in central China who had engaged in online gaming during the previous year ($n=441$, 33.1%, females; mean age 3.64, SD 0.76, years). PA, loneliness, and IGD symptoms were assessed using the Physical Activity Rating Scale-3, the three-item short form of the University of California, Los Angeles Loneliness Scale, and the nine-item *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* IGD Checklist, respectively, at three 6-month intervals: wave 1 (baseline, T1), wave 2 (T2), and wave 3 (T3). A random intercept (RI) cross-lagged panel model (CLPM) and multigroup analyses were conducted.

Results: RI-CLPM indicated that at the within-person level, higher PA predicted decreased subsequent loneliness (T1→T2: $\beta=-0.12$, 95% CI -0.16 to -0.08 ; T2→T3: $\beta=-0.13$, 95% CI -0.17 to -0.08) and IGD symptoms (T1→T2: $\beta=-0.10$, 95% CI -0.16 to -0.04 , $P=.009$; T2→T3: $\beta=-0.10$, 95% CI -0.16 to -0.04). Increased loneliness ($\beta=-0.22$, 95% CI -0.26 to -0.18) and increased IGD symptoms ($\beta=-0.08$, 95% CI -0.12 to -0.04) each predicted later reductions in PA, indicating a mutually reinforcing cycle. Increased loneliness also predicted an increase in subsequent IGD symptoms (T1→T2: $\beta=0.14$, 95% CI 0.11 - 0.17), and increased IGD symptoms, in turn, predicted greater loneliness in the next wave (T1→T2: $\beta=0.18$, 95% CI 0.14 - 0.22). Multigroup models indicated that the protective effect of PA on later loneliness and the prospective effect of loneliness on subsequent IGD symptoms were stronger among girls than boys. In males, PA significantly predicted loneliness ($\beta=-0.08$ to -0.09 , 95% CI -0.15 to -0.01), and loneliness significantly predicted IGD symptoms ($\beta=0.09$, 95% CI 0.03 - 0.15). In females, the cross-lagged effects from PA to loneliness were significant but stronger ($\beta=-0.17$, 95% CI -0.23 to -0.11), and the paths from loneliness to later IGD symptoms were also significant ($\beta=0.16$, 95% CI 0.11 - 0.21).

Conclusions: PA, loneliness, and IGD symptoms are reciprocally linked in adolescent gamers. By using RI-CLPM to distinguish within-person changes from stable between-person differences, this study extends prior research based mainly on cross-sectional designs or traditional CLPMs and provides a clearer understanding of the dynamic interplay among behavioral, emotional, and

gaming-related factors. The findings highlight that interventions aiming to prevent IGD symptoms should simultaneously promote PA and reduce loneliness, with particular attention to sex-specific patterns, especially in girls.

(*JMIR Serious Games* 2026;14:e87847) doi:[10.2196/87847](https://doi.org/10.2196/87847)

KEYWORDS

Chinese adolescents; sexual differences; internet gaming disorder symptoms; loneliness; reciprocal association; physical activity

Introduction

Background

Internet gaming disorder (IGD) is a persistent and maladaptive pattern of online gaming associated with significant functional impairments and is increasingly recognized as a major public health concern [1]. IGD is characterized by impaired control over gaming, prioritization of gaming over other activities despite negative consequences, and persistent gaming behavior despite functional impairment [2]. In population-based research, these manifestations are commonly examined in terms of IGD symptoms, which reflect the severity of problematic gaming behaviors [3-7]. Higher levels of IGD symptoms have been linked to a range of detrimental outcomes, including psychological difficulties, behavioral problems, and social consequences [8-11]. Given the widespread accessibility of online games and the developmental vulnerability of adolescence [12,13], understanding the longitudinal mechanisms underlying IGD symptoms and identifying relevant risk and protective factors during this period is particularly important. However, the dynamic interplay among behavioral and emotional factors related to IGD symptoms remains insufficiently understood.

Longitudinal Association Between Physical Activity and IGD Symptoms

The displacement hypothesis posits that time is limited, so engagement in one activity reduces the time available for others [14]. In this study, this mainly concerned adolescents' leisure time outside of structured school activities. In Chinese schools, mobile phone use during class is restricted, and gaming typically occurs after school or during other leisure periods [15]. Therefore, greater engagement in physical activity (PA) may reduce opportunities for extended gaming and, in turn, lower the risk for IGD symptoms [16,17]. In addition, PA may promote healthier lifestyle routines, better emotional regulation, and alternative sources of enjoyment and social interaction, thereby reducing adolescents' reliance on gaming as a primary leisure activity [18-20]. At the same time, this relationship may be bidirectional. Excessive gaming or more severe IGD symptoms may reduce participation in PA by consuming leisure time, lowering motivation for physical movement, and disrupting daily routines or social patterns [21-23]. Taken together, these mechanisms suggest a potentially bidirectional relationship between PA and IGD symptoms.

Accumulating empirical evidence indicates a negative association between PA and IGD symptoms. Cross-sectional investigations across diverse cultural contexts, such as Australia, Taiwan, Portugal, Germany, and China, have consistently shown that higher PA is associated with decreased IGD symptoms [16,17,24-26]. Systematic reviews and meta-analyses have

further indicated that exercise-based interventions can reduce gaming addiction scores and may be among the most effective interventions for adolescent internet addiction [18,27]. Longitudinal findings have provided further, although mixed, evidence. A Swiss cohort study found bidirectional associations between video gaming disorder symptoms and sport or exercise participation over 15 months [22], and another longitudinal study reported that higher PA is associated with decreased IGD symptoms [28]. However, other studies have found that IGD symptoms predict later PA but PA does not predict later IGD symptoms [29] or that gaming time predicts PA trajectories but not vice versa [30]. Notably, most prior longitudinal studies have relied on the traditional cross-lagged panel model (CLPM), which cannot disentangle within-person fluctuations from stable between-person differences [31]. Consequently, the within-person longitudinal association between PA and IGD symptoms remains insufficiently understood.

Longitudinal Association Between Loneliness and IGD Symptoms

Loneliness is a distressing emotional state arising from the perceived discrepancy between desired and actual social relationships [32]. Rather than reflecting objective social isolation, it represents a subjective sense of lacking meaningful connections with others [33]. According to the model of compensatory internet use [34], negative life situations may motivate individuals to go online to regulate negative emotions, and problematic online behaviors arise from attempts to compensate for adverse offline experiences [34]. Online gaming may temporarily relieve loneliness by providing social interaction, achievement, and escape from a negative affect [35]. However, when such compensatory use becomes habitual, it may develop into maladaptive patterns characteristic of increased IGD symptoms [36]. Over time, reliance on gaming as a coping strategy may reduce opportunities for meaningful offline interactions and further intensify loneliness [33]. Likewise, increased IGD symptoms may increase loneliness by reducing real-life social engagement, substituting online interactions for offline relationships, and fostering social withdrawal [37]. These mechanisms suggest a dynamic, bidirectional relationship between loneliness and IGD symptoms.

Empirical studies generally support this perspective. Cross-sectional studies across university and adolescent samples have consistently shown a positive association between loneliness and IGD symptoms [38-42]. Meta-analytic evidence has further identified loneliness as a risk factor of IGD symptoms and internet use disorder symptoms more broadly [43,44]. Longitudinal research has also provided support. Baseline loneliness significantly predicted later IGD symptoms among Chinese university students [45], loneliness predicted

increases in gaming problems at both within-person and between-person levels [46], and a two-wave longitudinal study of Dutch adolescent gamers identified loneliness to be both an antecedent and a consequence of pathological gaming [47]. However, most longitudinal studies have used traditional structural equation modeling (SEM) or CLPM, both of which are limited in distinguishing within-person fluctuations from stable between-person differences [48]. As a result, the dynamic longitudinal association between loneliness and IGD symptoms remains insufficiently understood. Furthermore, much of the existing evidence focuses on general internet addiction rather than IGD symptoms specifically, and research among adolescent gamers remains relatively scarce.

Longitudinal Association Between PA and Loneliness

The bidirectional association between PA and loneliness can be understood within the framework of basic psychological needs theory (BPNT), a subtheory of self-determination theory [49]. BPNT posits that the fulfillment of three innate psychological needs, namely autonomy, competence, and relatedness, is essential for psychological well-being [50,51]. In this context, PA may help satisfy the need for relatedness by providing opportunities for social interaction, belonging, and shared achievement [52]. Regular participation in exercise or group-based sports may thus alleviate loneliness through enhanced social connectedness and emotional support [53]. Conversely, persistent loneliness may undermine these needs by reducing social motivation and self-regulatory capacity, thereby decreasing engagement in PA over time [54]. This reciprocal process suggests that PA and loneliness may influence each other dynamically across time.

Empirical findings generally support this proposition. Cross-sectional research consistently reveals a negative association between PA and loneliness across diverse populations [55-57]. Systematic reviews support this claim [58,59]. Longitudinal evidence further reports the bidirectional association between PA and loneliness among adults and older adults [60,61]. A recent study using a random intercept cross-lagged panel model (RI-CLPM) further demonstrated small but significant within-person bidirectional effects between PA and loneliness [62]. However, most existing longitudinal studies have focused on adult or older populations, with adolescents receiving far less attention. Additionally, many studies have relied on traditional CLPMs. Moreover, relatively few studies have simultaneously examined PA, loneliness, and IGD symptoms within an integrated longitudinal framework. To address these limitations, this study used an RI-CLPM to clarify the bidirectional relationship between PA and loneliness among adolescents.

Potential Sexual Differences

Emerging evidence suggests that sex may moderate the associations among PA, loneliness, and IGD symptoms. Epidemiological studies consistently show that male adolescents have a higher prevalence of IGD than females [63,64], and males with more severe IGD symptoms appear to be more sensitive to gaming-related rewards, whereas females tend to exhibit stronger affective dysregulation and mood-related symptoms [64]. Sexual differences are also evident in PA levels,

with males generally reporting higher engagement in PA than females [65]. Notably, when PA levels are low, males report slightly higher internet addiction than females, but as PA increases, males show a sharper decline in addiction symptoms, suggesting that the protective role of PA may be particularly pronounced among males [65]. At the same time, loneliness and its correlates may operate differently across sexes. Prior studies have shown that the correlation between loneliness and internet addiction tends to be weaker among males [30]. Furthermore, longitudinal evidence shows that moderate-to-vigorous PA negatively predicts gaming time among males but not females [30]. Moreover, sexual differences have been observed in the link between PA and psychological distress [66], where the negative association between PA and loneliness is more pronounced among females [67]. These findings indicate that sex may moderate the associations among PA, loneliness, and IGD symptoms. However, few longitudinal studies have examined whether such sex-specific patterns remain stable over time in youth populations.

Aims of This Study

To address these gaps, this study adopted a three-wave longitudinal design and an RI-CLPM to examine the temporal associations between PA, loneliness, and IGD symptoms among Chinese adolescent gamers. By distinguishing stable between-person differences from within-person fluctuations, this approach clarified the dynamic reciprocal relationships among these variables over time. Four hypotheses were proposed:

- Hypothesis (H)1: PA and IGD symptoms are reciprocally associated.
- H2: Loneliness and IGD symptoms are reciprocally associated.
- H3: PA and loneliness are reciprocally associated.
- H4: These within-person reciprocal associations differ between male and female adolescents.

Methods

Study Design

This 1-year longitudinal study was conducted from March 2024 to March 2025, with data collected at three 6-month intervals: wave 1 (baseline [T1], March 2024), wave 2 (T2, September 2024), and wave 3 (T3, March 2025).

Participants and Procedures

Inclusion and Exclusion Criteria

The target population comprised adolescents who had engaged in online gaming during the preceding 12 months, thereby ensuring that all participants were active gamers rather than nongaming peers. Eligible participants were students from grades 7 and 8. Grade 9 students were excluded because they typically prepare for the high school entrance examination and may graduate or transfer to different schools, which would make longitudinal follow-up difficult. Additionally, focusing on grades 7 and 8 allowed us to capture adolescents in early adolescence, a critical developmental stage for the emergence of problematic gaming behaviors and social-emotional changes [11,68].

Sampling Procedures

The study was conducted in a school-based educational setting, and all participating schools were urban public middle schools located in central China. Participants were recruited using convenience sampling in collaboration with school administrators, and recruitment procedures included announcements in classrooms and direct invitations by trained field researchers.

Participant Characteristics

At baseline (T1), 1332 adolescents ($n=441$, 33.1%, females; mean age 13.64, SD 0.76, years) who met the inclusion criteria completed the survey. By T2, 1240 (93.1%) students had contributed two waves of data. By the final wave (T3), 1138 students remained in the study, yielding a retention rate of 85.4%. Participant attrition across waves was mainly attributable to common circumstances in school-based longitudinal surveys, such as student absence on the day of data collection, transfers to other schools during the follow-up period, or incomplete questionnaires that could not be matched across waves.

Sample Size, Power, and Precision

The sample size was determined using the number of eligible students available in the participating schools during the recruitment period. Given the longitudinal design and the use of an RI-CLPM, the available sample size was considered adequate to support estimation of within-person longitudinal associations. In addition, all participants who provided valid data at least once were retained in the main analyses to maximize statistical power and reduce bias associated with attrition.

Ethical Considerations

Ethical approval for the study was obtained from the Institutional Ethics Committee of Wuhan College (approval #20240306). Prior to each survey administration, students and their parents (or legal guardians) were fully informed about the purpose of the study, data collection procedures, and measures taken to guarantee confidentiality. Parents/guardians and students provided written informed consent. Participation was strictly voluntary, and both students and parents/guardians were assured that refusal or withdrawal would not have any negative consequences on the students' academic records or school evaluations. To enable longitudinal matching of responses across waves, students were asked to provide their school IDs, which were encrypted, stored securely, and removed from the analytical dataset once the linkage was completed, thereby preserving anonymity. An information sheet describing participants' rights was distributed prior to survey completion. Participants did not receive any financial compensation for participation. No identifiable images or personal information of participants are included in the manuscript or [Multimedia Appendix 1](#).

Measures and Covariates

Background Variables

Background information collected included age, sex, self-reported academic performance, parental educational level, perceived family financial situations, single-parent family status, and gaming time.

IGD Symptoms

IGD symptoms were measured using the nine-item *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (DSM-5) IGD Checklist [2], which assesses the presence of core IGD symptoms over the past 12 months (yes/no response). The nine criteria are preoccupation with gaming, withdrawal, tolerance, loss of control, giving priority to gaming over other activities, persistence despite negative consequences, deception regarding gaming time, escapism, and functional impairment due to gaming. Although the DSM-5 proposes a diagnostic threshold of five or more symptoms for identifying probable IGD, this study treated IGD as a continuous indicator of symptom severity, which is common practice in population-based research examining variability in problematic gaming behaviors [3-7]. The Chinese version of this checklist has been validated among adolescents and has demonstrated satisfactory psychometric properties [69]. Cronbach α was .70 at T1, .75 at T2, and .81 at T3 in this study.

Loneliness

Loneliness was assessed using the three-item short form of the University of California, Los Angeles (UCLA) Loneliness Scale [70]. This brief measure captures core aspects of subjective social isolation, including feelings of lacking companionship, being left out, and experiencing isolation. Items were rated on a 4-point Likert scale ranging from 1 (never) to 4 (often), with higher scores reflecting greater loneliness. Previous research has demonstrated satisfactory psychometric properties of this instrument among Chinese adolescents [42]. Cronbach α was .86 at T1, .88 at T2, and .91 at T3 in this study.

Physical Activity

PA was assessed using the Physical Activity Rating Scale-3 (PARS-3), a three-item self-report instrument evaluating exercise intensity, duration, and frequency [71]. Each dimension is rated on a 5-point scale, and a composite score is derived according to the following formula: exercise intensity \times (exercise duration - 1) \times exercise frequency. This yields a possible range from 0 to 100. Higher total scores correspond to greater engagement in PA. Previous research has provided evidence of satisfactory psychometric properties of the scale among Chinese populations [26]. Cronbach α was .76 at T1, .79 at T2, and .80 at T3 in this study.

Covariates

In the RI-CLPM, sex, age, self-reported academic performance, perceived family financial situation, and single-parent family status were included as time-invariant covariates at the between-person level, whereas weekly gaming time was included as a time-varying covariate at the within-person level.

Data Analysis

Preliminary Analyses

Attrition analyses were performed to examine potential differences between participants who completed all waves and those lost to follow-up, using chi-square tests for categorical variables and independent-sample t tests for continuous variables. Pearson correlations were computed to examine bivariate associations among the main study variables. For

descriptive statistics and correlation analyses, listwise deletion was applied to handle missing data.

Measurement Invariance

Longitudinal measurement invariance was tested for PA, loneliness, and IGD symptoms. Configural invariance was first evaluated to establish a baseline model with acceptable fit, indicated by comparative fit index (CFI) ≥ 0.90 , and both root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) ≤ 0.08 [72]. Subsequent tests of metric and scalar invariance were based on changes in fit indices, with $\Delta\text{CFI} \leq 0.01$ and $\Delta\text{RMSEA} \leq 0.015$ indicating invariance across the three waves [72]. In addition, measurement invariance across sex was examined using the same sequential procedure (configural, metric, and scalar invariance) to ensure that the constructs were comparable between males and females.

Random Intercept Cross-Lagged Panel Modeling

An RI-CLPM was used to examine the within-person associations among PA, loneliness, and IGD symptoms. All variables were treated as continuous variables. Model construction began with a fully unconstrained specification, in which all autoregressive and cross-lagged pathways were estimated freely. Constraints were then introduced by equating autoregressive and cross-lagged paths across adjacent waves to test whether the longitudinal effects were stable over time. Once the optimal model fit was identified, covariates were included in the model. Sex, age, self-reported academic performance, perceived family financial situation, and single-parent family status were included as time-invariant covariates on the random intercepts (RIs), while weekly gaming time was included as a time-varying covariate at the within-person level. Model fit was evaluated using multiple indices, including $\chi^2(df) \leq 3$, both CFI

and Tucker-Lewis index (TLI) ≥ 0.90 , and both RMSEA and SRMR ≤ 0.08 [73].

Multigroup Analysis

To explore potential sexual differences in these associations, a multigroup RI-CLPM was conducted. A series of models, each constraining a specific path, were compared to an unconstrained model in which all paths were freely estimated. Significant differences between sexes were identified using Wald tests ($P < .05$).

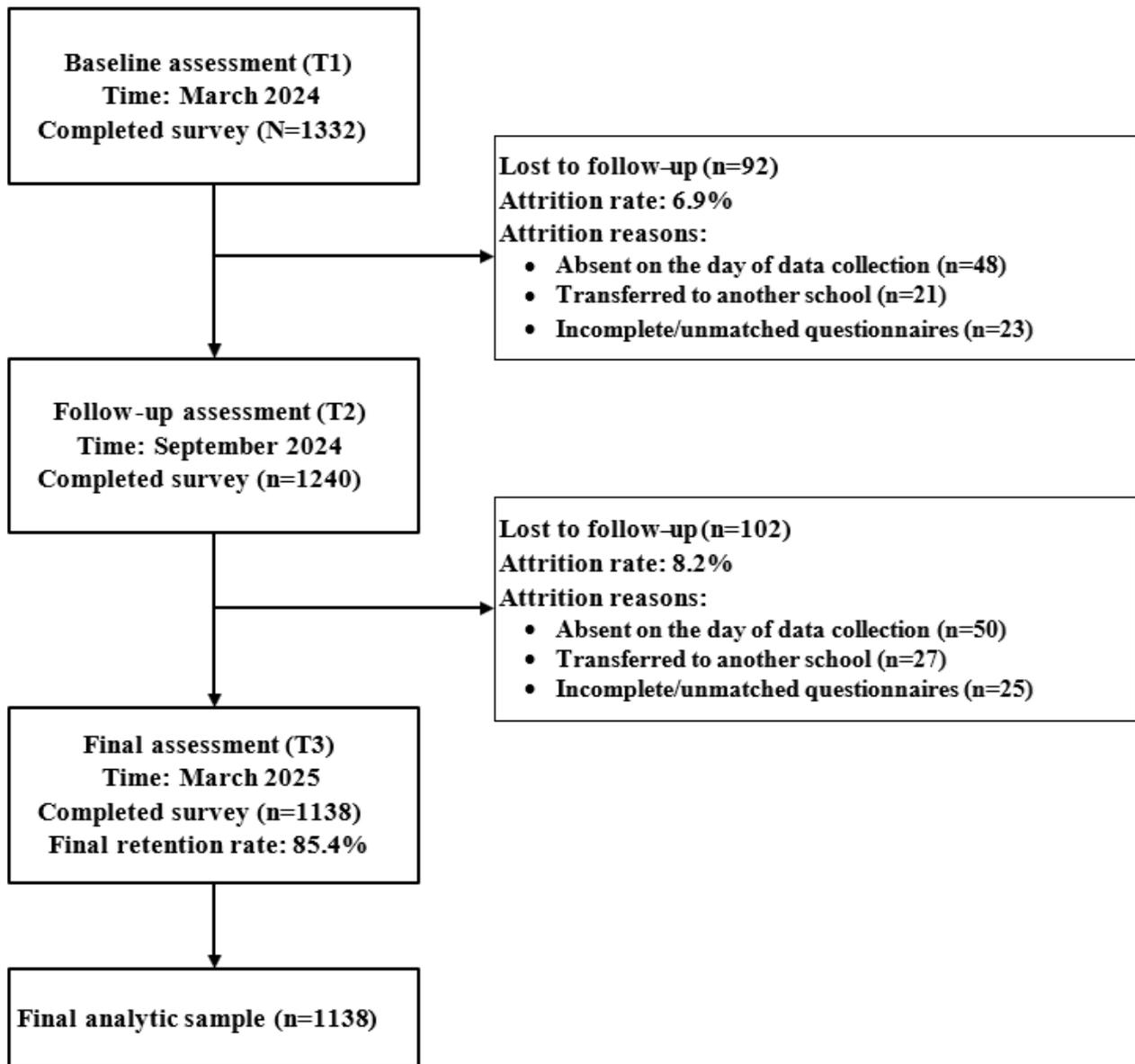
Software and Missing Data Handling

Preliminary analyses were conducted in IBM SPSS 26.0, and all SEM analyses were performed in Mplus 8.3 using robust maximum likelihood estimation (MLR). Prior to the main analyses, Little's Missing Completely at Random (MCAR) test was conducted to examine the missing data mechanism. The results indicated that the missing data were consistent with the assumption of MCAR ($\chi^2_{18} = 12.30, P = .83$). Therefore, missing data were handled using the full information maximum likelihood (FIML) method, which uses all available information to produce unbiased parameter estimates under MCAR assumptions. The FIML method provides unbiased parameter estimates by using all available information, thereby maximizing statistical power [74]. Statistical significance was defined as two-tailed $P < .05$.

Results

Attrition Analyses

A detailed overview of participant enrollment, retention, and attrition across each wave is provided in [Figure 1](#).

Figure 1. Participant flow diagram across the three waves of the study.

As presented in Table 1, the two groups did not differ significantly with respect to self-reported academic performance, parental education, perceived family financial status, single-parent household status, weekly gaming time, PA, or

loneliness. However, significant group differences were observed for sex ($P=.03$), age ($P=.001$), and IGD symptoms ($P=.02$).

Table 1. Attrition analyses comparing adolescent gamers who completed all three waves of the longitudinal study and those lost to follow-up (N=1332)^a.

Characteristics	Follow-up (n=1138)	Lost to follow-up (n=194)	P value
Sex, n (%)			.03
Male	745 (65.6)	142 (73.6)	— ^b
Female	390 (34.4)	51 (26.4)	—
Self-reported academic performance, n (%)			
Bottom 20%	168 (14.8)	18 (9.3)	.19
21st-40th percentile	241 (21.2)	41 (21.1)	—
41st-60th percentile (average)	288 (25.3)	53 (27.3)	—
61st-80th percentile	259 (22.8)	42 (21.6)	—
Top 20%	182 (16.0)	40 (20.6)	—
Father's educational level, n (%)			.05
Middle school or below	640 (56.2)	105 (54.1)	—
High school	284 (25.0)	66 (34.0)	—
College or above	214 (18.8)	23 (11.9)	—
Mother's educational level, n (%)			.84
Middle school or below	582 (51.1)	98 (50.5)	—
High school	351 (30.8)	60 (30.9)	—
College or above	205 (18.1)	36 (18.6)	—
Perceived family financial situation, n (%)			.65
Very poor/poor	149 (13.1)	30 (15.5)	—
Average	782 (68.7)	128 (66.0)	—
Good/very good	207 (18.2)	36 (18.5)	—
Single-parent family status, n (%)			.58
No	863 (75.8)	151 (77.8)	—
Yes	84 (7.4)	16 (8.2)	—
Not reported	191 (16.8)	27 (13.9)	—
Gaming time per week (hours) , n (%)			.55
<1	123 (10.8)	20 (10.3)	—
1-2	253 (22.2)	44 (22.7)	—
2-3	314 (27.6)	55 (28.4)	—
3-4	256 (22.5)	42 (21.6)	—
>4	192 (16.9)	33 (17.0)	—
Age (years), mean (SD)	13.61 (0.76)	13.81 (0.71)	.001
PA ^c (score 0-100), mean (SD)	75.27 (27.46)	74.35 (27.99)	.67
Loneliness (score 3-12), mean (SD)	4.61 (2.21)	4.88 (2.07)	.12
IGD ^d (score 0-9), mean (SD)	1.57 (1.76)	1.90 (1.86)	.02

^aMean (SD) for continuous variables and n (%) for categorical variables.

^bNot applicable.

^cPA: physical activity.

^dIGD: internet gaming disorder.

Pearson Correlations

Table 2 demonstrates the bivariate correlations. IGD symptoms were consistently and negatively associated with PA across all three waves, with correlation coefficients ranging from -0.26

to -0.19 . Similarly, loneliness exhibited significant negative correlations with PA at each wave, with r values between -0.44 and -0.25 . In contrast, IGD symptoms and loneliness were positively correlated over time, with coefficients ranging from 0.17 to 0.29 .

Table 2. Pearson correlation analyses between PA^a, loneliness, and IGD^b symptoms among adolescents.

Number	Variable	1	2	3	4	5	6	7	8	9
1	PA at T1 ^c	1	— ^d	—	—	—	—	—	—	—
2	PA at T2 ^e	0.42 ^f	1	—	—	—	—	—	—	—
3	PA at T3 ^g	0.37 ^f	0.43 ^f	1	—	—	—	—	—	—
4	Loneliness at T1	-0.39^f	-0.27^f	-0.26^f	1	—	—	—	—	—
5	Loneliness at T2	-0.29^f	-0.36^f	-0.27^f	0.58 ^f	1	—	—	—	—
6	Loneliness at T3	-0.25^f	-0.32^f	-0.44^f	0.48 ^f	0.59 ^f	1	—	—	—
7	IGD symptoms at T1	-0.23^f	-0.19^f	-0.21^f	0.25 ^f	0.17 ^f	0.19 ^f	1	—	—
8	IGD symptoms at T2	-0.23^f	-0.23^f	-0.22^f	0.24 ^f	0.27 ^f	0.29 ^f	0.51 ^f	1	—
9	IGD symptoms at T3	-0.25^f	-0.21^f	-0.26^f	0.20 ^f	0.25 ^f	0.29 ^f	0.44 ^f	0.55 ^f	1

^aPA: physical activity.

^bIGD: internet gaming disorder.

^cT1: wave 1.

^dNot applicable.

^eT3: wave 2.

^f $P < .001$.

^gT4: wave 3.

Measurement Invariance Test

Table 3 presents the results of longitudinal measurement invariance testing for PA, loneliness, and IGD symptoms. The configural invariance models showed acceptable fit, with the CFI and TLI exceeding 0.90 and the RMSEA and SRMR below 0.08, indicating that the basic factor structures of these constructs were stable across the three waves. Metric invariance was supported, as changes in fit indices ($\Delta CFI \leq 0.01$; $\Delta RMSEA \leq 0.015$) fell within recommended thresholds, suggesting that the factor loadings remained consistent over

time. Scalar invariance was also established, with minimal declines in model fit ($\Delta CFI \leq 0.01$; $\Delta RMSEA \leq 0.015$), indicating equivalence of item intercepts across measurement occasions. In addition, measurement invariance across sexes was examined using the same sequential procedure (configural, metric, and scalar invariance). The results similarly supported measurement invariance between males and females, indicating that the constructs were comparable across sexes. Detailed results of measurement invariance tests across sexes are presented in Table S1 in [Multimedia Appendix 1](#).

Table 3. Longitudinal invariance test of PA^a, loneliness, and IGD^b symptoms.

Variable and model	CFI ^c	TLI ^d	RMSEA ^e	SRMR ^f	Δ CFI	Δ RMSEA	Δ SRMR
PA							
Configural invariance	0.954	0.947	0.038	0.033	— ^g	—	—
Metric invariance	0.952	0.939	0.042	0.035	0.002	0.004	0.002
Scalar invariance	0.951	0.937	0.042	0.038	0.001	0.002	0.003
Loneliness							
Configural invariance	0.996	0.989	0.044	0.021	—	—	—
Metric invariance	0.996	0.992	0.038	0.023	0.000	0.006	0.002
Scalar invariance	0.995	0.992	0.038	0.023	0.001	0.000	0.000
IGD symptoms							
Configural invariance	0.941	0.929	0.036	0.034	—	—	—
Metric invariance	0.936	0.927	0.036	0.038	0.005	0	0.004
Scalar invariance	0.931	0.925	0.039	0.04	0.005	0.003	0.002

^aPA: physical activity.

^bIGD: internet gaming disorder.

^cCFI: comparative fit index.

^dTLI: Tucker-Lewis index.

^eRMSEA: root mean square error of approximation.

^fSRMR: standardized root mean square residual.

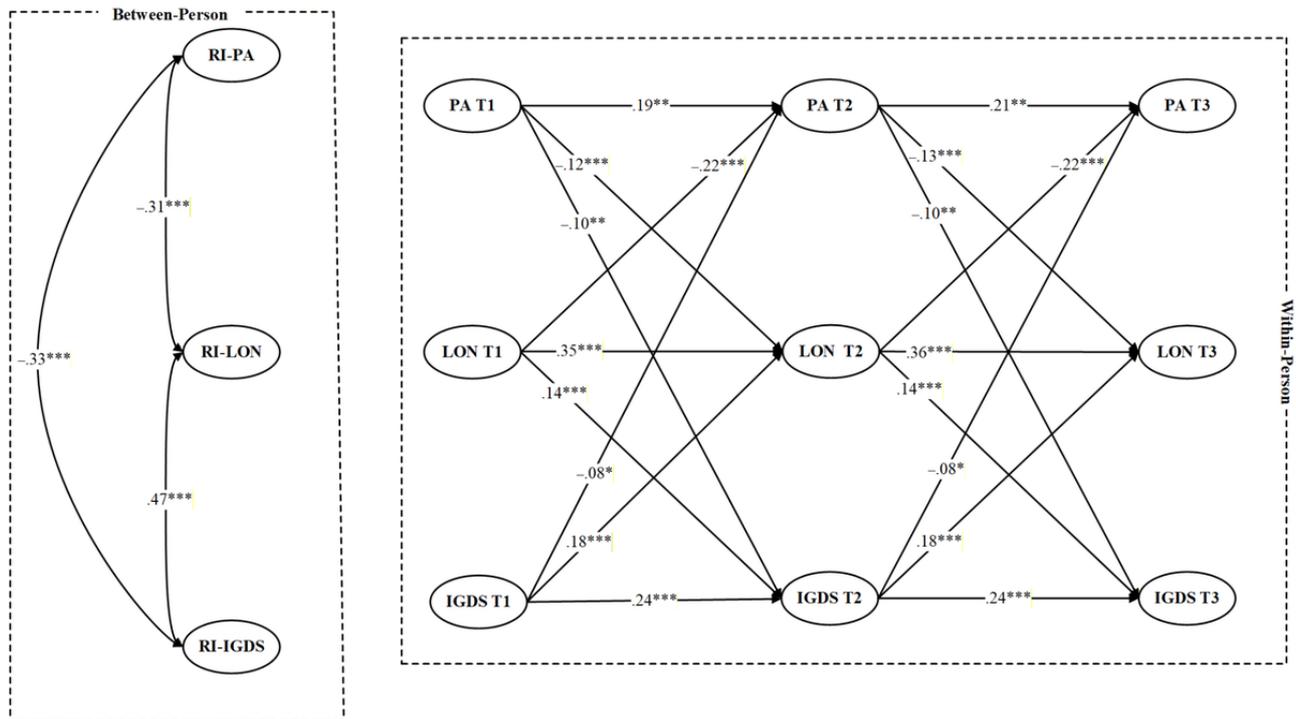
^gNot applicable.

RI-CLPM Results

A fully unconstrained model showed acceptable fit to the data ($\chi^2_8=35.12$, CFI=0.985, TLI=0.974, RMSEA=0.043, SRMR=0.030). Subsequently, a model constraining autoregressive and cross-lagged paths to be equal across adjacent waves was estimated. The chi-square difference test indicated that imposing these equality constraints does not significantly

worsen model fit ($\Delta\chi^2_9=15.35$, $P=.08$). Therefore, the more parsimonious constrained model was retained as the final model ($\chi^2_{17}=50.47$, CFI=0.989, TLI=0.977, RMSEA=0.042, SRMR=0.030). The results of model fit comparison are presented in Table S2 in [Multimedia Appendix 1](#). The results of the RI-CLPM examining longitudinal associations among PA, loneliness, and IGD symptoms are summarized in [Figure 2](#).

Figure 2. Results of the RI-CLPM (missing data handled via the FIML approach). Standardized estimates were reported. FIML: full information maximum likelihood; IGDS: internet gaming disorder symptoms; LON: loneliness; PA: physical activity; RI-CLPM: random intercept cross-lagged panel model; RI-IGDS: random intercept of internet gaming disorder symptoms; RI-LON: random intercept of loneliness; RI-PA: random intercept of physical activity. * $P < .05$, ** $P < .01$, *** $P < .001$.



Autoregressive paths for each variable were significant across waves, indicating substantial temporal stability for PA, loneliness, and IGD symptoms. Cross-lagged effects revealed significant reciprocal associations among PA, loneliness, and IGD symptoms across the three measurement waves. Specifically, higher PA predicted decreased subsequent loneliness (T1→T2: $\beta = -0.12$, 95% CI -0.16 to -0.08 , $P < .001$; T2→T3: $\beta = -0.13$, 95% CI -0.17 to -0.08 , $P < .001$), whereas increased loneliness predicted lower PA at the following waves (T1→T2: $\beta = -0.22$, 95% CI -0.26 to -0.18 , $P < .001$; T2→T3: $\beta = -0.22$, 95% CI -0.26 to -0.18 , $P < .001$). Similarly, higher PA was associated with decreased subsequent IGD symptoms (T1→T2: $\beta = -0.10$, 95% CI -0.16 to -0.04 , $P = .009$; T2→T3: $\beta = -0.10$, 95% CI -0.16 to -0.04 , $P = .009$), while increased IGD symptoms predicted decreases in PA over time (T1→T2: $\beta = -0.08$, 95% CI -0.12 to -0.04 , $P = .02$; T2→T3: $\beta = -0.08$, 95% CI -0.12 to -0.04 , $P = .02$). In addition, increased loneliness predicted increases in subsequent IGD symptoms (T1→T2: $\beta = 0.14$, 95% CI 0.11 - 0.17 , $P < .001$; T2→T3: $\beta = 0.14$, 95% CI 0.11 - 0.17 , $P < .001$), and increased IGD symptoms, in turn, predicted increased loneliness at the next wave (T1→T2: $\beta = 0.18$, 95% CI 0.14 - 0.22 , $P < .001$; T2→T3: $\beta = 0.18$, 95% CI 0.14 - 0.22 , $P < .001$).

Within-person residual correlations revealed that PA was negatively associated with both loneliness and IGD symptoms at all three waves, with standardized coefficients ranging from -0.33 to -0.28 (all $P < .001$). In contrast, loneliness and IGD symptoms were positively correlated at the within-person level across waves, with coefficients ranging from 0.21 to 0.38 (all $P < .001$).

At the between-person level, latent trait correlations further supported these associations. The latent factor of PA was negatively associated with the latent traits of loneliness ($r = -0.31$, 95% CI -0.45 to -0.17 , $P < .001$) and IGD symptoms ($r = -0.33$, 95% CI -0.49 to -0.17 , $P < .001$), indicating that adolescents with higher overall levels of PA across the study period tended to report consistently decreased loneliness and IGD symptoms. Moreover, loneliness and IGD symptoms were strongly and positively associated at the between-person level ($r = 0.47$, 95% CI 0.35 - 0.59 , $P < .001$), suggesting a stable covariation between these two constructs over time.

The effects of time-invariant covariates on the RIs are reported in Table S3 in [Multimedia Appendix 1](#). Sex was significantly associated with the RI of PA ($\beta = -0.08$, 95% CI -0.13 to -0.02 , $P = .005$) and IGD symptoms ($\beta = -0.08$, 95% CI -0.15 to -0.01 , $P = .001$), indicating that females (0=male, 1=female) reported slightly lower overall levels of PA and IGD symptoms across the study period. In contrast, age, self-reported academic performance, perceived family financial situation, and single-parent family status were not significantly associated with the RIs (all $P > .05$).

The time-varying effects of gaming time on PA, loneliness, and IGD symptoms are presented in Table S4 in [Multimedia Appendix 1](#). Gaming time was negatively associated with PA at each wave (T1: $\beta = -0.11$, 95% CI -0.19 to -0.03 , $P < .001$; T2: $\beta = -0.09$, 95% CI -0.15 to -0.03 , $P = .001$; T3: $\beta = -0.06$, 95% CI -0.10 to -0.02 , $P = .04$), suggesting that higher gaming engagement is concurrently associated with lower PA. In addition, gaming time was positively associated with loneliness at each wave (T1: $\beta = 0.07$, 95% CI 0.01 - 0.13 , $P = .009$; T2: $\beta = 0.09$, 95% CI 0.03 - 0.15 , $P = .001$; T3: $\beta = 0.06$, 95% CI

0.02-0.10, $P=.02$), suggesting that higher gaming engagement is concurrently associated with increased loneliness. Moreover, gaming time was positively associated with IGD symptoms at each wave (T1: $\beta=0.10$, 95% CI 0.04-0.16, $P<.001$; T2: $\beta=0.06$, 95% CI 0.01-0.12, $P=.01$; T3: $\beta=0.08$, 95% CI 0.02-0.14, $P=.004$), suggesting that higher gaming engagement is concurrently associated with increased IGD symptom severity.

Sexual Differences

Multigroup RI-CLPM was conducted to examine whether the longitudinal associations among PA, loneliness, and IGD symptoms varied by sex. In the first step, an unconstrained

model allowing all cross-lagged and autoregressive paths to vary freely across males and females was estimated and demonstrated acceptable model fit: $\chi^2_8=21.46$, CFI=0.971, TLI=0.954, RMSEA=0.038, SRMR=0.036. In the second step, a constrained model was tested in which all cross-lagged paths were set equal across groups. This model also exhibited good fit: $\chi^2_{18}=60.24$, CFI=0.942, TLI=0.915, RMSEA=0.077, SRMR=0.021. However, the chi-square difference test revealed a significant decline in model fit when the cross-lagged paths were constrained ($\Delta\chi^2_{10}=38.78$, $P<.001$), suggesting that the model results significantly differed by sex. Figures 3 and 4 present the results.

Figure 3. Results of the RI-CLPM for male adolescent gamers (missing data handled via the FIML approach). Standardized estimates were reported. FIML: full information maximum likelihood; IGDS: internet gaming disorder symptoms; LON: loneliness; PA: physical activity; RI-CLPM: random intercept cross-lagged panel model; RI-IGDS: random intercept of internet gaming disorder symptoms; RI-LON: random intercept of loneliness; RI-PA: random intercept of physical activity. * $P<.05$, ** $P<.01$, *** $P<.001$.

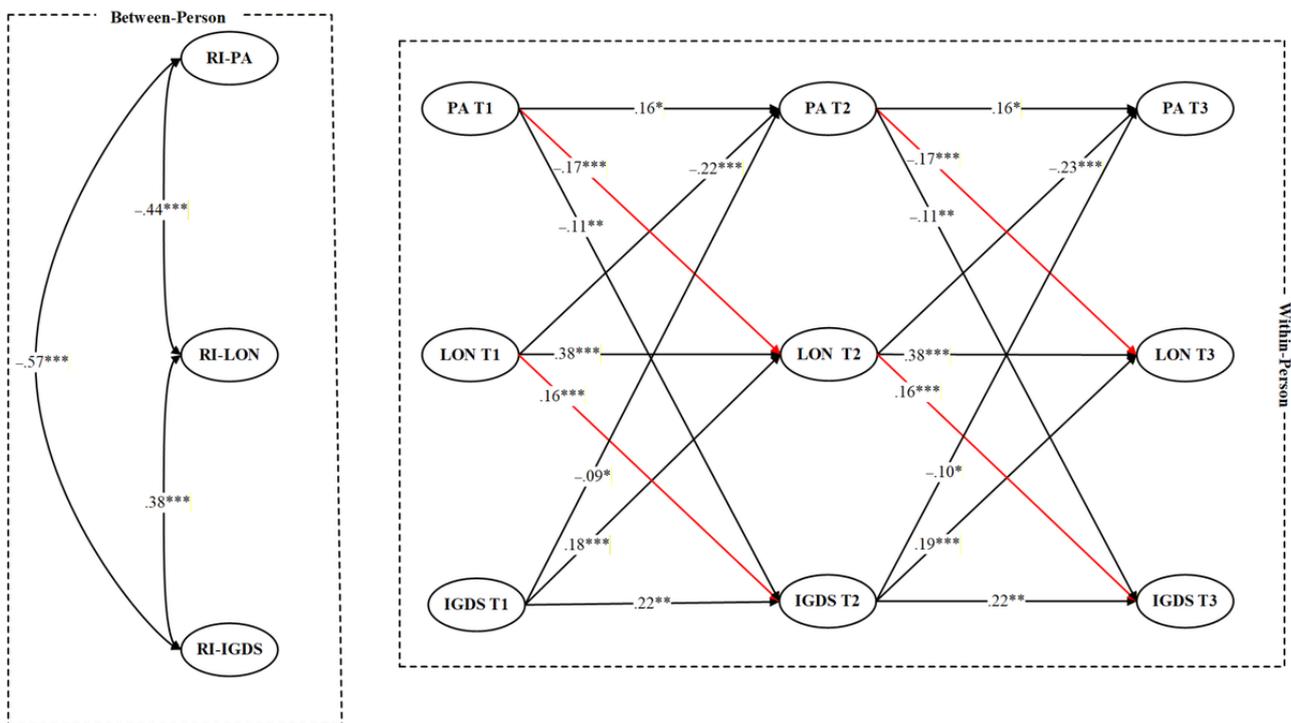
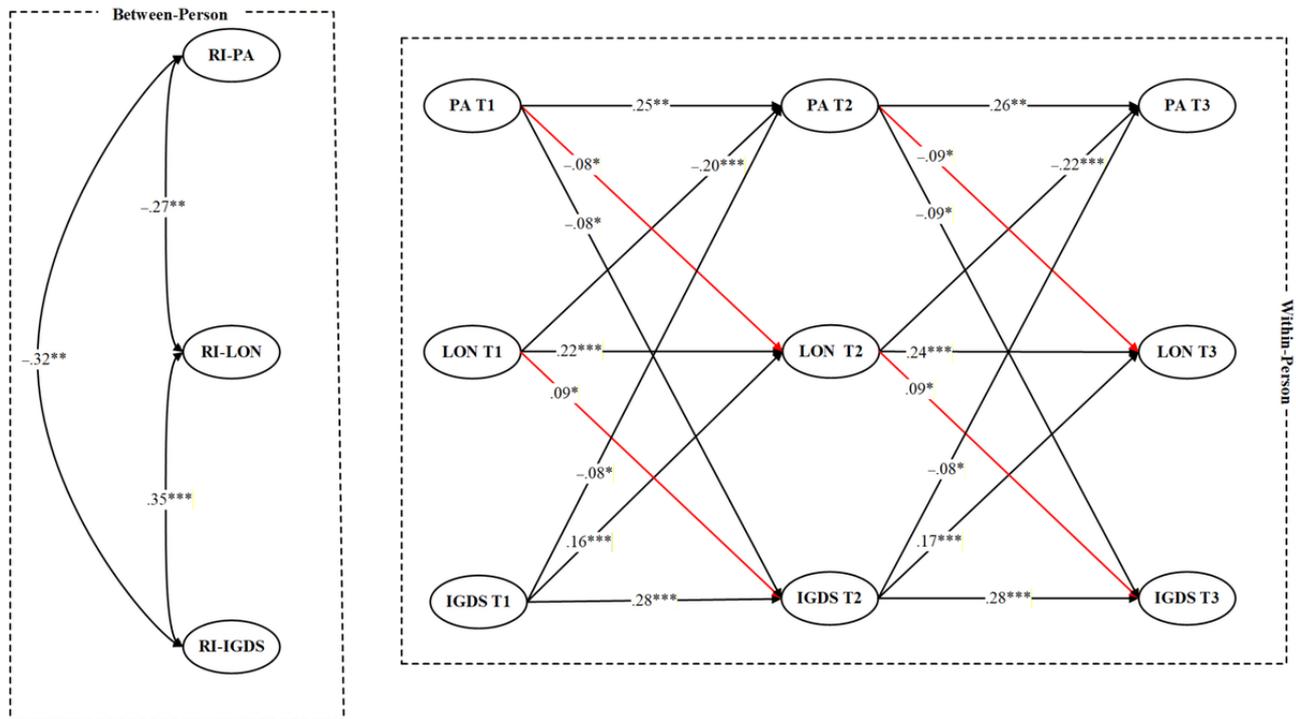


Figure 4. Results of the RI-CLPM for female adolescent gamers (missing data handled via the FIML approach). Standardized estimates were reported. FIML: full information maximum likelihood; IGDS: internet gaming disorder symptoms; LON: loneliness; PA: physical activity; RI-CLPM: random intercept cross-lagged panel model; RI-IGDS: random intercept of internet gaming disorder symptoms; RI-LON: random intercept of loneliness; RI-PA: random intercept of physical activity. * $P < .05$, ** $P < .01$, *** $P < .001$.



Specifically, sexual differences emerged in the path from PA to loneliness. A Wald test of parameter constraints indicated a significant difference between females and males (Wald test estimate=6.27, $P = .01$). For males, PA significantly predicted loneliness (T1→T2: $\beta = -0.08$, 95% CI -0.15 to -0.01 , $P = .03$; T2→T3: $\beta = -0.09$, 95% CI -0.16 to -0.02 , $P = .01$). For females, the cross-lagged effects from PA to loneliness were significant but stronger (T1→T2: $\beta = -0.17$, 95% CI -0.23 to -0.11 , $P < .001$; T2→T3: $\beta = -0.17$, 95% CI -0.23 to -0.11 , $P < .001$). Sexual differences were also found in the pathway from loneliness predicting IGD symptoms (Wald test estimate=5.87, $P = .02$). Among males, loneliness significantly predicted IGD symptoms (T1→T2: $\beta = 0.09$, 95% CI 0.03 - 0.15 , $P = .02$; T2→T3: $\beta = 0.09$, 95% CI 0.03 - 0.15 , $P = .02$). Among females, the paths from loneliness to later IGD symptoms were also significant (T1→T2: $\beta = 0.16$, 95% CI 0.11 - 0.21 , $P < .001$; T2→T3: $\beta = 0.16$, 95% CI 0.11 - 0.21 , $P < .001$).

For the remaining cross-lagged paths, no significant sexual differences were observed. Specifically, PA negatively predicted subsequent IGD symptoms for both males (T1→T2: $\beta = -0.08$, 95% CI -0.15 to -0.01 , $P = .02$; T2→T3: $\beta = -0.09$, 95% CI -0.16 to -0.02 , $P = .01$) and females (T1→T2: $\beta = -0.11$, 95% CI -0.17 to -0.05 , $P = .004$; T2→T3: $\beta = -0.11$, 95% CI -0.17 to -0.05 , $P = .004$). IGD symptoms also predicted decreases in PA across time among males (T1→T2: $\beta = -0.08$, 95% CI -0.14 to -0.02 , $P = .03$; T2→T3: $\beta = -0.09$, 95% CI -0.15 to -0.03 , $P = .02$) and females (T1→T2: $\beta = -0.09$, 95% CI -0.15 to -0.03 , $P = .02$; T2→T3: $\beta = -0.08$, 95% CI -0.14 to -0.02 , $P = .03$). In addition, IGD symptoms predicted subsequent increases in loneliness among both males (T1→T2: $\beta = 0.16$, 95% CI 0.11 - 0.21 , $P < .001$; T2→T3: $\beta = 0.17$, 95% CI 0.12 - 0.22 , $P < .001$) and females (T1→T2: $\beta = 0.18$, 95% CI 0.13 - 0.23 , $P < .001$; T2→T3: $\beta = 0.19$,

95% CI 0.14 - 0.24 , $P < .001$). Likewise, loneliness predicted later decreases in PA for males (T1→T2: $\beta = -0.20$, 95% CI -0.25 to -0.15 , $P < .001$; T2→T3: $\beta = -0.21$, 95% CI -0.26 to -0.16 , $P < .001$) and females (T1→T2: $\beta = -0.22$, 95% CI -0.27 to -0.17 , $P < .001$; T2→T3: $\beta = -0.23$, 95% CI -0.28 to -0.18 , $P < .001$).

Sensitivity Analyses

To ensure that the main findings were not influenced by missing data or analytic choices, additional analyses were carried out. A complete case analysis (listwise deletion approach) was applied, including only participants who participated in all three assessment waves ($n = 1138$, 85.4%). Although this reduced the sample size, the pattern and size of the cross-lagged paths among PA, loneliness, IGD symptoms were almost identical to those from the FIML model. The similar pattern also appeared in the sex-specific models, indicating that the main conclusions were not sensitive to data exclusion. Figures S1-S3 in [Multimedia Appendix 2](#) present these results. This procedure follows current recommendations for longitudinal research and aligns with previous applications of RI-CLPMs in developmental studies [75,76].

Discussion

Principal Findings

This study used a three-wave RI-CLPM to investigate the within-person associations between PA, loneliness, and IGD symptoms among Chinese adolescent gamers. The results revealed bidirectional and longitudinal relationships between PA, loneliness, and IGD symptoms. Multigroup RI-CLPM analyses indicated that the effect of PA on loneliness is stronger among females and that the association between loneliness and

subsequent IGD symptoms is also stronger for females. These findings highlight the dynamic nature of these associations, emphasizing the importance of considering temporal and sex-specific patterns in understanding fluctuations in these behaviors over time.

Reciprocal Association Between PA and IGD Symptoms

This study revealed clear bidirectional within-person associations between PA and IGD symptoms across time, supporting H1. Consistent with previous studies, this pattern indicates a mutually reinforcing process between PA and IGD symptoms [28,77]. Importantly, the cross-lagged effect size, although statistically significant, was modest (0.08-0.10), aligning with recent methodological benchmarks proposed by Orth et al [78], who noted that cross-lagged effects in longitudinal models are typically small, with $\beta=0.03$, 0.07, and 0.12 representing small, medium, and large effects, respectively. From this perspective, the observed effects, although modest, indicate meaningful within-person short-term dynamics across waves without implying long-term accumulation, which is beyond the RI-CLPM framework [79].

The finding that PA prospectively predicts decreased IGD symptoms suggests that engaging in PA may reduce adolescents' risk for problematic gaming by providing structured, rewarding offline activities that compete with gaming involvement [80]. Participation in team or group-based sports also satisfies needs for relatedness and competence, offering offline sources of reward and belonging that may otherwise be sought in gaming environments [81,82]. In addition, PA may support emotion regulation and self-control, reducing reliance on gaming as a coping strategy [83,84]. The reverse pathway was also observed, with increased IGD symptoms predicting lower subsequent PA, suggesting that problematic gaming may constrain adolescents' active engagement [85]. As gaming time expands, it may displace opportunities for physical movement, disturb daily routines, and foster social withdrawal from offline peer contexts, collectively diminishing the motivation and energy required for regular exercise [86]. These within-person reciprocal associations suggest a short-term reinforcing cycle between PA and IGD symptoms.

By applying an RI-CLPM, the study clarified that these associations reflect within-person temporal dynamics rather than stable between-person differences. These findings highlight the potential value of promoting regular PA as part of strategies aimed at reducing problematic gaming among adolescents [87]. Importantly, sustaining increases in PA over longer periods can be challenging among adolescents. Therefore, interventions may benefit from combining PA promotion with additional strategies, such as structured reductions in gaming time, parental monitoring, and school- or community-based programs that encourage balanced leisure activities [88,89].

Reciprocal Association Between Loneliness and IGD Symptoms

This study revealed consistent bidirectional within-person associations between loneliness and IGD symptoms over time, supporting H2. The finding that loneliness predicts increased IGD symptoms is consistent with compensatory internet use

theory [34]. Online gaming may provide immediate gratification, social interaction, and emotional relief, temporarily alleviating feelings of loneliness [90]. However, as adolescents continue to seek solace in gaming, it can lead to more persistent problematic gaming behaviors. This process may be further compounded by the lack of emotional regulation that often accompanies loneliness, which leaves adolescents more vulnerable to maladaptive coping strategies, such as excessive gaming [91]. Equally important is the reverse path: increased IGD symptoms predicting increased loneliness over time. Excessive gaming may reduce offline social interaction, contributing to increased social isolation and loneliness [92,93]. This pattern suggests a reinforcing cycle in which loneliness promotes gaming, while gaming further intensifies loneliness. As problematic gaming often substitutes for meaningful social connections, it further diminishes adolescents' opportunities to foster real-world relationships [94].

Furthermore, the use of RI-CLPMs clarifies that these associations reflect within-person temporal dynamics rather than stable between-person differences. This approach strengthens the conclusion that interventions targeting loneliness and IGD symptoms need to address their dynamic and reciprocal nature. Consistent with previous research, these findings suggest that interventions should address both emotional distress and gaming behaviors [45]. This highlights the importance of integrating interventions that focus on emotion regulation and social connection, such as promoting offline social activities, while simultaneously addressing problematic gaming behaviors [95]. Strengthening offline social engagement and coping skills may help mitigate the negative impacts of both loneliness and IGD symptoms [96]. As a practical implication, fostering social relationships and coping skills in offline settings, such as through group activities or community support programs, could serve as an effective way to mitigate the negative impacts of both loneliness and IGD symptoms on adolescents [97].

Reciprocal Association Between PA and Loneliness

This study revealed clear bidirectional within-person associations between PA and loneliness across time, supporting H3. The reciprocal relationship highlights the importance of understanding how emotional well-being and PA influence each other over time, particularly in adolescence [98]. The finding that PA predicts decreased loneliness is consistent with prior research emphasizing the role of structured, socially engaging activities in alleviating loneliness [58]. PA offers adolescents a sense of competence and achievement, which enhances self-esteem and buffers against emotional distress [83]. This is particularly important during adolescence, a period in which social validation and peer acceptance are key to emotional well-being [99]. Conversely, the finding that loneliness predicts reduced PA suggests that emotional distress plays a significant role in adolescents' engagement in healthy behaviors. Loneliness can undermine adolescents' motivation to participate in PA, as the emotional pain of isolation may lead them to withdraw from social settings, including sports or group activities [98]. In line with BPNT, when the need for relatedness is not met, adolescents are less likely to engage in activities that require social interaction, such as PA [100].

From a practical perspective, the findings suggest that promoting regular PA could be an effective strategy for reducing loneliness in adolescents. Interventions that incorporate group-based sports or other socially engaging physical activities could help adolescents build stronger social connections, thereby reducing feelings of isolation [101]. At the same time, addressing loneliness through social support or community-building activities may also encourage adolescents to re-engage in PA, ultimately improving both their emotional well-being and physical health [102]. This highlights the importance of integrating interventions that target both emotional distress (loneliness) and behavioral patterns, such as excessive gaming, while simultaneously promoting PA, to break the self-reinforcing cycle between these factors.

Concurrent Associations of Gaming Time With PA, Loneliness, and IGD Symptoms

Gaming time was concurrently linked with adolescents' PA, loneliness, and IGD symptoms. These patterns suggest that gaming engagement may function as a shared behavioral context in which reduced PA, increased loneliness, and problematic gaming co-occur. This finding aligns with research showing that intensive gaming can restructure adolescents' daily routines, displacing PA while simultaneously shifting social interaction from offline contexts to online environments [103-105]. Thus, gaming time may represent a proximal behavioral condition that connects lifestyle behaviors and psychosocial experiences, complementing the longer-term within-person dynamics identified by the RI-CLPM analyses.

Sexual Differences

Multigroup RI-CLPM analyses revealed significant sexual differences in the paths from PA to loneliness and from loneliness to IGD symptoms, supporting H4. Specifically, the effects of these two pathways were stronger among females. Several potential explanations may help account for these sexual differences. First, sex-related socialization and developmental psychology suggest that adolescent females are more likely to prioritize interpersonal relationships and social connectedness than males [106,107]. As such, feelings of loneliness may elicit more intense emotional distress in females, leading them to seek social compensation through online interactions, including gaming [90]. This is consistent with research indicating that females often experience higher interpersonal sensitivity and engage in more internalizing responses to social stress, which could make them more likely to use gaming or other online avenues to cope with loneliness [5]. Second, the motivations for gaming may differ between sexes. Although males tend to engage in gaming for achievement and competition, females are more likely to seek social affiliation through gaming [108]. Therefore, if females' gaming is more socially motivated, loneliness may act as a stronger precursor to problematic gaming behaviors in females. Our findings support this pattern, indicating that loneliness functions as a stronger prospective risk factor for IGD symptoms among adolescent females. Similarly, Lacko et al [109] reported sex-specific within-person effects using an RI-CLPM, showing that increases in social gaming were associated with decreased loneliness among boys but increased loneliness among girls. Although their study

focused on social gaming rather than IGD symptoms, their findings suggest that gaming-related behaviors may have different psychosocial implications for male and female adolescents.

Practically, these findings underscore the importance of sex-sensitive interventions [110]. For females, interventions should prioritize enhancing social connectedness, providing peer support, and promoting emotion-focused coping strategies alongside PA to address the stronger impact of PA on loneliness [95]. For males, programs may benefit from emphasizing PA as an alternative reward system and offering structured, competitive activities to reduce the appeal of gaming and mitigate its negative impact [95].

Limitations

This study has several limitations. First, variables, especially PA, were assessed using self-report questionnaires, which may be subject to response biases, such as social desirability or recall bias. Future studies could incorporate multi-informant reports or objective measures (eg, accelerometers) to improve accuracy. Second, although the three-wave design enabled examination of temporal dynamics, the short intervals and limited number of waves restrict the ability to capture longer-term developmental processes. More frequent assessments across longer periods are needed in future studies. Third, although cross-lagged effects were statistically significant and fell within the medium-to-large range [78], their magnitudes should still be interpreted with caution in the light of the large sample size and the reliance on self-reported measures. Fourth, although missing data analyses suggested that the missing-at-random assumption was reasonable and FIML estimation was applied, differential attrition related to sex, age, and IGD symptoms may still introduce potential bias. Fifth, the sample consisted of Chinese adolescent gamers from urban schools, which may limit generalizability to other cultural or contextual settings, particularly rural populations. Future research should consider including more diverse samples and examining potential urban-rural differences. Finally, the study did not collect data on the types of games played (eg, competitive, social, or cooperative), which could have provided additional insights into how different gaming experiences influence the observed outcomes.

Conclusion

This study used a three-wave longitudinal design to examine the within-person dynamics between PA, loneliness, and IGD symptoms among Chinese adolescent gamers. The results revealed bidirectional relationships between these variables, reflecting a mutually reinforcing cycle. Importantly, sexual differences were observed in the strength of these associations, with the effects of PA on loneliness and loneliness on IGD symptoms being stronger for females. By using an RI-CLPM to distinguish within-person changes from stable between-person differences, this study extends prior research based mainly on cross-sectional designs or traditional CLPMs and provides a clearer understanding of how behavioral and emotional factors jointly shape adolescent problematic gaming. The results suggest that interventions should be sex sensitive, addressing both PA

and loneliness simultaneously to effectively mitigate the risk of IGD symptoms, particularly among females.

Acknowledgments

We thank all the participants who took part in the study.

Data Availability

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

Funding

No financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Authors' Contributions

CH was responsible for writing—original draft, validation, methodology, investigation, formal analysis, and conceptualization; XM and XW for writing—original draft, project administration, methodology, investigation, data curation, and conceptualization; and LJH for writing—review and editing, supervision, resources, project administration, methodology, investigation, data curation, and conceptualization.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Measurement invariance across sex; model fit comparison between unconstrained model and constrained model; effects of time-invariant covariates on the random intercepts in the RI-CLPM; and concurrent effects of gaming time on PA, loneliness, and IGD symptoms across the three waves in the RI-CLPM. IGD: internet gaming disorder; PA: physical activity; RI-CLPM: random intercept cross-lagged panel model.

[[DOCX File, 24 KB - games_v14i1e87847_app1.docx](#)]

Multimedia Appendix 2

Results of the RI-CLPM for all participants, male adolescent gamers, and female adolescent gamers. RI-CLPM: random intercept cross-lagged panel model.

[[DOCX File, 207 KB - games_v14i1e87847_app2.docx](#)]

References

1. Yu Y, Yen SHY, Wang DB, Wu AMS, Chen JH, Zhang G, et al. School climate and school identification as determinants of internet gaming disorder among Chinese adolescent internet gamers: cross-sectional mediation study. *JMIR Serious Games* 2024 Nov 12;12:e50418 [[FREE Full text](#)] [doi: [10.2196/50418](https://doi.org/10.2196/50418)] [Medline: [39622695](https://pubmed.ncbi.nlm.nih.gov/39622695/)]
2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*. Washington, DC: American Psychiatric Association; 2013.
3. Dang L, Zhai J, Kam CCS, Su X, Wu AMS. Gaming for self-searching? A three-wave longitudinal investigation on the correlations and predictions among self-concept clarity, player-avatar identification, and internet gaming disorder. *J Youth Adolesc* 2026 Jan 17;55(1):14-26 [[FREE Full text](#)] [doi: [10.1007/s10964-025-02258-w](https://doi.org/10.1007/s10964-025-02258-w)] [Medline: [40960706](https://pubmed.ncbi.nlm.nih.gov/40960706/)]
4. Poon LYJ, Tsang HWH, Chan TYJ, Man SWT, Ng LY, Wong YLE, et al. Psychometric properties of the internet gaming disorder scale-short-form (IGDS9-SF): systematic review. *J Med Internet Res* 2021 Oct 18;23(10):e26821 [[FREE Full text](#)] [doi: [10.2196/26821](https://doi.org/10.2196/26821)] [Medline: [34661543](https://pubmed.ncbi.nlm.nih.gov/34661543/)]
5. Yu Y, Mo PK, Zhang J, Li J, Lau JT. Impulsivity, self-control, interpersonal influences, and maladaptive cognitions as factors of internet gaming disorder among adolescents in China: cross-sectional mediation study. *J Med Internet Res* 2021 Oct 27;23(10):e26810 [[FREE Full text](#)] [doi: [10.2196/26810](https://doi.org/10.2196/26810)] [Medline: [34704960](https://pubmed.ncbi.nlm.nih.gov/34704960/)]
6. Yu Y, Li S, Lyu X, Zhang J, Li J, Lau JTF. A cross-lagged panel model of the bidirectional relationships between gaming-specific maladaptive cognitions and internet gaming disorder among adolescents in China. *Addict Behav* 2026 Jan;172:108518 [[FREE Full text](#)] [doi: [10.1016/j.addbeh.2025.108518](https://doi.org/10.1016/j.addbeh.2025.108518)] [Medline: [41075582](https://pubmed.ncbi.nlm.nih.gov/41075582/)]
7. Zheng P, Tao Z, Fang L, Yao T, Zhou L, Wang Y, et al. Longitudinal association between internet gaming disorder and school refusal among adolescents using a random intercept cross-lagged panel model: three-wave prospective cohort study. *JMIR Serious Games* 2026 Mar 04;14:e89619 [[FREE Full text](#)] [doi: [10.2196/89619](https://doi.org/10.2196/89619)] [Medline: [41813117](https://pubmed.ncbi.nlm.nih.gov/41813117/)]

8. Anderson CA, Sakamoto A, Gentile DA, Ihori N, Shibuya A, Yukawa S, et al. Longitudinal effects of violent video games on aggression in Japan and the United States. *Pediatrics* 2008 Nov;122(5):e1067-e1072 [FREE Full text] [doi: [10.1542/peds.2008-1425](https://doi.org/10.1542/peds.2008-1425)] [Medline: [18977956](https://pubmed.ncbi.nlm.nih.gov/18977956/)]
9. Rahul R, Kareem SA, Gowda KC, Sajna SJ. Impact of internet gaming disorder on mental well-being: prevalence and psychological distress among adolescents and young adults. *Psychiatry Res* 2025 Jul;349:116512 [FREE Full text] [doi: [10.1016/j.psychres.2025.116512](https://doi.org/10.1016/j.psychres.2025.116512)] [Medline: [40318564](https://pubmed.ncbi.nlm.nih.gov/40318564/)]
10. Teng Z, Griffiths MD, Nie Q, Xiang G, Guo C. Parent-adolescent attachment and peer attachment associated with internet gaming disorder: a longitudinal study of first-year undergraduate students. *J Behav Addict* 2020 Apr 01;9(1):116-128 [FREE Full text] [doi: [10.1556/2006.2020.00011](https://doi.org/10.1556/2006.2020.00011)] [Medline: [32359235](https://pubmed.ncbi.nlm.nih.gov/32359235/)]
11. Tsui YY, Cheng C. Internet gaming disorder, risky online behaviour, and mental health in Hong Kong adolescents: the beneficial role of psychological resilience. *Front Psychiatry* 2021;12:722353 [FREE Full text] [doi: [10.3389/fpsy.2021.722353](https://doi.org/10.3389/fpsy.2021.722353)] [Medline: [34721101](https://pubmed.ncbi.nlm.nih.gov/34721101/)]
12. da Silva Pinho A, Céspedes Izquierdo V, Lindström B, van den Bos W. Youths' sensitivity to social media feedback: a computational account. *Sci Adv* 2024 Oct 25;10(43):eadp8775 [FREE Full text] [doi: [10.1126/sciadv.adp8775](https://doi.org/10.1126/sciadv.adp8775)] [Medline: [39441931](https://pubmed.ncbi.nlm.nih.gov/39441931/)]
13. Maza MT, Fox KA, Kwon S, Flannery JE, Lindquist KA, Prinstein MJ, et al. Association of habitual checking behaviors on social media with longitudinal functional brain development. *JAMA Pediatr* 2023 Feb 01;177(2):160-167 [FREE Full text] [doi: [10.1001/jamapediatrics.2022.4924](https://doi.org/10.1001/jamapediatrics.2022.4924)] [Medline: [36595277](https://pubmed.ncbi.nlm.nih.gov/36595277/)]
14. Lizandra J, Devis-Devis J, Valencia-Peris A, Tomás JM, Peiró-Velert C. Screen time and moderate-to-vigorous physical activity changes and displacement in adolescence: a prospective cohort study. *Eur J Sport Sci* 2019 Jun;19(5):686-695 [FREE Full text] [doi: [10.1080/17461391.2018.1548649](https://doi.org/10.1080/17461391.2018.1548649)] [Medline: [30550370](https://pubmed.ncbi.nlm.nih.gov/30550370/)]
15. Quicktake: MOE blocks phones in schools. *China Law Translate*. 2021. URL: <https://www.chinalawtranslate.com/quicktake-moe-blocks-phones-in-schools/> [accessed 2026-03-14]
16. Marques A, Calmeiro L, Loureiro N, Frasquilho D, de Matos MG. Health complaints among adolescents: associations with more screen-based behaviours and less physical activity. *J Adolesc* 2015 Oct;44:150-157 [FREE Full text] [doi: [10.1016/j.adolescence.2015.07.018](https://doi.org/10.1016/j.adolescence.2015.07.018)] [Medline: [26275746](https://pubmed.ncbi.nlm.nih.gov/26275746/)]
17. Spengler S, Mess F, Woll A. Do media use and physical activity compete in adolescents? Results of the MoMo study. *PLoS One* 2015;10(12):e0142544 [FREE Full text] [doi: [10.1371/journal.pone.0142544](https://doi.org/10.1371/journal.pone.0142544)] [Medline: [26629688](https://pubmed.ncbi.nlm.nih.gov/26629688/)]
18. Aryani A, Kusuma K, Saputra M. Physical exercise intervention for gaming disorder: a systematic review of clinical trials. *Int J Ment Health* 2024 May 31;53(4):438-452 [FREE Full text] [doi: [10.1080/00207411.2024.2347672](https://doi.org/10.1080/00207411.2024.2347672)]
19. Fu Q, Li L, Li Q, Wang J. The effects of physical activity on the mental health of typically developing children and adolescents: a systematic review and meta-analysis. *BMC Public Health* 2025 Apr 23;25(1):1514 [FREE Full text] [doi: [10.1186/s12889-025-22690-8](https://doi.org/10.1186/s12889-025-22690-8)] [Medline: [40269876](https://pubmed.ncbi.nlm.nih.gov/40269876/)]
20. Hu Y, Liu B. The impact of physical exercise on college students' mental health through emotion regulation and self-efficacy. *Sci Rep* 2025 Sep 29;15(1):33548 [FREE Full text] [doi: [10.1038/s41598-025-18352-9](https://doi.org/10.1038/s41598-025-18352-9)] [Medline: [41022943](https://pubmed.ncbi.nlm.nih.gov/41022943/)]
21. Haug E, Mæland S, Lehmann S, Bjørknes R, Fadnes LT, Sandal GM, et al. Increased gaming during COVID-19 predicts physical inactivity among youth in Norway—a two-wave longitudinal cohort study. *Front Public Health* 2022;10:812932 [FREE Full text] [doi: [10.3389/fpubh.2022.812932](https://doi.org/10.3389/fpubh.2022.812932)] [Medline: [35237551](https://pubmed.ncbi.nlm.nih.gov/35237551/)]
22. Henchoz Y, Studer J, Deline S, N'Goran AA, Baggio S, Gmel G. Video gaming disorder and sport and exercise in emerging adulthood: a longitudinal study. *Behav Med* 2016;42(2):105-111 [FREE Full text] [doi: [10.1080/08964289.2014.965127](https://doi.org/10.1080/08964289.2014.965127)] [Medline: [25258243](https://pubmed.ncbi.nlm.nih.gov/25258243/)]
23. Potoczny W, Herzog-Krzywoszańska R, Krzywoszański Ł. Regular physical activity and life satisfaction: unpacking the roles of self-control and emotion regulation. *Appl Sci* 2025 Jan;15(4):1878 [FREE Full text] [doi: [10.3390/app15041878](https://doi.org/10.3390/app15041878)]
24. Liew L, Stavropoulos V, Adams B, Burleigh T, Griffiths M. Internet gaming disorder: the interplay between physical activity and user–avatar relationship. *Behav Inf Technol* 2018 Apr 30;37(6):558-574 [FREE Full text] [doi: [10.1080/0144929x.2018.1464599](https://doi.org/10.1080/0144929x.2018.1464599)]
25. Saffari M, Huang C, Huang P, Chang Y, Chen J, Poon WC, et al. Mediating roles of weight stigma and physical activity avoidance in the associations between severity of gaming disorder and levels of physical activity among young adults. *J Behav Addict* 2025 Mar 28;14(1):289-303 [FREE Full text] [doi: [10.1556/2006.2024.00083](https://doi.org/10.1556/2006.2024.00083)] [Medline: [39853322](https://pubmed.ncbi.nlm.nih.gov/39853322/)]
26. Yang G, Li Y, Liu S, Liu C, Jia C, Wang S. Physical activity influences the mobile phone addiction among Chinese undergraduates: the moderating effect of exercise type. *J Behav Addict* 2021 Sep 17;10(3):799-810 [FREE Full text] [doi: [10.1556/2006.2021.00059](https://doi.org/10.1556/2006.2021.00059)] [Medline: [34546969](https://pubmed.ncbi.nlm.nih.gov/34546969/)]
27. Gao W, Zhou H, Zou X, Li M, Lin X. The impact of exercise and exercise-based combined therapies on adolescent internet addiction: a systematic review and network meta-analysis. *Curr Psychol* 2025 Mar 24;44(8):7095-7110 [FREE Full text] [doi: [10.1007/s12144-025-07719-w](https://doi.org/10.1007/s12144-025-07719-w)]
28. Nowotarski A, Rothen S, Kasina F, Zullino D, Thorens G. Physical activity as a predictor of internet gaming disorder in a Swiss Male Cohort (C-SURF): l'activité physique comme prédicteur des troubles liés aux jeux vidéo en ligne dans une cohorte de jeunes hommes suisses (C-SURF). *Can J Psychiatry* 2024 Dec;69(12):855-860 [FREE Full text] [doi: [10.1177/07067437241293979](https://doi.org/10.1177/07067437241293979)] [Medline: [39513191](https://pubmed.ncbi.nlm.nih.gov/39513191/)]

29. Huang P, Chen J, Potenza MN, Griffiths MD, Pakpour AH, Chen J, et al. Temporal associations between physical activity and three types of problematic use of the internet: a six-month longitudinal study. *J Behav Addict* 2022 Dec 27;11(4):1055-1067 [FREE Full text] [doi: [10.1556/2006.2022.00084](https://doi.org/10.1556/2006.2022.00084)] [Medline: [36427199](https://pubmed.ncbi.nlm.nih.gov/36427199/)]
30. Hygen BW, Belsky J, Stenseng F, Steinsbekk S, Wichstrøm L, Skalicka V. Longitudinal relations between gaming, physical activity, and athletic self-esteem. *Comput Hum Behav* 2022 Jul;132:107252 [FREE Full text] [doi: [10.1016/j.chb.2022.107252](https://doi.org/10.1016/j.chb.2022.107252)]
31. Hamaker EL, Kuiper RM, Grasman RPPP. A critique of the cross-lagged panel model. *Psychol Methods* 2015 Mar;20(1):102-116 [FREE Full text] [doi: [10.1037/a0038889](https://doi.org/10.1037/a0038889)] [Medline: [25822208](https://pubmed.ncbi.nlm.nih.gov/25822208/)]
32. Goldman N, Khanna D, El Asmar ML, Qualter P, El-Osta A. Addressing loneliness and social isolation in 52 countries: a scoping review of national policies. *BMC Public Health* 2024 May 01;24(1):1207 [FREE Full text] [doi: [10.1186/s12889-024-18370-8](https://doi.org/10.1186/s12889-024-18370-8)] [Medline: [38693471](https://pubmed.ncbi.nlm.nih.gov/38693471/)]
33. Lu H, Ren Z, Zhen P, Su Z, Wu Y. The longitudinal association between family functioning and problematic social media use among Chinese university students: mediation via loneliness and a subgroup analysis by sex. *Addict Behav* 2025 Jul;166:108337 [FREE Full text] [doi: [10.1016/j.addbeh.2025.108337](https://doi.org/10.1016/j.addbeh.2025.108337)] [Medline: [40121924](https://pubmed.ncbi.nlm.nih.gov/40121924/)]
34. Kardefelt-Winther D. A conceptual and methodological critique of internet addiction research: towards a model of compensatory internet use. *Comput Hum Behav* 2014 Feb;31:351-354 [FREE Full text] [doi: [10.1016/j.chb.2013.10.059](https://doi.org/10.1016/j.chb.2013.10.059)]
35. Nowland R, Necka EA, Cacioppo JT. Loneliness and social internet use: pathways to reconnection in a digital world? *Perspect Psychol Sci* 2018 Jan;13(1):70-87 [FREE Full text] [doi: [10.1177/1745691617713052](https://doi.org/10.1177/1745691617713052)] [Medline: [28937910](https://pubmed.ncbi.nlm.nih.gov/28937910/)]
36. Ren Z, Lu H, Zheng P, Lu H, Su Z, Huang Z, et al. Reciprocal relationships among problematic social media use, depression, and social anxiety of Chinese university students: disentangling between- and within-person effects. *Addict Behav* 2026 Jan;172:108515 [FREE Full text] [doi: [10.1016/j.addbeh.2025.108515](https://doi.org/10.1016/j.addbeh.2025.108515)] [Medline: [41056673](https://pubmed.ncbi.nlm.nih.gov/41056673/)]
37. Lee S, Jeong E, Choi J, Park M. Social intelligence and pathological gaming: a longitudinal study of the associations among negative emotions, social intelligence, aggression, and pathological gaming in adolescents. *Front Psychiatry* 2024;15:1353969 [FREE Full text] [doi: [10.3389/fpsy.2024.1353969](https://doi.org/10.3389/fpsy.2024.1353969)] [Medline: [38903650](https://pubmed.ncbi.nlm.nih.gov/38903650/)]
38. Li L, Niu Z, Griffiths MD, Wang W, Chang C, Mei S. A network perspective on the relationship between gaming disorder, depression, alexithymia, boredom, and loneliness among a sample of Chinese university students. *Technol Soc* 2021 Nov;67:101740 [FREE Full text] [doi: [10.1016/j.techsoc.2021.101740](https://doi.org/10.1016/j.techsoc.2021.101740)]
39. Mihalíková V, Dědová M. Association between internet gaming disorder and loneliness: the mediating effect of suicidal and self-injurious behavior in adolescents. *JESP* 2024 Mar 15;14(76)(1):188-201 [FREE Full text] [doi: [10.51865/jesp.2024.1.18](https://doi.org/10.51865/jesp.2024.1.18)]
40. Wang M, Zeng N, Zheng H, Du X, Potenza MN, Dong G. Altered effective connectivity from the pregenual anterior cingulate cortex to the laterobasal amygdala mediates the relationship between internet gaming disorder and loneliness. *Psychol Med* 2022 Mar;52(4):737-746 [FREE Full text] [doi: [10.1017/S0033291720002366](https://doi.org/10.1017/S0033291720002366)] [Medline: [32684185](https://pubmed.ncbi.nlm.nih.gov/32684185/)]
41. Yu Y, Mo PKH, Zhang J, Li J, Lau JTF. Maladaptive cognitions, loneliness, and social anxiety as potential moderators of the association between internet gaming time and internet gaming disorder among adolescent internet gamers in China. *Addict Behav* 2022 Jun;129:107239 [FREE Full text] [doi: [10.1016/j.addbeh.2022.107239](https://doi.org/10.1016/j.addbeh.2022.107239)] [Medline: [35092885](https://pubmed.ncbi.nlm.nih.gov/35092885/)]
42. Yu Y, Fong VWI, Ng JH, Wang Z, Tian X, Lau JTF. The associations between loneliness, hopelessness, and self-control and internet gaming disorder among university students who were men who have sex with men: cross-sectional mediation study. *J Med Internet Res* 2023 Jan 17;25:e43532 [FREE Full text] [doi: [10.2196/43532](https://doi.org/10.2196/43532)] [Medline: [36649059](https://pubmed.ncbi.nlm.nih.gov/36649059/)]
43. Mestre-Bach G, Paiva U, San Martín Iniguez L, Beranuy M, Martín-Vivar M, Mallorqui-Bagué N, et al. The association between internet-use-disorder symptoms and loneliness: a systematic review and meta-analysis with a categorical approach. *Psychol Med* 2025 Mar 12;55:e77 [FREE Full text] [doi: [10.1017/S0033291725000376](https://doi.org/10.1017/S0033291725000376)] [Medline: [40071718](https://pubmed.ncbi.nlm.nih.gov/40071718/)]
44. Zhuang X, Zhang Y, Tang X, Ng TK, Lin J, Yang X. Longitudinal modifiable risk and protective factors of internet gaming disorder: a systematic review and meta-analysis. *J Behav Addict* 2023 Jun 29;12(2):375-392 [FREE Full text] [doi: [10.1556/2006.2023.00017](https://doi.org/10.1556/2006.2023.00017)] [Medline: [37224007](https://pubmed.ncbi.nlm.nih.gov/37224007/)]
45. Gao B, Cai Y, Zhao C, Qian Y, Zheng R, Liu C. Longitudinal associations between loneliness and online game addiction among undergraduates: a moderated mediation model. *Acta Psychol (Amst)* 2024 Mar;243:104134 [FREE Full text] [doi: [10.1016/j.actpsy.2024.104134](https://doi.org/10.1016/j.actpsy.2024.104134)] [Medline: [38228073](https://pubmed.ncbi.nlm.nih.gov/38228073/)]
46. Vuorinen I, Savolainen I, Sirola A, Oksanen A. The impacts of stress and loneliness on gambling and gaming problems: a nationwide longitudinal study. *Int J Soc Psychiatry* 2024 Nov;70(7):1325-1332 [FREE Full text] [doi: [10.1177/00207640241264661](https://doi.org/10.1177/00207640241264661)] [Medline: [39049602](https://pubmed.ncbi.nlm.nih.gov/39049602/)]
47. Lemmens JS, Valkenburg PM, Peter J. Psychosocial causes and consequences of pathological gaming. *Comput Hum Behav* 2011 Jan;27(1):144-152 [FREE Full text] [doi: [10.1016/j.chb.2010.07.015](https://doi.org/10.1016/j.chb.2010.07.015)]
48. Ren Z, Lu H, Ma X, Min G, Liu L, Wu L, et al. Longitudinal association between problematic social media use and sleep problems among adolescents: a random intercept cross-lagged panel model and sex differences. *J Youth Adolesc* 2026 Jan 26;55(1):119-134 [FREE Full text] [doi: [10.1007/s10964-025-02310-9](https://doi.org/10.1007/s10964-025-02310-9)] [Medline: [41452419](https://pubmed.ncbi.nlm.nih.gov/41452419/)]
49. 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 [FREE Full text] [doi: [10.1037//0003-066x.55.1.68](https://doi.org/10.1037//0003-066x.55.1.68)] [Medline: [11392867](https://pubmed.ncbi.nlm.nih.gov/11392867/)]

50. Li X, Zhu J, Yu H, Lin J. Effect of perceived autonomy supports on exercise persistence for adolescents: an integrated model based on basic psychological needs theory and the theory of planned behavior. *Front Psychol* 2025;16:1692940 [FREE Full text] [doi: [10.3389/fpsyg.2025.1692940](https://doi.org/10.3389/fpsyg.2025.1692940)] [Medline: [41459258](https://pubmed.ncbi.nlm.nih.gov/41459258/)]
51. Ntoumanis N, Moller AC. Self-determination theory informed research for promoting physical activity: contributions, debates, and future directions. *Psychol Sport Exerc* 2025 Sep;80:102879 [FREE Full text] [doi: [10.1016/j.psychsport.2025.102879](https://doi.org/10.1016/j.psychsport.2025.102879)] [Medline: [40383282](https://pubmed.ncbi.nlm.nih.gov/40383282/)]
52. 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 Jun 22;9:78 [FREE Full text] [doi: [10.1186/1479-5868-9-78](https://doi.org/10.1186/1479-5868-9-78)] [Medline: [22726453](https://pubmed.ncbi.nlm.nih.gov/22726453/)]
53. Sebastião E, Mirda D. Group-based physical activity as a means to reduce social isolation and loneliness among older adults. *Aging Clin Exp Res* 2021 Jul;33(7):2003-2006 [FREE Full text] [doi: [10.1007/s40520-020-01722-w](https://doi.org/10.1007/s40520-020-01722-w)] [Medline: [33387363](https://pubmed.ncbi.nlm.nih.gov/33387363/)]
54. Zhou F, Wang W, Wu J, Nie Y, Shao C, Qiu W, et al. Body image and loneliness as mediators of the relationship between physical activity and exercise self-efficacy in college students. *Sci Rep* 2025 Aug 21;15(1):30782 [FREE Full text] [doi: [10.1038/s41598-025-16307-8](https://doi.org/10.1038/s41598-025-16307-8)] [Medline: [40841721](https://pubmed.ncbi.nlm.nih.gov/40841721/)]
55. Wang K, Li Y, Yang Y, Zhang T, Luo J. The role of loneliness and learning burnout in the regulation of physical exercise on internet addiction in Chinese college students. *BMC Public Health* 2023 Oct 12;23(1):1994 [FREE Full text] [doi: [10.1186/s12889-023-16783-5](https://doi.org/10.1186/s12889-023-16783-5)] [Medline: [37828509](https://pubmed.ncbi.nlm.nih.gov/37828509/)]
56. Xu J, Tang L. The relationship between physical exercise and problematic internet use in college students: the chain-mediated role of self-control and loneliness. *BMC Public Health* 2024 Jun 27;24(1):1719 [FREE Full text] [doi: [10.1186/s12889-024-19226-x](https://doi.org/10.1186/s12889-024-19226-x)] [Medline: [38937729](https://pubmed.ncbi.nlm.nih.gov/38937729/)]
57. Zhu X, Niu H, Zhu Q, Chen W. Physical activity and mobile phone addiction among Chinese college students: a chain mediation model of social anxiety and loneliness. *Front Psychol* 2025;16:1545440 [FREE Full text] [doi: [10.3389/fpsyg.2025.1545440](https://doi.org/10.3389/fpsyg.2025.1545440)] [Medline: [40330309](https://pubmed.ncbi.nlm.nih.gov/40330309/)]
58. Ahn J, Falk EB, Kang Y. Relationships between physical activity and loneliness: a systematic review of intervention studies. *Curr Res Behav Sci* 2024;6:100141 [FREE Full text] [doi: [10.1016/j.crbeha.2023.100141](https://doi.org/10.1016/j.crbeha.2023.100141)]
59. Pels F, Kleinert J. Loneliness and physical activity: a systematic review. *Int Rev Sport Exerc Psychol* 2016 Jun 07;9(1):231-260 [FREE Full text] [doi: [10.1080/1750984x.2016.1177849](https://doi.org/10.1080/1750984x.2016.1177849)]
60. Gu S, Li S. Bidirectional dynamics: a cross-lagged examination of physical exercise and loneliness in older adults. *J Affect Disord* 2025 Nov 01;388:119635 [FREE Full text] [doi: [10.1016/j.jad.2025.119635](https://doi.org/10.1016/j.jad.2025.119635)] [Medline: [40494500](https://pubmed.ncbi.nlm.nih.gov/40494500/)]
61. Rodeiro J, Olaya B, Haro JM, Gabarrell-Pascuet A, Ayuso-Mateos JL, Francia L, et al. The longitudinal relationship among physical activity, loneliness, and mental health in middle-aged and older adults: results from the Edad con Salud cohort. *Ment Health Phys Act* 2025 Mar;28:100667 [FREE Full text] [doi: [10.1016/j.mhpa.2024.100667](https://doi.org/10.1016/j.mhpa.2024.100667)]
62. Groß D, Haffa J. Unpacking the temporal dynamics of physical activity and loneliness: a 1-year interval analysis. *Am J Epidemiol* 2025 Aug 05;194(8):2134-2139 [FREE Full text] [doi: [10.1093/aje/kwaf077](https://doi.org/10.1093/aje/kwaf077)] [Medline: [40239987](https://pubmed.ncbi.nlm.nih.gov/40239987/)]
63. Fam JY. Prevalence of internet gaming disorder in adolescents: a meta-analysis across three decades. *Scand J Psychol* 2018 Oct;59(5):524-531 [FREE Full text] [doi: [10.1111/sjop.12459](https://doi.org/10.1111/sjop.12459)] [Medline: [30004118](https://pubmed.ncbi.nlm.nih.gov/30004118/)]
64. Yu Y, Mo PKH, Zhang J, Li J, Lau JTF. Why is internet gaming disorder more prevalent among Chinese male than female adolescents? The role of cognitive mediators. *Addict Behav* 2021 Jan;112:106637 [FREE Full text] [doi: [10.1016/j.addbeh.2020.106637](https://doi.org/10.1016/j.addbeh.2020.106637)] [Medline: [32919322](https://pubmed.ncbi.nlm.nih.gov/32919322/)]
65. Zhihao D, Tao W, Yingjie S, Feng Z. The influence of physical activity on internet addiction among Chinese college students: the mediating role of self-esteem and the moderating role of gender. *BMC Public Health* 2024 Apr 01;24(1):935 [FREE Full text] [doi: [10.1186/s12889-024-18474-1](https://doi.org/10.1186/s12889-024-18474-1)] [Medline: [38561700](https://pubmed.ncbi.nlm.nih.gov/38561700/)]
66. Jiang W, Luo J, Guan H. Gender difference in the relationship of physical activity and subjective happiness among Chinese university students. *Front Psychol* 2021 Dec 7;12:800515 [FREE Full text] [doi: [10.3389/fpsyg.2021.800515](https://doi.org/10.3389/fpsyg.2021.800515)] [Medline: [34950093](https://pubmed.ncbi.nlm.nih.gov/34950093/)]
67. Tubić T, Modrić T, Sekulić D, Bianco A, Radjo I, Drid P. Loneliness in sports active and non-active school-age children: can sport protect children against loneliness? *Front Psychiatry* 2022;13:1063714 [FREE Full text] [doi: [10.3389/fpsyg.2022.1063714](https://doi.org/10.3389/fpsyg.2022.1063714)] [Medline: [36713906](https://pubmed.ncbi.nlm.nih.gov/36713906/)]
68. Wang R, Yang S, Yan Y, Tian Y, Wang P. Internet gaming disorder in early adolescents: gender and depression differences in a latent growth model. *Healthcare (Basel)* 2021 Sep 09;9(9):1188 [FREE Full text] [doi: [10.3390/healthcare9091188](https://doi.org/10.3390/healthcare9091188)] [Medline: [34574962](https://pubmed.ncbi.nlm.nih.gov/34574962/)]
69. Yang X, Jiang X, Wu AMS, Ma L, Cai Y, Wong KM, et al. Validation of the Internet Gaming Disorder Symptoms Checklist based on the Fifth Edition of the Diagnostic and Statistical Manual of Mental Disorders in Chinese adolescents. *Child Psychiatry Hum Dev* 2023 Feb;54(1):26-33 [FREE Full text] [doi: [10.1007/s10578-021-01213-7](https://doi.org/10.1007/s10578-021-01213-7)] [Medline: [34338962](https://pubmed.ncbi.nlm.nih.gov/34338962/)]
70. Russell DW. UCLA Loneliness Scale (version 3): reliability, validity, and factor structure. *J Pers Assess* 1996 Feb;66(1):20-40 [FREE Full text] [doi: [10.1207/s15327752jpa6601_2](https://doi.org/10.1207/s15327752jpa6601_2)] [Medline: [8576833](https://pubmed.ncbi.nlm.nih.gov/8576833/)]
71. Liang D. The relationship between stress level and physical exercise for college students. *Chin Mental Health J* 1994;8(1):5-6.

72. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equation Modell Multidiscip* 2002 Apr;9(2):233-255 [[FREE Full text](#)] [doi: [10.1207/S15328007SEM0902_5](https://doi.org/10.1207/S15328007SEM0902_5)]
73. McDonald RP, Ho MR. Principles and practice in reporting structural equation analyses. *Psychol Methods* 2002;7(1):64-82 [[FREE Full text](#)] [doi: [10.1037/1082-989X.7.1.64](https://doi.org/10.1037/1082-989X.7.1.64)]
74. Enders C, Bandalos D. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Struct Equation Modell Multidiscip* 2001 Jul 1;8(3):430-457 [[FREE Full text](#)] [doi: [10.1207/s15328007sem0803_5](https://doi.org/10.1207/s15328007sem0803_5)]
75. Nicholson JS, Deboeck PR, Howard W. Attrition in developmental psychology: a review of modern missing data reporting and practices. *Int J Behav Dev* 2015 Dec 17;41(1):143-153 [[FREE Full text](#)] [doi: [10.1177/0165025415618275](https://doi.org/10.1177/0165025415618275)]
76. Zhou J, Li T, Li S, Yang P, Gong X. Bidirectional relations between belief in a just world and internalizing and externalizing problems among Chinese secondary vocational students. *J Youth Adolesc* 2025 Oct 17;54(10):2644-2659 [[FREE Full text](#)] [doi: [10.1007/s10964-025-02259-9](https://doi.org/10.1007/s10964-025-02259-9)] [Medline: [40960704](https://pubmed.ncbi.nlm.nih.gov/40960704/)]
77. Mahmud S, Jobayer MAA, Salma N, Mahmud A, Tamanna T. Online gaming and its effect on academic performance of Bangladeshi university students: a cross-sectional study. *Health Sci Rep* 2023 Dec;6(12):e1774 [[FREE Full text](#)] [doi: [10.1002/hsr2.1774](https://doi.org/10.1002/hsr2.1774)] [Medline: [38107152](https://pubmed.ncbi.nlm.nih.gov/38107152/)]
78. Orth U, Meier LL, Bühler JL, Dapp LC, Krauss S, Messerli D, et al. Effect size guidelines for cross-lagged effects. *Psychol Methods* 2024 Apr;29(2):421-433 [[FREE Full text](#)] [doi: [10.1037/met0000499](https://doi.org/10.1037/met0000499)] [Medline: [35737548](https://pubmed.ncbi.nlm.nih.gov/35737548/)]
79. Mulder JD, Hamaker EL. Three extensions of the random intercept cross-lagged panel model. *Struct Equation Modell Multidiscip* 2020 Aug 11;28(4):638-648 [[FREE Full text](#)] [doi: [10.1080/10705511.2020.1784738](https://doi.org/10.1080/10705511.2020.1784738)]
80. Ibabe I, Albertos A, Lopez-Del Burgo C. Leisure time activities in adolescents predict problematic technology use. *Eur Child Adolesc Psychiatry* 2024 Jan 15;33(1):279-289 [[FREE Full text](#)] [doi: [10.1007/s00787-023-02152-5](https://doi.org/10.1007/s00787-023-02152-5)] [Medline: [36790555](https://pubmed.ncbi.nlm.nih.gov/36790555/)]
81. Eime RM, Young JA, Harvey JT, Charity MJ, Payne WR. A systematic review of the psychological and social benefits of participation in sport for children and adolescents: informing development of a conceptual model of health through sport. *Int J Behav Nutr Phys Act* 2013 Aug 15;10:98 [[FREE Full text](#)] [doi: [10.1186/1479-5868-10-98](https://doi.org/10.1186/1479-5868-10-98)] [Medline: [23945179](https://pubmed.ncbi.nlm.nih.gov/23945179/)]
82. Xing Z, Ge C. The relationship between physical exercise and social adjustment in Chinese university students: the sequential mediating effect of peer attachment and self-esteem. *Front Psychol* 2025;16:1525811 [[FREE Full text](#)] [doi: [10.3389/fpsyg.2025.1525811](https://doi.org/10.3389/fpsyg.2025.1525811)] [Medline: [40432802](https://pubmed.ncbi.nlm.nih.gov/40432802/)]
83. Peng B, Chen W, Wang H, Yu T. How does physical exercise influence self-efficacy in adolescents? A study based on the mediating role of psychological resilience. *BMC Psychol* 2025 Mar 21;13(1):285 [[FREE Full text](#)] [doi: [10.1186/s40359-025-02529-y](https://doi.org/10.1186/s40359-025-02529-y)]
84. Wu J, Zhu L, Dong X, Sun Z, Cai K, Shi Y, et al. Relationship between physical activity and emotional regulation strategies in early adulthood: mediating effects of cortical thickness. *Brain Sci* 2022 Sep 08;12(9):1210 [[FREE Full text](#)] [doi: [10.3390/brainsci12091210](https://doi.org/10.3390/brainsci12091210)] [Medline: [36138946](https://pubmed.ncbi.nlm.nih.gov/36138946/)]
85. Salmensalo M, Ruotsalainen H, Hylkilä K, Kääriäinen M, Konttila J, Männistö M, et al. Associations between digital gaming behavior and physical activity among Finnish vocational students. *J Public Health (Berl)* 2022 Dec 11;32(1):53-63 [[FREE Full text](#)] [doi: [10.1007/s10389-022-01788-y](https://doi.org/10.1007/s10389-022-01788-y)]
86. Khan A, Uddin R, Burton NW. Insufficient physical activity in combination with high screen time is associated with adolescents' psychosocial difficulties. *Int Health* 2018 Jul 01;10(4):246-251 [[FREE Full text](#)] [doi: [10.1093/inthealth/ihy019](https://doi.org/10.1093/inthealth/ihy019)] [Medline: [29579220](https://pubmed.ncbi.nlm.nih.gov/29579220/)]
87. García-Ortiz C, Lorenzo-González M, Fernández-Sánchez J, Solano-Lizcano V, Del Coso J, Collado-Mateo D. Recommendations for physical exercise as a strategy to reduce problematic use of the internet and digital devices: a perspective. *Int J Environ Res Public Health* 2025 May 10;22(5):753 [[FREE Full text](#)] [doi: [10.3390/ijerph22050753](https://doi.org/10.3390/ijerph22050753)] [Medline: [40427868](https://pubmed.ncbi.nlm.nih.gov/40427868/)]
88. Lai NM, Lim YS, Chaiyakunapruk N, Lee SWH, Liabsuetrakul T, Sabbir M, et al. Review article: The effectiveness of school-based interventions for reducing screen time - a systematic review and meta-analysis. *Child Adolesc Ment Health* 2025 Sep;30(3):223-237 [[FREE Full text](#)] [doi: [10.1111/camh.70022](https://doi.org/10.1111/camh.70022)] [Medline: [40660635](https://pubmed.ncbi.nlm.nih.gov/40660635/)]
89. Tock WL, Gauvin L, Hudon F, Tremblay F, Bélanger RE, Turcotte-Tremblay A. Scoping review of interventions aimed at promoting healthy screen use among adolescents. *BMJ Open* 2025 Sep 21;15(9):e103772 [[FREE Full text](#)] [doi: [10.1136/bmjopen-2025-103772](https://doi.org/10.1136/bmjopen-2025-103772)] [Medline: [40976666](https://pubmed.ncbi.nlm.nih.gov/40976666/)]
90. Yue Z, Bickham D, Schwamm S, Rich M. Who they play with matters: an exploration of behavioral pathways linking loneliness, social anxiety, and problematic gaming. *J Adolesc Health* 2026 Feb;78(2):294-301 [[FREE Full text](#)] [doi: [10.1016/j.jadohealth.2025.09.029](https://doi.org/10.1016/j.jadohealth.2025.09.029)] [Medline: [41171243](https://pubmed.ncbi.nlm.nih.gov/41171243/)]
91. Kim J, Choi J, Lee S, Jeong E. The influence of negative emotions on adolescents' pathological gaming: the role of self-control in the associations with aggression, academic stress, and pathological gaming. *Front Psychol* 2025 Jul 2;16:1539162 [[FREE Full text](#)] [doi: [10.3389/fpsyg.2025.1539162](https://doi.org/10.3389/fpsyg.2025.1539162)] [Medline: [40672844](https://pubmed.ncbi.nlm.nih.gov/40672844/)]
92. Barreto Carvalho C, Moura Cabral J, Teixeira M, Cordeiro F, Costa R, Moura Arroz A. "Belonging without being": relationships between problematic gaming, internet use, and social group attachment in adolescence. *Comput Hum Behav* 2023 Dec;149:107932 [[FREE Full text](#)] [doi: [10.1016/j.chb.2023.107932](https://doi.org/10.1016/j.chb.2023.107932)]

93. Bowditch L, Chapman J, Signal T, Naweed A. Gamers in real life: internet gamer typologies and their association with problematic gaming and health outcomes. *Hum Behav Emerging Technol* 2025 Apr 9;2025(1):2368683 [[FREE Full text](#)] [doi: [10.1155/hbe2/2368683](https://doi.org/10.1155/hbe2/2368683)]
94. Wang Y, Zeng Y. Relationship between loneliness and internet addiction: a meta-analysis. *BMC Public Health* 2024 Mar 19;24(1):858 [[FREE Full text](#)] [doi: [10.1186/s12889-024-18366-4](https://doi.org/10.1186/s12889-024-18366-4)] [Medline: [38504216](https://pubmed.ncbi.nlm.nih.gov/38504216/)]
95. Yayla ?, Dombak K, Diril S, Düşünceli B, Çelik E, Yildirim M. Difficulties in emotion regulation as a mediator and gender as a moderator in the relationship between problematic digital gaming and life satisfaction among adolescents. *Behav Sci (Basel)* 2025 Aug 12;15(8):1092 [[FREE Full text](#)] [doi: [10.3390/bs15081092](https://doi.org/10.3390/bs15081092)] [Medline: [40867449](https://pubmed.ncbi.nlm.nih.gov/40867449/)]
96. Ji W, Ding L, Li X, Wu Y, Li H. Why is problematic social media use more prevalent among Chinese adolescents with probable depression than their non-depressed peers? A longitudinal examination of the mediating roles of peer acceptance and loneliness. *Addict Behav* 2026 Jun;177:108640 [[FREE Full text](#)] [doi: [10.1016/j.addbeh.2026.108640](https://doi.org/10.1016/j.addbeh.2026.108640)] [Medline: [41698284](https://pubmed.ncbi.nlm.nih.gov/41698284/)]
97. Lai B, Young R, Craig M, Chaviano K, Swanson-Kimani E, Wozow C, et al. Improving social isolation and loneliness among adolescents with physical disabilities through group-based virtual reality gaming: feasibility pre-post trial study. *JMIR Form Res* 2023 Dec 06;7:e47630 [[FREE Full text](#)] [doi: [10.2196/47630](https://doi.org/10.2196/47630)] [Medline: [38055309](https://pubmed.ncbi.nlm.nih.gov/38055309/)]
98. Rangul V, Sund ER, Ingul JM, Rimehaug T, Pape K, Kvaløy K. Exploring the link between physical activity, sports participation, and loneliness in adolescents before and into the COVID-19 pandemic: the HUNT study, Norway. *Int J Environ Res Public Health* 2024 Oct 25;21(11):1417 [[FREE Full text](#)] [doi: [10.3390/ijerph21111417](https://doi.org/10.3390/ijerph21111417)] [Medline: [39595684](https://pubmed.ncbi.nlm.nih.gov/39595684/)]
99. Lu H, Yu Y, Wang DB, Wu AM, Chen JH, Zhang G, et al. Association between interpersonal resources and mental health professional help-seeking among Chinese adolescents with probable depression: mediations via personal resources and active coping. *BMC Psychiatry* 2024 Nov 21;24(1):840 [[FREE Full text](#)] [doi: [10.1186/s12888-024-06271-4](https://doi.org/10.1186/s12888-024-06271-4)] [Medline: [39574049](https://pubmed.ncbi.nlm.nih.gov/39574049/)]
100. Vansteenkiste M, Ryan RM, Soenens B. Basic psychological need theory: advancements, critical themes, and future directions. *Motiv Emot* 2020 Jan 21;44(1):1-31 [[FREE Full text](#)] [doi: [10.1007/s11031-019-09818-1](https://doi.org/10.1007/s11031-019-09818-1)]
101. Mahon A, Holt C, Hibbert M, Jago R, Tinner L, Kidger J. A scoping review of physical activity interventions to support the mental health of students aged 16 to 25. *Discov Public Health* 2025 Sep 12;22(1):538 [[FREE Full text](#)] [doi: [10.1186/s12982-025-00933-8](https://doi.org/10.1186/s12982-025-00933-8)] [Medline: [40950509](https://pubmed.ncbi.nlm.nih.gov/40950509/)]
102. Francisco-Sánchez A, Martínez-León S, García-Pérez A, Báez-Hernández JA, Rodríguez-Álvaro M, García-Hernández AM. Impact of a multidimensional community-based intervention on the feeling of unwanted loneliness and its consequences: a quasi-experimental study. *Healthcare (Basel)* 2025 Jun 18;13(12):1465 [[FREE Full text](#)] [doi: [10.3390/healthcare13121465](https://doi.org/10.3390/healthcare13121465)] [Medline: [40565491](https://pubmed.ncbi.nlm.nih.gov/40565491/)]
103. Fan H, Yan J, Yang Z, Liang K, Chen S. Cross-sectional associations between screen time and the selected lifestyle behaviors in adolescents. *Front Public Health* 2022 Sep 27;10:932017 [[FREE Full text](#)] [doi: [10.3389/fpubh.2022.932017](https://doi.org/10.3389/fpubh.2022.932017)] [Medline: [36238246](https://pubmed.ncbi.nlm.nih.gov/36238246/)]
104. Haug E. Adolescents' screen-based media use and the relationship with moderate-to-vigorous physical activity, sports club participation and active commuting. *Scand J Public Health* 2025 Apr;53(Suppl 1):42-51 [[FREE Full text](#)] [doi: [10.1177/14034948241293603](https://doi.org/10.1177/14034948241293603)] [Medline: [39673495](https://pubmed.ncbi.nlm.nih.gov/39673495/)]
105. Zablotzky B, Ng AE, Black LI, Haile G, Bose J, Jones JR, et al. Associations between screen time use and health outcomes among US teenagers. *Prev Chronic Dis* 2025 Jul 10;22:E38 [[FREE Full text](#)] [doi: [10.5888/pcd22.240537](https://doi.org/10.5888/pcd22.240537)] [Medline: [40638804](https://pubmed.ncbi.nlm.nih.gov/40638804/)]
106. Carter M. Gender socialization and identity theory. *Soc Sci* 2014 May 12;3(2):242-263 [[FREE Full text](#)] [doi: [10.3390/socsci3020242](https://doi.org/10.3390/socsci3020242)]
107. Hoominfar E. Gender socialization. In: *Gender Equality*. Cham: Springer Nature; 2021:645-654.
108. Gutman LM, Codiroli McMaster N. Gendered pathways of internalizing problems from early childhood to adolescence and associated adolescent outcomes. *J Abnorm Child Psychol* 2020 May;48(5):703-718 [[FREE Full text](#)] [doi: [10.1007/s10802-020-00623-w](https://doi.org/10.1007/s10802-020-00623-w)] [Medline: [32040796](https://pubmed.ncbi.nlm.nih.gov/32040796/)]
109. Lacko D, Kyslík F, Smahel D, Machackova H. Gaming together, feeling better—or feeling worse? How social video gaming impacts loneliness and depressive mood differently for boys and girls. *Comput Hum Behav* 2025 Nov;172:108752 [[FREE Full text](#)] [doi: [10.1016/j.chb.2025.108752](https://doi.org/10.1016/j.chb.2025.108752)]
110. Marraudino M, Bonaldo B, Vitiello B, Bergui GC, Panzica G. Sexual differences in internet gaming disorder (IGD): from psychological features to neuroanatomical networks. *J Clin Med* 2022 Feb 16;11(4):1018 [[FREE Full text](#)] [doi: [10.3390/jcm11041018](https://doi.org/10.3390/jcm11041018)] [Medline: [35207293](https://pubmed.ncbi.nlm.nih.gov/35207293/)]

Abbreviations

BPNT: basic psychological needs theory

CFI: comparative fit index

CLPM: cross-lagged panel model

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

FIML: full information maximum likelihood
IGD: internet gaming disorder
MCAR: Missing Completely at Random
PA: physical activity
RI: random intercept
RI-CLPM: random intercept cross-lagged panel model
RMSEA: root mean square error of approximation
SEM: structural equation modeling
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index
UCLA: University of California, Los Angeles

Edited by S Brini; submitted 15.Nov.2025; peer-reviewed by D Lacko, T Baranowski; comments to author 05.Mar.2026; revised version received 08.Apr.2026; accepted 08.Apr.2026; published 27.Apr.2026.

Please cite as:

Hu C, Meng X, Wang X, Huang J

Reciprocal Within-Person Dynamics Between Internet Gaming Disorder Symptoms, Physical Activity, and Loneliness Among Chinese Adolescent Gamers: Three-Wave Prospective Cohort Study

JMIR Serious Games 2026;14:e87847

URL: <https://games.jmir.org/2026/1/e87847>

doi: [10.2196/87847](https://doi.org/10.2196/87847)

PMID:

©Chang Hu, Xuan Meng, Xi Wang, Jilin Huang. Originally published in JMIR Serious Games (<https://games.jmir.org>), 27.Apr.2026. 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.

Correction: Tongue Muscle Training App for Middle-Aged and Older Adults Incorporating Flow-Based Gameplay: Design and Feasibility Pilot Study

Kuan-Chu Su¹, MA; Ko-Chiu Wu², PhD; Kuei-Ru Chou³, PhD; Chia-Hsu Huang², MDes

¹College of Design, National Taipei University of Technology, Taipei, Taiwan

²Department of Interaction Design, National Taipei University of Technology, Rm.701-4, Design Building, No.1, Sec.3, Chung-hsiao E. Rd, Taipei, Taiwan

³School of Nursing, Taipei Medical University, Taipei, Taiwan

Corresponding Author:

Ko-Chiu Wu, PhD

Department of Interaction Design, National Taipei University of Technology, Rm.701-4, Design Building, No.1, Sec.3, Chung-hsiao E. Rd, Taipei, Taiwan

Related Article:

Correction of: <https://games.jmir.org/2025/1/e53045>

Abstract

(*JMIR Serious Games* 2026;14:e90410) doi:[10.2196/90410](https://doi.org/10.2196/90410)

In “Tongue Muscle Training App for Middle-Aged and Older Adults Incorporating Flow-Based Gameplay: Design and Feasibility Pilot Study” [1], the authors made one change.

The institutional affiliation of author KCS has been changed from the following:

Department of Interaction Design, National Taipei University of Technology, Taipei, Taiwan

The affiliation now reads :

College of Design, National Taipei University of Technology, Taipei, Taiwan

The correction will appear in the online version of the paper on the JMIR Publications website, 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.

Reference

1. Su KC, Wu KC, Chou KR, Huang CH. Tongue muscle training app for middle-aged and older adults incorporating flow-based gameplay: design and feasibility pilot study. *JMIR Serious Games* 2025 Jan 9;13:e53045. [doi: [10.2196/53045](https://doi.org/10.2196/53045)] [Medline: [39791331](https://pubmed.ncbi.nlm.nih.gov/39791331/)]

Submitted 27.Dec.2025; this is a non-peer-reviewed article; accepted 29.Dec.2025; published 16.Jan.2026.

Please cite as:

Su KC, Wu KC, Chou KR, Huang CH

Correction: Tongue Muscle Training App for Middle-Aged and Older Adults Incorporating Flow-Based Gameplay: Design and Feasibility Pilot Study

JMIR Serious Games 2026;14:e90410

URL: <https://games.jmir.org/2026/1/e90410>

doi: [10.2196/90410](https://doi.org/10.2196/90410)

© Kuan-Chu Su, Ko-Chiu Wu, Kuei-Ru Chou, Chia-Hsu Huang. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 16.Jan.2026. 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.

Corrigenda and Addenda

Correction: Children's Improvement After Language and Rhythm Training With the Digital Medical Device Poppins for Dyslexia: Single-Arm Intervention Study

Charline Grossard^{1,2}, PhD; Melanie Descamps², PhD; Hugues Pellerin¹, MA; François Vonthron², MA; David Cohen¹, Prof Dr Med

¹Department of Child and Adolescent Psychiatry, Pitié-Salpêtrière Hospital, Paris, France

²Poppins, Paris, France

Corresponding Author:

Charline Grossard, PhD

Department of Child and Adolescent Psychiatry

Pitié-Salpêtrière Hospital

47-83 boulevard de l'Hôpital

Paris, 75013

France

Phone: 33 142162383

Email: charline.grossard@aphp.fr

Related Article:

Correction of: <https://games.jmir.org/2025/1/e76435>

(*JMIR Serious Games* 2026;14:e93737) doi:[10.2196/93737](https://doi.org/10.2196/93737)

In "Children's Improvement After Language and Rhythm Training With the Digital Medical Device Poppins for Dyslexia: Single-Arm Intervention Study" [1], the authors made one correction.

In [Table 2](#), the score for the subtest comprehension at T2 has been revised from:

15.45

To:

5.45

[Table 2](#) has been revised to the following:

Table 2. Comparison of participants' scores on the proposed tasks at pretest (T1) and posttest (T2).

Tasks	T2 (n=37), mean (SD)	T1 (n=38), mean (SD)	Difference (95% CI)	P value
Words read in 2 minutes (EVAL2M)				
Words correctly read	101.49 (33.86)	88.39 (34.00)	11.46 (8.75 to 14.25)	<.001
Words read	110.46 (32.43)	98.50 (32.54)	10.26 (6.77 to 13.85)	<.001
Phoneme deletion (BALE)				
Total score	16.58 (3.38)	13.54 (4.51)	2.87 (1.64 to 4.09)	<.001
Total time (seconds)	205.19 (101.37)	213.14 (102.56)	-5.37 (-31.97 to 21)	.69
Text with no meaning (Evalouette)				
Words read	105.59 (42.94)	93.21 (38.76)	10.95 (4.87 to 17.19)	<.001
Words correctly read	94.49 (42.80)	79.76 (38.92)	13.25 (7.78 to 18.86)	<.001
Phonological discrimination (Evaleo 6-15)				
Total score	21.84 (3.95)	20.21 (3.86)	1.61 (0.2 to 3.06)	.03
Text comprehension 4th and 5th grades				
Number of reading errors	5.50 (5.73)	4.50 (4.88)	1.00 (-0.72 to 2.67)	.25
Reading time (seconds)	97.76 (41.51)	85.29 (37.46)	12.46 (4.96 to 19.76)	.001
Comprehension score	5.45 (1.50)	4.75 (1.12)	0.70 (0.06 to 1.32)	.03

The correction will appear in the online version of the paper on the JMIR Publications website, 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.

Reference

- Grossard C, Descamps M, Pellerin H, Vonthron F, Cohen D. Children's Improvement After Language and Rhythm Training With the Digital Medical Device Poppins for Dyslexia: Single-Arm Intervention Study. *JMIR Serious Games* 2025 Aug 01;13:e76435 [FREE Full text] [doi: [10.2196/76435](https://doi.org/10.2196/76435)] [Medline: [40750096](https://pubmed.ncbi.nlm.nih.gov/40750096/)]

Submitted 18.Feb.2026; this is a non-peer-reviewed article; accepted 26.Feb.2026; published 12.Mar.2026.

Please cite as:

Grossard C, Descamps M, Pellerin H, Vonthron F, Cohen D

Correction: Children's Improvement After Language and Rhythm Training With the Digital Medical Device Poppins for Dyslexia: Single-Arm Intervention Study

JMIR Serious Games 2026;14:e93737

URL: <https://games.jmir.org/2026/1/e93737>

doi: [10.2196/93737](https://doi.org/10.2196/93737)

PMID:

©Charline Grossard, Melanie Descamps, Hugues Pellerin, François Vonthron, David Cohen. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 12.Mar.2026. 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.

Research Design Processes in Serious Games for Adolescent Mental Health: Systematic Review

S T Roosj, de¹, MSc; D A Kuipers¹, PhD; J T Prins², PhD; J P E N Pierie³, Prof Dr

¹Lectorate of Design Driven Innovation, School of Design, NHL Stenden University of Applied Sciences, Rengerslaan 8-10, Leeuwarden, The Netherlands

²Wenckebach Institute, University Medical Center Groningen, Groningen, The Netherlands

³Faculty of Medical Sciences, University Medical Center Groningen, Groningen, The Netherlands

Corresponding Author:

S T Roosj, de, MSc

Lectorate of Design Driven Innovation, School of Design, NHL Stenden University of Applied Sciences, Rengerslaan 8-10, Leeuwarden, The Netherlands

Abstract

Background: Serious games are increasingly recognized as effective tools in adolescent mental health interventions, providing engaging platforms for emotional regulation, skill development, and behavioral change. However, the ways in which core theoretical concepts such as transfer, boundary crossing, and models of reality are incorporated into serious game designs are not consistently described in the literature. Clarifying how these concepts are addressed is important for understanding how game-based learning may connect to real-world health care practice.

Objective: This systematic review aims to examine how serious games for adolescent health care are designed to support learning and facilitate outcomes. Specifically, it examines how the design incorporates constructs of transfer, boundary crossing, and models of reality, and how these elements are represented across published studies.

Methods: We conducted a systematic search across 5 databases (PubMed, Scopus, ERIC, PsycINFO, and EMBASE) covering publications up to 2023. Studies were included if they involved serious games targeting adolescents with behavioral or developmental health concerns. Titles and abstracts were screened independently by 2 reviewers, with disagreements resolved by a third party. A qualitative analytical framework was applied to identify elements of design, with a particular focus on transfer, boundary crossing, and models of reality.

Results: Thirty-three studies met the inclusion criteria. Figural transfer was identified in 24 studies, while literal transfer was identified in 10 studies. Among boundary-crossing mechanisms, reflection occurred most frequently (22 studies), whereas transformation was observed in 3 studies. Causal and procedural models of reality were most commonly identified as primary model types, whereas relational and structural models were more often reported as secondary. Explicit design rationales were infrequently reported across studies.

Conclusions: This review demonstrates that serious games for adolescent mental health most frequently emphasize reflective and representational forms of learning. Across the reviewed studies, theoretical constructs related to transfer, boundary crossing, and models of reality were often implicitly embedded rather than explicitly articulated. The proposed analytical framework offers a structured approach for analyzing these design characteristics and may support designers, researchers, and health care professionals in more explicitly aligning design choices with intended learning mechanisms and real-world applications.

(*JMIR Serious Games* 2026;14:e77173) doi:[10.2196/77173](https://doi.org/10.2196/77173)

KEYWORDS

serious games; serious gaming; adolescent mental health; learning transfer; boundary crossing; model of reality; design rationale; game-based learning; health care innovation

Introduction

Context and Rationale

Playing video games is an integral part of adolescent culture, serving not only as entertainment but also as a platform for skill development and identity exploration [1]. For adolescents with behavioral or developmental needs, gaming provides structured environments to practice competencies that support personal

and social growth. Therapeutic games can enhance emotional regulation, problem-solving, and executive function, making them valuable for interventions addressing these challenges [2]. Health care professionals increasingly explore how these skills can be applied in therapeutic contexts, particularly through digital interventions that expand access to psychological support [3]. However, further research is needed to understand how in-game learning effectively transfers to real-world applications and how game design supports this process.

A growing body of research examines the role of serious games in mental health interventions. Systematic reviews and meta-analyses have shown their effectiveness in addressing anxiety [4], depression [5,6], emotional regulation [7], and broader psychological well-being [8,9]. Serious games targeting adolescents show particular promise for fostering engagement and skill development among populations who may resist traditional forms of therapy [10]. Yet, existing reviews primarily evaluate outcomes rather than examining how their design rationale enables learning and behavioral change. Recent work has begun to highlight the need for design-focused evaluation frameworks in serious gaming interventions [11], including the integration of psychological, motivational, and contextual design principles [12]. Understanding these mechanisms is critical for developing scalable, evidence-based interventions that integrate meaningfully into health care practice.

Health Care and Serious Gaming

Adolescents with behavioral or developmental challenges often engage deeply with video games, offering both opportunities for therapeutic skill development and challenges for implementation in health care contexts [13]. Many abilities cultivated through gameplay, such as executive functioning, emotional regulation, and social communication, align closely with health care objectives [14]. Serious games, distinct from entertainment games, are intentionally designed with educational or therapeutic intent to foster learning and behavioral change through structured feedback and interaction [15].

For adolescents experiencing anxiety, social stress, or executive dysfunction, game environments provide opportunities for cognitive rehearsal and emotional self-regulation [16]. The structured and controllable nature of gameplay can reduce stress and facilitate gradual exposure, particularly among individuals with autism spectrum disorder [17]. Integrating biofeedback features, such as heart rate monitoring, further enhances awareness of physiological states and supports self-management [3]. Multiplayer and cooperative formats extend the therapeutic potential by promoting communication and collaboration, aligning with broader educational and clinical goals [18].

From a motivational perspective, serious gaming, or the process of applying game-based interaction in structured interventions, benefits from self-determination theory, emphasizing autonomy, competence, and relatedness [19]. Well-designed games can support all 3 dimensions, fostering intrinsic motivation and engagement. Because players differ in what motivates them, adaptable and inclusive designs are essential [20]. Despite growing evidence for their value, the integration of serious games into health care remains limited. Participatory design, involving adolescents, caregivers, and clinicians, can enhance contextual relevance and therapeutic alignment [21]. In this sense, serious games can function as mediating tools that connect education, therapy, and everyday life, establishing conditions for what boundary-crossing theory [22] describes as transitions between distinct domains of practice.

Boundary Crossing and Learning Mechanisms

Serious games designed for therapeutic or educational purposes often simulate or abstract real-world challenges to foster learning

and behavioral change [23]. Understanding how these games enable the application of skills beyond gameplay remains complex [24,25]. Drawing on model-based reasoning, effective learning occurs when interactive representations balance abstraction with real-world fidelity, enabling players to engage in both conceptual exploration and situated practice [26,27]. Within this context, serious games can act as boundary objects or artifacts that mediate between distinct domains such as health care, education, and daily life [28].

Boundary crossing theory helps explain how serious games facilitate transitions between these domains. Akkerman and Bakker [22] identify 4 mechanisms of boundary crossing: identification, coordination, reflection, and transformation, which enable knowledge movement across contexts. In serious games, these mechanisms appear when players assume new roles, collaborate across disciplines, or reflect on in-game experiences to link them with real-world contexts [29]. Through these processes, players act as boundary crossers, integrating insights, perspectives, and strategies from simulated environments into their personal or professional realities.

Designing for boundary crossing requires balancing structure and flexibility. Overly standardized mechanics can constrain learning, while excessive abstraction risks disconnecting gameplay from therapeutic goals. Scaffolding methods, such as reflective dialogue, adaptive feedback, and structured debriefing, can strengthen the link between play and practice [30,31]. When aligned with health care objectives, serious games can thus serve as both learning tools and vehicles for systemic integration [32].

Transfer and Model of Reality

The ability to apply knowledge and skills from one context to another is central to the educational and therapeutic values of serious games [33,34]. In health care, successful transfer determines whether insights gained during gameplay lead to meaningful behavioral or cognitive changes in real-world settings [35,36]. Two complementary types of transfer are often distinguished: literal transfer, which involves the direct application of learned skills in similar contexts, and figural transfer, which requires abstraction and adaptation to new or unfamiliar situations [33]. While literal transfer supports practice through realism, figural transfer fosters generalizable learning and adaptability [34].

Boundary crossing and transfer are interdependent. Mechanisms such as reflection and coordination provide the cognitive scaffolding that allows experiences within a game to extend beyond its boundaries. Thus, serious games do not merely teach within a virtual system but prepare players to act beyond it.

The model of reality represented in a game further defines how transfer occurs. Each serious game embodies a representational model, causal, relational, procedural, or structural, that shapes how aspects of the real world are simplified or transformed during play. As research by Wenzler [37] notes, these models mediate the relationship between simulated and actual experiences, influencing both fidelity and interpretive depth. Complementing this, the Bogost [38] concept of procedural rhetoric highlights how the logic of rules and mechanics

communicates meaning, shaping how players understand complex systems. Together, these frameworks explain how serious games model reality in ways that make transfer possible and meaningful.

For adolescents, such design choices are crucial. Games provide safe environments for exploring identity, practicing regulation, and testing social behaviors [39,40]. By combining representational fidelity with reflective abstraction, serious games can translate in-game learning into durable real-world competencies [41].

Aim and Research Questions

This systematic review examines how serious games in health care incorporate the constructs of transfer, boundary crossing, and models of reality within their design rationale and implementation. Unlike prior reviews that focus on outcomes, this study analyzes the conceptual coherence of design approaches, identifying how theoretical frameworks inform game development and affect the connection between simulated and real-world practice.

The study aims to clarify how design choices facilitate learning transfer, enable boundary-crossing conditions across educational and clinical domains, and represent reality through varying degrees of abstraction and fidelity.

We proposed the following research questions for this study:

- RQ1: To what extent do publications on serious games in health care incorporate design elements related to learning transfer, boundary crossing, and models of reality?
- RQ2: How are these 3 constructs conceptualized and operationalized within the design and development of serious games for health care?
- RQ3: What conceptual or methodological gaps remain in current design rationales, and how might addressing them strengthen the translational potential of serious games in health care?

Methods

This systematic review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 [42] guidelines to ensure transparency and reproducibility (Checklist 1). The review was not registered beforehand. The study focused on identifying serious game artifacts explicitly used in adolescent health care relationships. These artifacts were defined as intentionally designed games aimed at supporting learning or behavioral change rather than entertainment or measurement tools.

Five databases (PubMed, Scopus, ERIC, PsycINFO, and EMBASE) were searched for potentially relevant studies. The databases were selected for their combined coverage of health, psychology, and social sciences. Search strategies were developed in consultation with an information specialist and refined through several trial searches to improve precision and comprehensiveness (see [Multimedia Appendix 1](#) for search keys). Searches included papers published from database inception until October 2023 and were conducted between April 3 and May 21, 2023.

The search strategy was built around 4 intersecting conceptual domains to ensure relevant publications: (1) the health care domain, including games addressing therapeutic, preventive, or educational objectives; (2) the serious game domain, which included artifacts designed with explicit learning intent; (3) the problem domain, referring to behavioral, emotional, or developmental challenges; and (4) the target audience domain, referring to adolescents in health-related contexts. Only studies located within the overlap of all 4 domains were included. This approach yielded 3997 records, which were reduced to 2296 after duplicate removal. The broader search and the choice in strategy to try to find these artifacts were intentional. As a framework for evaluating the design rationale of serious games does not yet exist, we expected any data to be implicit.

To systematically evaluate design rationale and learning mechanisms, a structured framework was applied as follows. Artifact identification, which identifies the serious game, its name, and targeted skills or competencies [43,44]; Objective, which identifies the intent of the artifact by describing the desired learning process; Boundary crossing, which assesses how the game functions as a boundary object, supporting learning mechanisms' identification, coordination, reflection, or transformation [29]; Transfer, which evaluates how learning extends from game to reality, distinguishing between literal and figural transfer [33]; and Model of reality, which categorizes how the game represents real-world systems through 4 types: causal, relational, procedural, and structural models.

To ensure consistency and reliability, the framework was pilot-tested on a subset of 5 randomly selected studies before full application. Both reviewers applied the criteria independently, after which definitions were refined through consensus.

For coding consistency, categories were defined as follows. Literal transfer refers to games that replicate real-world tasks or environments, enabling direct practice of target behaviors. Figural transfer refers to games using symbolic or metaphorical representations that require players to generalize learning beyond the game. Boundary-crossing mechanisms were coded according to Akkerman and Bakker's [22] framework: identification (recognizing new roles or perspectives), coordination (managing interactions across tasks or roles), reflection (linking in-game experiences to real-world contexts), and transformation (applying learning across contexts in ways that alter practice). Models of reality were categorized as causal (cause-effect logic), procedural (stepwise processes or rules), relational (social or interpersonal dynamics), or structural (conceptual or systemic relationships).

Results

Screening Process

Two reviewers independently screened the 2296 titles and abstracts using Rayyan (Qatar Computing Research Institute). Studies were excluded if they focused on simulation for data modeling, used games only as measurement tools (eg, eye-tracking or motor-tracking), and lacked peer review or methodological detail. Studies were included when they

addressed a health care–related challenge affecting adolescents’ social, emotional, cognitive, or developmental functioning and used a serious game as an active learning or therapeutic tool. Because many included populations do not share a single diagnostic label, both health care and the target audience were interpreted more broadly. Disagreements were resolved through a structured consensus process and by including a third reviewer.

Following this stage, 47 papers were reviewed in full, resulting in 33 included studies. An overview of the key characteristics of the 33 included studies is provided in [Table 1](#) (see [Table 2](#) for the coding table). The excluded papers comprised 9 conference abstracts without accompanying full papers, 2 unavailable manuscripts, 2 available only in French, and 2 publications about the same artifact and context. A PRISMA 2020 flow diagram ([Figure 1](#)) summarizes the study selection process from 3997 initial records to 33 included studies.

Table . Characteristics of the included studies (N=33).

Author and year	Title	Population (n; age range, y)	Study design	Health domain	Context
David et al (2022) [18]	Do improvements in therapeutic game-based skills transfer to real-life improvements in children's emotion-regulation abilities and mental health?	Children and adolescents (48; 10 - 16 y)	Feasibility/pilot study (secondary analysis of RCT ^a data)	Emotion regulation, resilience, and mental health	Digital game intervention in clinical/therapeutic context
Zhang et al (2018) [45]	Understanding performance and verbal-communication of children with ASD in a collaborative virtual environment	Children with ASD ^b and typically developing peers (28; NR ^c)	Feasibility/pilot study	Social communication in ASD	Collaborative virtual environment/lab
Bossavit and Parsons (2018) [46]	Outcomes for design and learning when teenagers with autism codesign a serious game: a pilot study	Teenagers with ASD (6; 11 - 15 y)	Feasibility/pilot study (participatory co-design)	Teamwork, social interaction, and learning design	School/co-design setting
Beach and Wendt (2015) [47]	Social interaction development through immersive virtual environments	Adolescents with high-functioning ASD (2; 15 - 18 y)	Qualitative pilot case study	Social interaction development	Immersive VR ^d /educational context
Wang and Xing (2022) [48]	Supporting youth with autism learning social competence: a comparison of game- and nongame-based activities in 3D virtual world	Adolescents with ASD (11; NR)	Comparative study	Social competence	3D virtual world/educational setting
Ke and Moon (2018) [49]	Virtual collaborative gaming as social skills training for high-functioning autistic children	Children with high-functioning ASD (8; 10 - 14 y)	Mixed method multiple case study	Social skills training	Virtual playground/online environment
Kee et al (2022) [50]	Virtual reality-based social skills training for children with autism spectrum disorder	Children with ASD (7; 10 - 14 y)	Single-case intervention study	Social communication skills	Desktop VR/clinical or educational context
Johnson et al (2022) [51]	Charisma™ virtual social training: a digital health platform and protocol	Children and adolescents with social difficulties (67; 9 - 17 y)	Feasibility/pilot intervention study	Social coaching and social cognition	Remote/hybrid digital health settings
Atherton and Cross (2021) [52]	The use of analog and digital games for autism interventions	Autistic individuals across reviewed studies (NR; NR)	Scoping review	Cognitive and social skills in ASD	Mixed analog/digital intervention contexts
Thomsen and Adjorlu (2021) [53]	A collaborative virtual reality supermarket training application to teach shopping skills to young individuals with autism spectrum disorder	Adolescents with ASD (8; NR)	Feasibility/pilot study	Daily living skills (shopping and money management)	Collaborative VR supermarket environment
Vallefuoco et al (2022) [54]	Design of a serious game for enhancing money use in teens with autism spectrum disorder	Target group: adolescents with ASD (NR; 13 - 19 y)	Design and development study	Money use and financial literacy	3D digital game, likely school/therapy linked

Author and year	Title	Population (n; age range, y)	Study design	Health domain	Context
Caria et al (2018) [55]	The design of web games for helping young high-functioning autistics in learning how to manage money	Adolescents and young adults with high-functioning ASD (6; NR)	Design and usability study	Practical money management	Web-based games in educational context
Ringland et al (2017) [56]	Making in Minecraft: a means of self-expression for youth with autism	Autistic children and youth in an online community (NR; NR)	Qualitative ethnographic study	Self-expression, creativity, and participation	Minecraft-based online/club settings
Kandalaf et al (2013) [57]	Virtual reality social cognition training for young adults with high-functioning autism	Young adults with high-functioning ASD (8; 18 - 26 y)	Intervention study (pre-post with follow-up)	Social cognition	VR lab/clinical research settings
Stichter et al (2014) [58]	iSocial: delivering the social competence intervention for adolescents (SCI-A) in a 3D virtual learning environment for youth with high functioning autism	Adolescents with high-functioning ASD (NR; 12 - 17 y)	Intervention study (pre-post)	Social competence	Virtual learning environment
Lahiri et al (2011) [59]	Design of a virtual reality based adaptive response technology for children with autism spectrum disorder	Adolescents with ASD (6; NR)	Design and usability study	Social communication/response to social cues	VR-based system in lab/clinical settings
Grynszpan et al (2007) [60]	Exploring the influence of task assignment and output modalities on computerized training for autism	Children with ASD and typically developing peers (NR; NR)	Experimental training study	Executive functions and communication	Computer-based training environment
García-Redondo et al (2019) [61]	Serious games and their effect improving attention in students with learning disabilities	Students with learning disabilities (NR; primary school age)	Intervention study	Attention and executive functioning	School-based game use
Kerns et al (2017) [62]	Attention and working memory training: a feasibility study in children with neurodevelopmental disorders	Children with neurodevelopmental disorders (NR; 6 - 13 y)	Feasibility/pilot study	Attention and working memory	Educational/clinical training context
Ghanouni et al (2020) [63]	Design elements during development of videogame programs for children with autism spectrum disorder: Stakeholders' viewpoints	Stakeholders incl. youth with ASD (26; youth 13 - 17 y)	Qualitative study	Design elements for ASD game interventions	Stakeholder workshops/design sessions
Dovis et al (2015) [64]	Improving executive functioning in children with ADHD: training multiple executive functions within the context of a computer game. a randomized double-blind placebo controlled trial	Children with ADHD ^e (89; 8 - 12 y)	RCT	Executive functioning	Home/clinical digital training

Author and year	Title	Population (n; age range, y)	Study design	Health domain	Context
Lahiri et al (2011) [65]	Design of a gaze-sensitive virtual social interactive system for children with autism.	Adolescents with ASD (6; NR)	Design and usability study	Social gaze and social communication	VR system in lab/clinical context
MacCormack and Freeman (2019) [66]	Part 2: the virtual environment social program for youths with autism spectrum disorder	Children with ASD (4; 11 - 13 y)	Feasibility/pilot intervention study	Social competence and play therapy	Clinic/community-based virtual program
Wang et al (2017) [67]	Fostering verbal and non-verbal social interactions in a 3D collaborative virtual learning environment: a case study of youth with autism spectrum disorders learning social competence in iSocial	Youth with ASD (11; NR)	Exploratory case study	Verbal and nonverbal social interaction	3D collaborative virtual learning environment
Lerman et al (2017) [68]	A clinic-based assessment for evaluating job-related social skills in adolescents and adults with autism	Adolescents and adults with ASD (8; 16 - 32 y)	Assessment study	Vocational social skills	Clinic-based simulated workplace
Elhaddadi et al (2021) [69]	Serious games to teach emotion recognition to children with autism spectrum disorders (ASD)	Children and adolescents with ASD (32; NR)	Intervention study (pre-post)	Emotion recognition	Computer-based/school or lab context
Amat et al (2021) [70]	Design of an interactive virtual reality system, InViRS, for joint attention practice in autistic children	Children with ASD and typically developing peers (18; NR)	Design and feasibility study	Joint attention practice	VR-based interactive system
Bossenbroek et al (2020) [71]	Efficacy of a virtual reality biofeedback game (DEEP) to reduce anxiety and disruptive classroom behavior: single-case study	Adolescents in special secondary education (8; mean=15, SD 1.83 y)	Single-case experimental design	Anxiety reduction and emotion regulation	VR biofeedback game in school/clinical context
Kahlon et al (2019) [72]	Virtual reality exposure therapy for adolescents with fear of public speaking: a non-randomized feasibility and pilot study	Adolescents with public speaking anxiety (27; 13 - 16 y)	Feasibility/pilot study	Social anxiety/fear of public speaking	VR exposure in clinical settings
Lahiri et al (2015) [73]	A physiologically informed virtual reality based social communication system for individuals with autism	Adolescents with ASD and typically developing peers (12; mean=15.9, SD 2.15 y)	Usability and proof-of-concept study	Social communication and engagement	VR-based social communication system
Fornasari et al (2013) [74]	Navigation and exploration of an urban virtual environment by children with autism spectrum disorder compared to children with typical development	Children with ASD and typically developing peers (32; 7 - 14 y)	Comparative experimental study	Urban navigation and spatial behavior	Virtual urban environment/lab

Author and year	Title	Population (n; age range, y)	Study design	Health domain	Context
Ringland (2019) [75]	“Autsome”: fostering an autistic identity in an online Minecraft community for youth with autism	Autistic children and youth in online community (NR; NR)	Qualitative ethnographic study	Autistic identity and social belonging	Online Minecraft community
Finke et al (2018) [16]	“To be quite honest, if it wasn’t for videogames I wouldn’t have a social life at all”: motivations of young adults with autism spectrum disorder for playing videogames as leisure	Young adults with ASD (10; 18 - 24 y)	Qualitative interview study	Leisure gaming, social life, and identity	Home/leisure gaming contexts

^aRCT: randomized controlled trial.

^bASD: autism spectrum disorder.

^cNR: not reported.

^dVR: virtual reality.

^eADHD: attention-deficit/hyperactivity disorder.

Figure 1. Selection process for review: PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram.

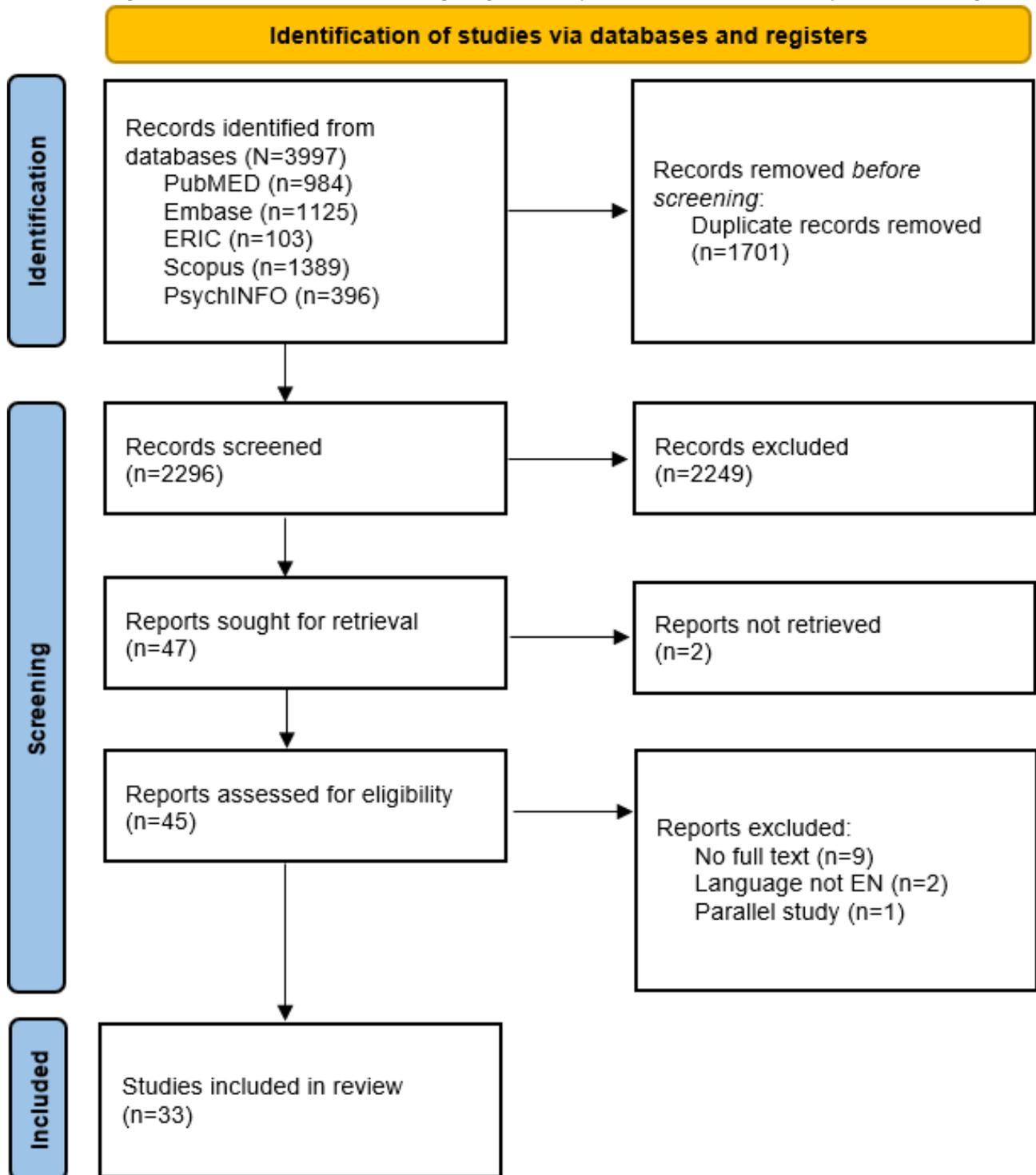


Table . Overview of design characteristics and analytical coding of included studies.

Author and year	Artifact	Objective	Learning mechanism	Transfer type	Model type
David et al (2022) [18]	REThink: emotional regulation and resilience for adolescents	Based on REBT ^a ; structured levels for emotion and cognition	Identification and reflection emphasized	Figural transfer. psychological fidelity	Relational, causal: REBT framework modeling emotional-cognitive feedback loops
Zhang et al (2018) [45]	CVE ^b : communication and collaboration skills for ASD ^c children	FSM ^d -based; real-time feedback for dynamic collaboration	Identification and coordination mechanisms present	Figural transfer. psychological fidelity via metaphorical social interaction scenarios	Causal, relational: FSM-based system structuring collaboration through cause-effect interactions
Bossavit and Parsons (2018) [46]	Geography-themed game: teamwork and geography knowledge for ASD adolescents	3T sandwich model; participatory design with tailored interaction modes	Identification and coordination in team-based play	Figural transfer: psychological fidelity with functional teamwork mechanics	Relational, procedural: team-based collaboration modeled through iterative design-feedback cycles
Beach and Wendt (2015) [47]	Immersive virtual environment: social interaction training for ASD students	Customized VR ^e scenarios based on social challenges	Identification and reflection mechanisms utilized	Figural transfer: psychological fidelity supporting skills transfer	Procedural, relational: guided VR social sequences replicating real-world interaction flows
Wang and Xing, (2022) [48]	3D virtual world game: collaborative social skill training for ASD youth	Scaffolded 3D gameplay; adaptive learning mechanics	Coordination through collaborative gameplay; reflection on interaction patterns	Figural transfer: fidelity via adaptive gameplay scenarios	Causal, relational: behavioral frameworks linking feedback to adaptive social competence
Ke and Moon, (2018) [49]	3D virtual playground: competition, role-play, and design for social interaction	Constructed in OpenSimulator; learner-adaptive	Identification via role-play and task adaptation	Figural transfer: physical fidelity to real-world context	Causal, procedural: role-play tasks reinforcing behavior through structured adaptation loops
Ke et al (2022) [50]	Desktop VR social skills program: social communication for ASD children	Scenario-based role-play; multimodal feedback	Coordination and reflection through tasks	Figural transfer from simulated to real	Procedural, relational: scenario-based communication tasks following stepwise learning processes
Johnson et al (2022) [51]	CHARISMA-VST ^f : low-immersion VR for pediatric social skill training	Strength-based coaching; remote-friendly	Reflection through peer role-play interactions	Figural transfer: psychological fidelity	Relational, causal: peer role-play and coaching systems modeling social cognition and feedback
Atherton and Cross (2021) [52]	Gamified interventions: digital and analog games for social and cognitive skills	Behavioral reinforcement; adaptive narratives	Reflection via joint engagement tasks	Figural transfer: blended fidelity with narrative integration	Causal, structural: cognitive-behavioral feedback integrated into narrative skill frameworks
Thomsen and Adjorlu (2021) [53]	VR supermarket training: shopping and money management for ASD adolescents	Co-designed with teachers; task-based learning	Coordination and transformation via tasks	Literal transfer: physical and psychological fidelity	Procedural, causal: shopping and money management simulated through task-based learning loops
Vallefuoco et al (2021) [54]	€UReka: 3D game for recognizing and handling money	Participatory design with multidisciplinary input; tailored challenges	Identification via realistic tasks; reflection through interactive learning	Literal transfer: psychological fidelity	Procedural, structural: iterative money-handling training structured within financial skill frameworks
Caria et al (2018) [55]	Web-based games for practical money management	Iterative co-design process; accessibility focus	Coordination via financial decision-making tasks	Literal transfer: psychological fidelity	Procedural, structural: decision-making tasks organized through educational usability structures

Author and year	Artifact	Objective	Learning mechanism	Transfer type	Model type
Ringland et al (2017) [56]	Minecraft-based activities for creativity and self-expression	Creative focus; adaptive gameplay for needs	Reflection through collaboration and shared space	Figural transfer: psychological fidelity via metaphorical engagement	Relational, structural: maker culture modeling social creativity within conceptual frameworks
Kandalajt et al (2013) [57]	VR-SCT ^g : virtual reality-based training for social cognition	Immersive VR; real-life scenario simulation	Reflection on cognition; transformation via practice	Figural transfer: psychological fidelity aligning with real skills	Relational, procedural: VR social cognition training reflecting interpersonal and sequential skill use
Stichter et al (2014) [58]	iSocial 3D VLE ^h : social competence training for adolescents	ABA ⁱ principles in collaborative virtual environments	Coordination and reflection via guided interaction	Figural transfer: psychological and physical fidelity for social skills	Causal, relational: ABA principles linking stimuli and responses within collaborative learning
Lahiri et al (2011) [59]	VR system for social communication tasks for ASD children	Real-time eye-gaze monitoring with adaptive tasks	Identification and reflection with social avatars	Figural transfer: psychological fidelity	Causal, procedural: adaptive gaze monitoring creating real-time feedback-based learning sequences
Grynszpan et al (2007) [60]	Multimodal games to enhance executive functions in ASD	Multimodal testing; task differentiation by domain	Coordination of spatial and pragmatic skills	Literal transfer: psychological fidelity tailored to user training	Structural, procedural: executive functions structured conceptually and trained through procedural tasks
García-Redondo et al (2019) [61]	Games based on the Gardner multiple intelligences for attention training	Narrative and cognitive reinforcement strategies	Reflection on cognitive strengths and interactive tasks	Figural transfer: psychological fidelity through narratives	Structural, causal: multiple intelligences mapped to causal reinforcement mechanisms for attention
Kerns et al (2017) [62]	Caribbean Quest: game-based attention and memory training	Metacognitive strategies integrated with gameplay	Reflection and transformation of cognitive skills	Figural and literal transfer via adaptive training: psychological fidelity	Structural, procedural: cognitive remediation organizing mental functions into adaptive training sequences
Ghanouni et al (2020) [63]	Stakeholder-informed videogame programs for ASD	Participatory design with stakeholder feedback	Reflection via user-centered design processes	Figural transfer for adaptable social skills: psychological fidelity	Relational, structural: stakeholder-based co-design modeling social systems within structured frameworks
Dovis et al (2015) [64]	Braingame Brian: game targeting multiple executive functions in ADHD ^j	Gamified EF ^k training integrating WM ^l , inhibition, and flexibility	Coordination through tasks; reflection on performance	Figural transfer: psychological fidelity	Structural, causal: executive functions integrated into structured cognitive frameworks with feedback
Lahiri et al (2011) [65]	VIGART ^m : VR-based system for real-time gaze interaction	Dynamic gaze-based feedback for social communication	Reflection and identification with gaze data	Literal transfer: physical fidelity	Causal, procedural: real-time gaze-based feedback linking attention to behavioral reinforcement
MacCormack and Freeman (2019) [66]	Minecraft-based structured/free play for social competence in ASD	Peer mediation, video modeling, and structured play methods	Coordination through mediated interactions	Figural transfer through gameplay scenarios	Relational, procedural: peer mediation and structured play modeling social competence through sequences
Wang et al (2017) [67]	iSocial: 3D CVLE ⁿ for ASD social competence	Narrative-embedded, goal-oriented, peer-supported tasks	Reflection in collaborative 3D environments	Figural transfer for social behaviors: psychological fidelity	Causal, relational. behavioral frameworks reinforcing social interaction in 3D virtual collaboration

Author and year	Artifact	Objective	Learning mechanism	Transfer type	Model type
Lerman et al (2017) [68]	Protocol for job-related social skill assessment in ASD	Simulated workplace tasks for skill assessment	Identification in work-simulated environments	Literal transfer of job-related skills: physical and psychological fidelity	Procedural, causal: simulated workplace tasks reflecting vocational training processes
Elhaddadi et al (2021) [69]	Multisensory game for teaching emotion recognition in ASD	Adaptation to cultural context and emotion-based tasks	Reflection through emotion recognition tasks	Literal transfer of emotion skills: psychological fidelity	Causal, structural: emotion recognition modeled through behavioral and cultural task frameworks
Amat et al (2021) [70]	VR-based game for gaze sharing and following in ASD	Real-time gaze tracking with adaptive feedback	Coordination through gaze-based interaction	Literal transfer via gaze coordination: physical fidelity	Causal, procedural: gaze-tracking feedback driving joint attention in sequential learning loops
Bossenbroek et al (2020) [71]	VR biofeedback game for anxiety and disruptive behavior reduction	Diaphragmatic breathing integrated with biofeedback	Reflection through self-awareness in VR	Figural transfer: psychological fidelity for emotion regulation	Causal, procedural: biofeedback linking physiological regulation to emotional response mechanisms
Kahlon et al (2019) [72]	VR exposure therapy simulating public speaking scenarios	Age-appropriate scenarios for exposure therapy	Reflection via structured exposure tasks	Figural transfer in social performance: psychological fidelity	Procedural, causal: exposure therapy modeled through sequential desensitization and behavioral outcomes
Lahiri et al (2015) [73]	VR-based system using physiological engagement for social communication	Adaptive response systems using engagement metrics	Reflection and coordination with biofeedback	Figural transfer tied to engagement cues: psychological fidelity	Causal, procedural: physiological feedback loops structuring adaptive engagement tasks
Fornasari et al (2013) [74]	Virtual environment simulation for urban navigation	Simplified navigation tasks for ASD; comparative evaluation	Reflection through exploration tasks in VE	Literal transfer of urban navigation skills: physical fidelity	Structural, procedural: simplified urban environments representing spatial systems for navigation practice
Ringland (2019) [75]	Minecraft-based online community for fostering autistic identity	Inclusive community promoting identity and safety	Coordination via shared gaming and identity	Figural transfer, psychological fidelity in building resilience	Relational, structural: online community fostering identity within structured digital interaction systems
Finke et al (2018) [16]	Videogaming as leisure activity for social and identity development	Leisure gaming to enhance social and communicative skills	Reflection on social experiences in gaming	Figural transfer: psychological fidelity emphasizing engagement	Relational, structural: gaming as social-identity context framed by interactional structures

^aREBT: rational emotive behavior therapy.

^bCVE: collaborative virtual environment.

^cASD: autism spectrum disorder.

^dFSM: finite state machine.

^eVR: virtual reality.

^fCHARISMA-VST: charisma virtual social training.

^gVR-SCT: virtual reality social cognition training.

^h3D-VLE: 3D virtual learning environment.

ⁱABA: applied behavior analysis.

^jADHD: attention-deficit/hyperactivity disorder.

^kEF: executive functions.

^lWM: working memory.

^mVIGART: virtual interactive gaze-based adaptive response technology.

ⁿCVLE: collaborative virtual learning environment.

Inter-rater reliability for the initial round of coding was $\kappa=0.55$, which reflects moderate agreement ($\kappa=0.41 - 0.60$) [76,77]. Moderate agreement is acceptable in qualitative content analyses that require inferential judgment, particularly when coding implicit or theory-driven categories, such as transfer types, boundary-crossing mechanisms, and model typologies. These constructs were often not explicitly described, making some variability in coder interpretation expected.

Transfer

Many games prioritized figural transfer ($n=24$), using metaphorical or symbolic representations to promote generalizable cognitive and emotional learning. Only 10 studies utilized literal transfer, replicating real-life tasks or environments. Examples include virtual reality supermarket training [53], which simulates grocery shopping to teach money management skills; Braingame Brian [64], which trains executive functions through structured gameplay; figural approaches, such as DEEP [71], which teaches emotional regulation through metaphorical underwater immersion; and iSocial [58], which develops social competence using virtual role-play, illustrating a focus on reflective learning rather than direct behavioral replication.

Despite these efforts, few studies incorporated explicit scaffolding or debriefing mechanisms to reinforce transfer, and longitudinal evaluations were rare. While short-term improvements in executive function and social skills were frequently measured, few studies investigated whether these abilities persisted or generalized beyond the intervention [72]. The overall dominance of figural transfer suggests that most games aim to develop reflective and conceptual understanding rather than direct behavioral transformation.

Boundary Crossing

The distribution of learning mechanisms was uneven: reflection appeared in 22 studies, coordination in 13, identification in 9, and transformation in 3. This distribution shows that most games promote self-awareness and perspective taking but rarely enable systemic changes or real-world applications. For instance, Minecraft-based communities [78] provided safe spaces for adolescents to explore identity and social belonging, while DEEP [71] encouraged users to monitor and regulate emotions. In contrast, CHARISMA-virtual social training [51] exemplified stronger coordination and transformation, linking coaching-based gameplay with clinical follow-up and remote engagement. While participatory design processes frequently involved caregivers, therapists, or educators, few studies explicitly articulated a boundary-crossing framework. Consequently, most serious games functioned as standalone interventions, rather than integrated components within health care systems.

Models of Reality

The conceptualization of models of reality varied significantly across the reviewed studies. Based on an analytical typology developed for this review, 4 types were identified: causal ($n=12$), procedural ($n=7$), relational ($n=9$), and structural ($n=5$). Causal and procedural models were most frequent overall (each appearing 19 times across primary and secondary codings),

reflecting a focus on feedback-driven learning and stepwise behavioral training. Relational and structural models typically appeared as secondary logics, emphasizing social interaction and conceptual scaffolding. Common examples include virtual reality exposure therapy [72], which applied a causal-procedural model to replicate public speaking scenarios, and iSocial [58], which integrated relational and structural logics to foster social competence through collaborative virtual environments. In contrast, some games oversimplified real-world complexity, such as urban navigation simulators [74], which lacked the unpredictability representative of real-life settings.

Alignment Between Model Type and Transfer

A cross-analysis of model types and transfer forms revealed strong internal alignment between theoretical constructs and design intent. Causal-figural ($n=9$) and relational-figural ($n=9$) pairings dominated, representing enabling configurations that support reflective and metaphorical learning. Conversely, procedural-literal ($n=4$) and structural-literal ($n=2$) combinations were less common but corresponded to direct skill training and conceptual knowledge transfer.

These results indicate that most games intentionally (or implicitly) align their learning mechanisms with their representational logic: figural models enable reflective and identity-building experiences, while literal designs facilitate direct behavioral rehearsal. The absence of relational-literal pairings further underscores that social learning is almost exclusively represented symbolically in adolescent mental health games.

Stakeholder Engagement

Stakeholder engagement was a consistent strength across many studies, ensuring relevance and usability through participatory design. For example, €UREKA [54], a 3D game for financial literacy, involved multidisciplinary teams in tailoring design challenges, while virtual interactive gaze-based adaptive response technology [65] used real-time adaptive feedback to improve social communication. However, broader collaboration with policymakers and institutional stakeholders was rare, limiting scalability and integration into health care systems.

From a design perspective, these 2 forms of engagement can be understood as “design in the small” and “design in the large.” The former refers to participation in the artifact-level design (mechanics, interface, and content), whereas the latter concerns the contextual integration of the artifact within health care or educational systems. While most studies effectively address design in the small, few extend their focus to design in the large, leaving questions of implementation, sustainability, and policy alignment unresolved. Future research should therefore emphasize multilevel stakeholder engagement, involving decision-makers and regulatory partners to facilitate scaling from pilot projects to institutionalized practice [53].

Summary

This review highlights distinct but interrelated design tendencies: (1) a preference for figural transfer and reflective boundary mechanisms, (2) a predominance of causal-procedural and relational-structural model clusters, and (3) a strong alignment

between representational logic and learning mechanisms. These findings suggest that serious games for adolescent mental health are predominantly designed as reflective learning tools, fostering emotional and cognitive awareness rather than direct behavioral transformation. However, their impact depends less on which configuration is used and more on how deliberately these design choices are made and articulated. When the mechanisms of transfer, boundary crossing, and the model of reality are explicitly aligned within the design process, they can actively enable boundary crossing conditions.

From a design perspective, this alignment operates across 2 levels: design in the small, focusing on the internal coherence of the artifact (mechanics, narrative, fidelity, and user experience), and design in the large, emphasizing how the artifact interacts with its broader institutional, educational, or clinical context. By making these theoretical dimensions explicit during both levels of design, developers and researchers can create serious games that not only model health care processes but also facilitate real-world transfer and systemic learning.

Discussion

Principal Findings

This review highlights the transformative potential of serious games in health care while identifying key areas for improvement. A major challenge is the need for a coherent design rationale that explicitly integrates transfer, boundary crossing, and models of reality into the process of game development. Without such a foundation, serious games risk remaining isolated interventions rather than scalable tools for systemic change. This risk is reflected in the reviewed studies by the limited occurrence of transformation mechanisms (3/33 studies), the predominance of figural transfer (24/33 studies), and the concentration of stakeholder engagement at the artifact level rather than at the level of institutional implementation. Recent meta-analyses confirm that although serious games show positive effects on engagement and short-term outcomes, long-term validation and theoretical consistency remain limited [79].

Learning Transfer

Among the reviewed studies, 24 employed figural transfer and 10 literal transfer, indicating that most serious games rely on symbolic or metaphorical learning rather than direct simulation. Figural transfer supports generalizable emotional and cognitive skills, whereas literal transfer offers task-specific realism. This emphasis on figural approaches suggests a preference for flexibility over precision but also highlights the need for scaffolding and debriefing mechanisms to reinforce learning. Drawing on educational and simulation-based research [80,81], future designs should integrate structured reflection and adaptive feedback to strengthen transfer beyond gameplay. This finding mirrors recent discussions in digital learning and simulation literature, emphasizing the importance of scaffolding mechanisms for durable learning outcomes [82,83].

Patterns in the data also reveal an alignment between transfer type and model typology: games employing causal or relational models tend to favor figural transfer. This finding suggests that

representational logic and learning design interact closely, underscoring the importance of explicit design decisions to ensure coherence between game mechanics, learning goals, and therapeutic intent. This alignment is further illustrated by the absence of relational–literal pairings and by the concentration of figural transfer within causal and relational primary models (see [Multimedia Appendix 2](#)).

Boundary Crossing

Boundary crossing was frequently observed but unevenly implemented. Reflection appeared in 22 studies, whereas transformation, representing deep cross-context learning, occurred in only 3 studies. Most games promoted perspective taking or self-awareness but stopped short of enabling collaborative learning or system-wide integration. This pattern corresponds with the low frequency of transformation as a boundary-crossing mechanism and suggests that learning is rarely designed to extend beyond the immediate intervention context. According to Akkerman and Bakker [29], transformation emerges when learning processes bridge distinct social or professional domains.

Explicitly defining intended boundary-crossing mechanisms within the design process could help serious games act as boundary objects, facilitating communication among clinicians, educators, and developers. This translational function positions games not only as therapeutic tools but also as mediators that enable shared understanding across sectors of health care.

Models of Reality

Analysis of model types revealed a dominance of causal and procedural frameworks, with relational and structural models often appearing as secondary. These choices reflect a focus on mechanistic and process-oriented representations, supporting procedural learning that may limit the representation of social and ethical complexity when relational or structural models are not explicitly incorporated. This aligns with Bogost's [38] notion of procedural rhetoric, where games express meaning through rules and system dynamics rather than narrative content.

While causal and procedural models effectively convey behavioral and process learning, they risk neglecting the interpersonal dimensions of health care. By contrast, relational and structural models, exemplified in iSocial [58] and CHARISMA-virtual social training [51], capture the social and contextual aspects of health practice. Integrating these model types can produce games that are both systematically rigorous and socially responsive, mirroring the real-world complexity of health care environments.

Stakeholder Engagement and Design Scales

Stakeholder engagement was a consistent strength but remained concentrated at the artifact level. Across the reviewed studies, participatory involvement most frequently informed game mechanics, content, or usability, while engagement with organizational, policy, or stakeholders was rarely reported. Most studies prioritized design in the small through participatory methods involving therapists, educators, or caregivers, but few extended collaboration to policy, administration, or implementation levels. Expanding engagement to design in the

large is essential for ensuring that serious games transition from prototypes to sustainable health care tools [84].

Synthesis and Implications

Across studies, serious games for adolescent health care predominantly support reflective and representational learning, emphasizing awareness and regulation rather than long-term behavioral transformation. However, the findings indicate that effectiveness depends less on any single design dimension than on their explicit alignment.

When transfer mechanisms, boundary-crossing strategies, and model logics are purposefully coordinated, serious games can actively enable boundary-crossing conditions. This allows learning to flow between digital environments, clinical practice, and educational systems. Such alignment strengthens both the internal validity of the game as a learning artifact and its external applicability as a health care intervention.

Based on the observed design patterns across the reviewed studies, these implications can be understood as a conceptual synthesis rather than a prescriptive model. In practice, this means designing serious games as multilevel systems that bridge theory and implementation, ensuring fidelity, feedback, and learner engagement, embedding cross-professional collaboration and evaluation, and aligning design with institutional and policy frameworks. With this approach, future developers and researchers can create serious games that not only demonstrate short-term efficacy but also achieve sustained, scalable, and clinically relevant impact in health care delivery.

Strengths and Limitations

This review advances serious game research by applying a structured multicriteria framework to examine design rationale, learning transfer, and real-world applicability. By analyzing 33 studies through the lenses of transfer, boundary crossing, and models of reality, this review provides one of the first quantified overviews of how theoretical constructs are applied in serious game design. The inclusion of typological and frequency analyses strengthens methodological transparency and highlights recurring design patterns.

However, several limitations should be acknowledged. The review is limited to published studies, introducing publication bias toward positive outcomes. While coding followed systematic criteria, some interpretive judgments were necessary to infer implicit design rationales. In addition, full PRISMA 2020 compliance and broader database coverage would enhance reproducibility. Finally, the scarcity of longitudinal studies constrains the understanding of sustained behavioral or cognitive impact. Despite these constraints, this review provides a solid foundation for future work. It shows that making design choices around transfer, boundary mechanisms, and models of reality explicit and aligned can help create more scalable, theoretically grounded health care games.

Future Research

Future research should pursue longitudinal studies to evaluate how learning from serious games endures in real-world health care settings. Establishing standardized methods for assessing

transfer, boundary crossing, and models of reality will improve comparability and design validation. Emphasis should be placed on explicit, theory-informed design alignment and interdisciplinary collaboration among developers, clinicians, and researchers. Broader stakeholder engagement, including policymakers, is vital for scaling interventions and integrating serious games into health care systems. These efforts can transform serious games from isolated innovations into sustainable, evidence-based tools that enhance learning, practice, and health care delivery.

Conclusions

Building on these insights, this review demonstrates the transformative potential of serious games in health care, particularly for individuals with developmental and behavioral needs [1,43]. By fostering skill acquisition, collaboration, and meaningful transfer between digital and real-world contexts, serious games offer promising tools for advancing emotional regulation, cognitive development, and social interaction. However, their scalability and long-term impact remain limited by inconsistently articulated design rationales, which hinder replicability and adaptation across health care settings [29,35].

By systematically examining learning transfer, boundary crossing, and models of reality, this review provides an evidence-based framework for understanding how design decisions influence learning outcomes and real-world applicability. The findings highlight that explicit alignment among these constructs supported by participatory co-design and interdisciplinary collaboration is essential for achieving sustainable health care innovation. Broadening stakeholder engagement to include clinicians, caregivers, policymakers, and end users can further strengthen contextual integration and scalability [64].

Future research should focus on developing standardized, theory-informed frameworks [33], implementing longitudinal and mixed methods evaluations [44], and establishing implementation guidelines that connect design in the small (artifact-level) with design in the large (systemic integration). Through such structured, evidence-based approaches, serious games can evolve from isolated prototypes into validated, scalable instruments that enhance clinical decision-making, therapeutic engagement, and professional development in modern health care. Future work should also focus on validating these design principles across larger datasets and contexts. Comparative and meta-analytic approaches could test how alignment among transfer, boundary crossing, and model typology predicts real-world outcomes, advancing serious game design as an evidence-based methodology in health care.

Key Takeaway

This review shows that serious games in adolescent health care predominantly support reflective and metaphorical learning, while rarely specifying how learning is expected to transfer beyond the game context. Making transfer mechanisms, boundary-crossing strategies, and underlying models of reality explicit during design could improve comparability, validation, and scalability of serious games in health care practice.

Acknowledgments

The authors would like to thank Olga van Dijk from the Hanzehogeschool Groningen (NL), who helped develop the search strategy and provided support during the review process. The authors used generative artificial intelligence tools (eg, Claude and Rayyan) to support language editing and clarity. All content was critically reviewed and revised by the authors.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search key appendix as referenced in the Methods.

[[DOCX File, 15 KB - games_v14i1e77173_app1.docx](#)]

Multimedia Appendix 2

Coding count and cross table.

[[DOCX File, 16 KB - games_v14i1e77173_app2.docx](#)]

Checklist 1

PRISMA 2020 guidelines.

[[DOCX File, 280 KB - games_v14i1e77173_app3.docx](#)]

References

1. Granic I, Lobel A, Engels R. The benefits of playing video games. *Am Psychol* 2014 Jan;69(1):66-78. [doi: [10.1037/a0034857](https://doi.org/10.1037/a0034857)] [Medline: [24295515](https://pubmed.ncbi.nlm.nih.gov/24295515/)]
2. Reynard S, Dias J, Mitic M, Schrank B, Woodcock KA. Digital interventions for emotion regulation in children and early adolescents: systematic review and meta-analysis. *JMIR Serious Games* 2022 Aug 19;10(3):e31456. [doi: [10.2196/31456](https://doi.org/10.2196/31456)] [Medline: [35984681](https://pubmed.ncbi.nlm.nih.gov/35984681/)]
3. Brezinka V. Computer games supporting cognitive behaviour therapy in children. *Clin Child Psychol Psychiatry* 2014 Jan;19(1):100-110. [doi: [10.1177/1359104512468288](https://doi.org/10.1177/1359104512468288)] [Medline: [23258925](https://pubmed.ncbi.nlm.nih.gov/23258925/)]
4. 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. [doi: [10.2196/games.9530](https://doi.org/10.2196/games.9530)] [Medline: [29490893](https://pubmed.ncbi.nlm.nih.gov/29490893/)]
5. Dias LPS, Barbosa JLV, Vianna HD. Gamification and serious games in depression care: a systematic mapping study. *Telemat Inform* 2018 Apr;35(1):213-224. [doi: [10.1016/j.tele.2017.11.002](https://doi.org/10.1016/j.tele.2017.11.002)]
6. 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 Dec 16;9(4):e30482. [doi: [10.2196/30482](https://doi.org/10.2196/30482)] [Medline: [34927589](https://pubmed.ncbi.nlm.nih.gov/34927589/)]
7. 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](https://doi.org/10.1089/g4h.2017.0108)] [Medline: [29424555](https://pubmed.ncbi.nlm.nih.gov/29424555/)]
8. David OA, Costescu C, Cardos R, Mogoșe C. How effective are serious games for promoting mental health and health behavioral change in children and adolescents? a systematic review and meta-analysis. *Child Youth Care Forum* 2020 Dec;49(6):817-838. [doi: [10.1007/s10566-020-09566-1](https://doi.org/10.1007/s10566-020-09566-1)]
9. Zayeni D, Raynaud JP, Revet A. Therapeutic and preventive use of video games in child and adolescent psychiatry: a systematic review. *Front Psychiatry* 2020;11:36. [doi: [10.3389/fpsy.2020.00036](https://doi.org/10.3389/fpsy.2020.00036)] [Medline: [32116851](https://pubmed.ncbi.nlm.nih.gov/32116851/)]
10. 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](https://doi.org/10.1016/j.psychres.2020.113585)] [Medline: [33303223](https://pubmed.ncbi.nlm.nih.gov/33303223/)]
11. De Jaegere E, van Heeringen K, Emmery P, Mommerency G, Portzky G. Effects of a serious game for adolescent mental health on cognitive vulnerability: pilot usability study. *JMIR Serious Games* 2024 May 9;12(1):e47513. [doi: [10.2196/47513](https://doi.org/10.2196/47513)] [Medline: [38725223](https://pubmed.ncbi.nlm.nih.gov/38725223/)]
12. Zeiler M, Vögl S, Prinz U, et al. Game design, effectiveness, and implementation of serious games promoting aspects of mental health literacy among children and adolescents: systematic review. *JMIR Ment Health* 2025 May 5;12:e67418. [doi: [10.2196/67418](https://doi.org/10.2196/67418)] [Medline: [40324175](https://pubmed.ncbi.nlm.nih.gov/40324175/)]
13. Bossavit B, Parsons S. "This is how I want to learn": high functioning autistic teens co-designing a serious game. Presented at: CHI'16; May 7-12, 2016. [doi: [10.1145/2858036.2858322](https://doi.org/10.1145/2858036.2858322)]

14. Vacca RA, Augello A, Gallo L, et al. Serious Games in the new era of digital-health interventions: a narrative review of their therapeutic applications to manage neurobehavior in neurodevelopmental disorders. *Neurosci Biobehav Rev* 2023 Jun;149:105156. [doi: [10.1016/j.neubiorev.2023.105156](https://doi.org/10.1016/j.neubiorev.2023.105156)] [Medline: [37019246](https://pubmed.ncbi.nlm.nih.gov/37019246/)]
15. Grynspan O, Weiss PLT, Perez-Diaz F, Gal E. Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. *Autism* 2014 May;18(4):346-361. [doi: [10.1177/1362361313476767](https://doi.org/10.1177/1362361313476767)] [Medline: [24092843](https://pubmed.ncbi.nlm.nih.gov/24092843/)]
16. Finke EH, Hickerson BD, Kremkow JMD. "To be quite honest, if it wasn't for videogames i wouldn't have a social life at all": motivations of young adults with autism spectrum disorder for playing videogames as leisure. *Am J Speech Lang Pathol* 2018 May 3;27(2):672-689. [doi: [10.1044/2017_AJSLP-17-0073](https://doi.org/10.1044/2017_AJSLP-17-0073)] [Medline: [29610898](https://pubmed.ncbi.nlm.nih.gov/29610898/)]
17. Kuo MH, Orsmond GI, Coster WJ, Cohn ES. Media use among adolescents with autism spectrum disorder. *Autism* 2014 Nov;18(8):914-923. [doi: [10.1177/1362361313497832](https://doi.org/10.1177/1362361313497832)] [Medline: [24142797](https://pubmed.ncbi.nlm.nih.gov/24142797/)]
18. David OA, Magurean S, Tomoiagă C. Do improvements in therapeutic game-based skills transfer to real life improvements in children's emotion-regulation abilities and mental health? a pilot study that offers preliminary validity of the RETHink in-game performance scoring. *Front Psychiatry* 2022;13:828481. [doi: [10.3389/fpsy.2022.828481](https://doi.org/10.3389/fpsy.2022.828481)] [Medline: [35386525](https://pubmed.ncbi.nlm.nih.gov/35386525/)]
19. Deci EL, Ryan RM. The "What" and "Why" of goal pursuits: human needs and the self-determination of behavior. *Psychol Inq* 2000 Oct;11(4):227-268. [doi: [10.1207/S15327965PLI1104_01](https://doi.org/10.1207/S15327965PLI1104_01)]
20. Yee N. Motivations for play in online games. *Cyberpsychol Behav* 2006 Dec;9(6):772-775. [doi: [10.1089/cpb.2006.9.772](https://doi.org/10.1089/cpb.2006.9.772)] [Medline: [17201605](https://pubmed.ncbi.nlm.nih.gov/17201605/)]
21. Gómez-Cambronero Á, Mann AL, Mira A, Doherty G, Casteleyn S. Smartphone-based serious games for mental health: a scoping review. *Multimed Tools Appl* 2024;83(36):84047-84094. [doi: [10.1007/s11042-024-18971-w](https://doi.org/10.1007/s11042-024-18971-w)] [Medline: [39553422](https://pubmed.ncbi.nlm.nih.gov/39553422/)]
22. Akkerman SF, Bakker A. Boundary crossing and boundary objects. *Rev Educ Res* 2011 Jun;81(2):132-169. [doi: [10.3102/0034654311404435](https://doi.org/10.3102/0034654311404435)]
23. Geurts JLA, Geurts RD, Vermeulen PAM. Policy games for strategic management. *Long Range Plann* 2007 Dec;40(6):535-558. [doi: [10.1016/j.lrp.2007.07.004](https://doi.org/10.1016/j.lrp.2007.07.004)]
24. Cartwright N. *Nature, the Artful Modeler: Lectures on Laws, Science, How Nature Arranges the World and How We Can Arrange It Better*: Open Court Publishing Company; 2019.
25. van Fraassen BC. *The Scientific Image*: Oxford University Press; 1980. [doi: [10.1093/0198244274.001.0001](https://doi.org/10.1093/0198244274.001.0001)]
26. Magnani L, Casadio C. *Model-Based Reasoning in Science and Technology: Logical, Epistemological, and Cognitive Issues*: Springer International Publishing; 2016. [doi: [10.1007/978-3-319-38983-7](https://doi.org/10.1007/978-3-319-38983-7)]
27. Swoyer C. Structural representation and surrogative reasoning. *Synthese* 1991 Jun;87(3):449-508. [doi: [10.1007/BF00499820](https://doi.org/10.1007/BF00499820)]
28. Star SL, Griesemer JR. Institutional ecology, `translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Soc Stud Sci* 1989 Aug;19(3):387-420. [doi: [10.1177/030631289019003001](https://doi.org/10.1177/030631289019003001)]
29. Wenger E. *Communities of Practice: Learning, Meaning, and Identity*: Cambridge University Press; 1998.
30. Engeström Y. Expansive learning at work: toward an activity theoretical reconceptualization. *J Educ Work* 2001 Feb;14(1):133-156. [doi: [10.1080/13639080020028747](https://doi.org/10.1080/13639080020028747)]
31. Akkerman SF, Meijer PC. A dialogical approach to conceptualizing teacher identity. *Teach Teach Educ* 2011 Feb;27(2):308-319. [doi: [10.1016/j.tate.2010.08.013](https://doi.org/10.1016/j.tate.2010.08.013)]
32. Spahl W, Motta V, Woodcock K, Rubeis G. Gamified digital mental health interventions for young people: scoping review of ethical aspects during development and implementation. *JMIR Serious Games* 2024 Nov 28;12(1):e64488. [doi: [10.2196/64488](https://doi.org/10.2196/64488)] [Medline: [39607995](https://pubmed.ncbi.nlm.nih.gov/39607995/)]
33. Perkins D, Salomon G. Transfer of learning. In: Husén T, Postlethwaite TN, editors. *International Encyclopedia of Education*, 2nd edition: Pergamon Press; 1992:425-441 URL: <https://jaymctighe.com/wp-content/uploads/2011/04/Transfer-of-Learning-Perkins-and-Salomon.pdf> [accessed 2025-09-19]
34. Royer JM. Theories of the transfer of learning. *Educ Psychol* 1979 Jan;14(1):53-69. [doi: [10.1080/00461527909529207](https://doi.org/10.1080/00461527909529207)]
35. Gee JP. *Why Video Games Are Good for Your Soul: Pleasure and Learning*: Common Ground Publishing; 2005.
36. Green CS, Bavelier D. Learning, attentional control, and action video games. *Curr Biol* 2012 Mar 20;22(6):R197-R206. [doi: [10.1016/j.cub.2012.02.012](https://doi.org/10.1016/j.cub.2012.02.012)] [Medline: [22440805](https://pubmed.ncbi.nlm.nih.gov/22440805/)]
37. Wenzler I. Models of reality: transfer point between perceived reality and simulated reality. In: Kriz WC, Kikkawa T, Wijse-Van Heeswijk M, Lukosch H, editors. *Legacy and Future Impact of Gaming Simulation Pioneers*: Springer Nature; 2025:135-150. [doi: [10.1007/978-981-96-4325-7_9](https://doi.org/10.1007/978-981-96-4325-7_9)]
38. Bogost I. *Persuasive Games: The Expressive Power of Videogames*: The MIT Press; 2007.
39. Turkle S. *Life on the Screen: Identity in the Age of the Internet*: Simon & Schuster Trade; 1995.
40. Bopp JA, Mekler ED, Opwis K. Negative emotion, positive experience?: emotionally moving moments in digital games. Presented at: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; May 7-12, 2016. [doi: [10.1145/2858036.2858227](https://doi.org/10.1145/2858036.2858227)]
41. Engelbrecht H, Van der Laan N, Van Enschoot R, et al. The effect of procedural rhetoric and narrative content in a narrative game-based fear appeal. *Int J Serious Games* 2025 May 16;12(2):233-255. [doi: [10.17083/ijsg.v12i2.886](https://doi.org/10.17083/ijsg.v12i2.886)]
42. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71. [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]

43. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. *Pediatrics* 2008 Aug;122(2):e305-e317. [doi: [10.1542/peds.2007-3134](https://doi.org/10.1542/peds.2007-3134)] [Medline: [18676516](https://pubmed.ncbi.nlm.nih.gov/18676516/)]
44. Sweller J. Cognitive load theory, learning difficulty, and instructional design. *Learn Instr* 1994 Jan;4(4):295-312. [doi: [10.1016/0959-4752\(94\)90003-5](https://doi.org/10.1016/0959-4752(94)90003-5)]
45. Zhang L, Warren Z, Swanson A, Weitlauf A, Sarkar N. Understanding performance and verbal-communication of children with ASD in a collaborative virtual environment. *J Autism Dev Disord* 2018 Aug;48(8):2779-2789. [doi: [10.1007/s10803-018-3544-7](https://doi.org/10.1007/s10803-018-3544-7)] [Medline: [29560561](https://pubmed.ncbi.nlm.nih.gov/29560561/)]
46. Bossavit B, Parsons S. Outcomes for design and learning when teenagers with autism codesign a serious game: a pilot study. *J Comput Assist Learn* 2018 Jun;34(3):293-305. [doi: [10.1111/jcal.12242](https://doi.org/10.1111/jcal.12242)]
47. Beach J, Wendt J. Social interaction development through immersive virtual environments. Presented at: International Conferences on Educational Technologies 2014; Dec 10-12, 2014 URL: <https://files.eric.ed.gov/fulltext/ED557325.pdf> [accessed 2025-09-19]
48. Wang X, Xing W. Supporting youth with autism learning social competence: a comparison of game-and nongame-based activities in 3d virtual world. *J Educ Comput Res* 2022 Mar;60(1):74-103. [doi: [10.1177/07356331211022003](https://doi.org/10.1177/07356331211022003)]
49. Ke F, Moon J. Virtual collaborative gaming as social skills training for high - functioning autistic children. *Brit J Educational Tech* 2018 Jul;49(4):728-741. [doi: [10.1111/bjet.12626](https://doi.org/10.1111/bjet.12626)]
50. Ke F, Moon J, Sokolikj Z. Virtual reality-based social skills training for children with autism spectrum disorder. *J Spec Educ Technol* 2022 Mar;37(1):49-62. [doi: [10.1177/0162643420945603](https://doi.org/10.1177/0162643420945603)]
51. Johnson M, Tate AM, Tate K, Laane SA, Chang Z, Chapman SB. Charisma™ virtual social training: a digital health platform and protocol. *Front Virtual Real* 2022 Nov 14;3. [doi: [10.3389/frvir.2022.1004162](https://doi.org/10.3389/frvir.2022.1004162)]
52. Atherton G, Cross L. The use of analog and digital games for autism interventions. *Front Psychol* 2021;12:669734. [doi: [10.3389/fpsyg.2021.669734](https://doi.org/10.3389/fpsyg.2021.669734)] [Medline: [34447328](https://pubmed.ncbi.nlm.nih.gov/34447328/)]
53. Thomsen LA, Adjorlu A. A collaborative virtual reality supermarket training application to teach shopping skills to young individuals with autism spectrum disorder. Presented at: 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); Mar 27 to Apr 1, 2021. [doi: [10.1109/VRW52623.2021.00015](https://doi.org/10.1109/VRW52623.2021.00015)]
54. Vallefucio E, Bravaccio C, Gison G, Pepino A. Design of a serious game for enhancing money use in teens with autism spectrum disorder. In: De Paolis LT, Arpaia P, Bourdot P, editors. *Augmented Reality, Virtual Reality, and Computer Graphics: 8th International Conference, AVR 2021, Virtual Event, September 7–10, 2021, Proceedings*. [doi: [10.1007/978-3-030-87595-4_25](https://doi.org/10.1007/978-3-030-87595-4_25)]
55. Caria S, Paternò F, Santoro C, Semucci V. The design of web games for helping young high-functioning autistics in learning how to manage money. *Mobile Netw Appl* 2018 Dec;23(6):1735-1748. [doi: [10.1007/s11036-018-1069-0](https://doi.org/10.1007/s11036-018-1069-0)]
56. Ringland KE, Boyd L, Faucett H, Cullen ALL, Hayes GR. Making in Minecraft: a means of self-expression for youth with autism. Presented at: IDC '17: Proceedings of the 2017 Conference on Interaction Design and Children; Jun 27-30, 2017. [doi: [10.1145/3078072.3079749](https://doi.org/10.1145/3078072.3079749)]
57. Kandalaf MR, Didehban N, Krawczyk DC, Allen TT, Chapman SB. Virtual reality social cognition training for young adults with high-functioning autism. *J Autism Dev Disord* 2013 Jan;43(1):34-44. [doi: [10.1007/s10803-012-1544-6](https://doi.org/10.1007/s10803-012-1544-6)] [Medline: [22570145](https://pubmed.ncbi.nlm.nih.gov/22570145/)]
58. Stichter JP, Laffey J, Galyen K, Herzog M. iSocial: delivering the Social Competence Intervention for Adolescents (SCI-A) in a 3D virtual learning environment for youth with high functioning autism. *J Autism Dev Disord* 2014 Feb;44(2):417-430. [doi: [10.1007/s10803-013-1881-0](https://doi.org/10.1007/s10803-013-1881-0)] [Medline: [23812663](https://pubmed.ncbi.nlm.nih.gov/23812663/)]
59. Lahiri U, Bekele E, Dohrmann E, Warren Z, Sarkar N. Design of a virtual reality based adaptive response technology for children with autism spectrum disorder. Presented at: 4th International Conference on Affective Computing and Intelligent Interaction (ACII 2011); Oct 9-12, 2011. [doi: [10.1007/978-3-642-24600-5_20](https://doi.org/10.1007/978-3-642-24600-5_20)]
60. Grynspan O, Martin JC, Nadel J. Exploring the influence of task assignment and output modalities on computerized training for autism. *Interact Stud* 2007 Jun 27;8(2):241-266. [doi: [10.1075/is.8.2.04gry](https://doi.org/10.1075/is.8.2.04gry)]
61. García-Redondo P, García T, Areces D, Núñez JC, Rodríguez C. Serious games and their effect improving attention in students with learning disabilities. *Int J Environ Res Public Health* 2019 Jul 11;16(14):2480. [doi: [10.3390/ijerph16142480](https://doi.org/10.3390/ijerph16142480)] [Medline: [31336804](https://pubmed.ncbi.nlm.nih.gov/31336804/)]
62. Kerns KA, Macoun S, MacSween J, Pei J, Hutchison M. Attention and working memory training: a feasibility study in children with neurodevelopmental disorders. *Appl Neuropsychol Child* 2017;6(2):120-137. [doi: [10.1080/21622965.2015.1109513](https://doi.org/10.1080/21622965.2015.1109513)] [Medline: [27049769](https://pubmed.ncbi.nlm.nih.gov/27049769/)]
63. Ghanouni P, Jarus T, Zwicker JG, Lucyshyn J, Fenn B, Stokley E. Design elements during development of videogame programs for children with autism spectrum disorder: stakeholders' viewpoints. *Games Health J* 2020 Apr;9(2):137-145. [doi: [10.1089/g4h.2019.0070](https://doi.org/10.1089/g4h.2019.0070)] [Medline: [31721617](https://pubmed.ncbi.nlm.nih.gov/31721617/)]
64. Dosis S, Van der Oord S, Wiers RW, Prins PJM. Improving executive functioning in children with ADHD: training multiple executive functions within the context of a computer game. a randomized double-blind placebo controlled trial. *PLOS ONE* 2015;10(4):e0121651. [doi: [10.1371/journal.pone.0121651](https://doi.org/10.1371/journal.pone.0121651)] [Medline: [25844638](https://pubmed.ncbi.nlm.nih.gov/25844638/)]

65. Lahiri U, Warren Z, Sarkar N. Design of a gaze-sensitive virtual social interactive system for children with autism. *IEEE Trans Neural Syst Rehabil Eng* 2011 Aug;19(4):443-452. [doi: [10.1109/TNSRE.2011.2153874](https://doi.org/10.1109/TNSRE.2011.2153874)] [Medline: [21609889](https://pubmed.ncbi.nlm.nih.gov/21609889/)]
66. MacCormack J, Freeman J. Part 2: The virtual environment social program for youths with autism spectrum disorder. *Int J Play Ther* 2019;28(4):218-237. [doi: [10.1037/pla0000093](https://doi.org/10.1037/pla0000093)]
67. Wang X, Laffey J, Xing W, Galyen K, Stichter J. Fostering verbal and non-verbal social interactions in a 3D collaborative virtual learning environment: a case study of youth with autism spectrum disorders learning social competence in iSocial. *Education Tech Research Dev* 2017 Aug;65(4):1015-1039. [doi: [10.1007/s11423-017-9512-7](https://doi.org/10.1007/s11423-017-9512-7)]
68. Lerman DC, White B, Grob C, Laudont C. A clinic-based assessment for evaluating job-related social skills in adolescents and adults with autism. *Behav Analysis Practice* 2017 Dec;10(4):323-336. [doi: [10.1007/s40617-017-0177-9](https://doi.org/10.1007/s40617-017-0177-9)]
69. Elhaddadi M, Maazouz H, Alami N, et al. Serious games to teach emotion recognition to children with autism spectrum disorders (ASD). *Acta Neuropsychol* 2021 Jan 13;19(1):81-92. [doi: [10.5604/01.3001.0014.7569](https://doi.org/10.5604/01.3001.0014.7569)]
70. Amat AZ, Zhao H, Swanson A, Weitlauf AS, Warren Z, Sarkar N. Design of an interactive virtual reality system, InViRS, for joint attention practice in autistic children. *IEEE Trans Neural Syst Rehabil Eng* 2021;29:1866-1876. [doi: [10.1109/TNSRE.2021.3108351](https://doi.org/10.1109/TNSRE.2021.3108351)] [Medline: [34460376](https://pubmed.ncbi.nlm.nih.gov/34460376/)]
71. 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. [doi: [10.2196/16066](https://doi.org/10.2196/16066)] [Medline: [32207697](https://pubmed.ncbi.nlm.nih.gov/32207697/)]
72. Kahlon S, Lindner P, Nordgreen T. Virtual reality exposure therapy for adolescents with fear of public speaking: a non-randomized feasibility and pilot study. *Child Adolesc Psychiatry Ment Health* 2019;13(1):47. [doi: [10.1186/s13034-019-0307-y](https://doi.org/10.1186/s13034-019-0307-y)] [Medline: [31890004](https://pubmed.ncbi.nlm.nih.gov/31890004/)]
73. Lahiri U, Bekele E, Dohrmann E, Warren Z, Sarkar N. A physiologically informed virtual reality based social communication system for individuals with autism. *J Autism Dev Disord* 2015 Apr;45(4):919-931. [doi: [10.1007/s10803-014-2240-5](https://doi.org/10.1007/s10803-014-2240-5)] [Medline: [25261247](https://pubmed.ncbi.nlm.nih.gov/25261247/)]
74. Fornasari L, Chittaro L, Ieronutti L, et al. Navigation and exploration of an urban virtual environment by children with autism spectrum disorder compared to children with typical development. *Res Autism Spectr Disord* 2013 Aug;7(8):956-965. [doi: [10.1016/j.rasd.2013.04.007](https://doi.org/10.1016/j.rasd.2013.04.007)]
75. Ringland KE. "Autosome": fostering an autistic identity in an online minecraft community for youth with autism. In: Taylor N, Christian-Lamb C, Martin M, Nardi B, editors. *Information in Contemporary Society: 14th International Conference, iConference 2019, Washington, DC, USA, March 31–April 3, 2019, Proceedings 2019*:132-143. [doi: [10.1007/978-3-030-15742-5_12](https://doi.org/10.1007/978-3-030-15742-5_12)]
76. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977 Mar;33(1):159-174. [doi: [10.2307/2529310](https://doi.org/10.2307/2529310)] [Medline: [843571](https://pubmed.ncbi.nlm.nih.gov/843571/)]
77. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med* 2012;22(3):276-282. [doi: [10.11613/BM.2012.031](https://doi.org/10.11613/BM.2012.031)]
78. Ringland KE, Wolf CT, Boyd LE, Baldwin MS, Hayes GR. Would you be mine: appropriating Minecraft as an assistive technology for youth with autism. Presented at: ASSETS '16: Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility; Oct 23-26, 2016. [doi: [10.1145/2982142.2982172](https://doi.org/10.1145/2982142.2982172)]
79. Eve Z, Turner M, Di Basilio D, et al. Therapeutic games to reduce anxiety and depression in young people: a systematic review and exploratory meta-analysis of their use and effectiveness. *Clin Psychology Psychoth* 2024 Jan;31(1):e2938. [doi: [10.1002/cpp.2938](https://doi.org/10.1002/cpp.2938)]
80. Vygotsky LS. *Mind in Society: Development of Higher Psychological Processes*: Harvard University Press; 1978.
81. Fanning RM, Gaba DM. The role of debriefing in simulation-based learning. *Simul Healthc* 2007;2(2):115-125. [doi: [10.1097/SIH.0b013e3180315539](https://doi.org/10.1097/SIH.0b013e3180315539)] [Medline: [19088616](https://pubmed.ncbi.nlm.nih.gov/19088616/)]
82. El Hussein MT, Cuncannon A. Nursing students' transfer of learning from simulated clinical experiences into clinical practice: a scoping review. *Nurse Educ Today* 2022 Sep;116:105449. [doi: [10.1016/j.nedt.2022.105449](https://doi.org/10.1016/j.nedt.2022.105449)] [Medline: [35777294](https://pubmed.ncbi.nlm.nih.gov/35777294/)]
83. Barnett SM, Ceci SJ. When and where do we apply what we learn? A taxonomy for far transfer. *Psychol Bull* 2002 Jul;128(4):612-637. [doi: [10.1037/0033-2909.128.4.612](https://doi.org/10.1037/0033-2909.128.4.612)] [Medline: [12081085](https://pubmed.ncbi.nlm.nih.gov/12081085/)]
84. Sanders EBN, Stappers PJ. Co-creation and the new landscapes of design. *CoDesign* 2008 Mar;4(1):5-18. [doi: [10.1080/15710880701875068](https://doi.org/10.1080/15710880701875068)]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by E Oliveira; submitted 08.May.2025; peer-reviewed by C Obianyo, C Zhang, C Aygün; revised version received 17.Dec.2025; accepted 19.Dec.2025; published 10.Apr.2026.

Please cite as:

Rooij, de ST, Kuipers DA, Prins JT, Pierie JPEN

Research Design Processes in Serious Games for Adolescent Mental Health: Systematic Review

JMIR Serious Games 2026;14:e77173

URL: <https://games.jmir.org/2026/1/e77173>

doi: [10.2196/77173](https://doi.org/10.2196/77173)

© S T Rooij, de, D A Kuipers, J T Prins, J P E N Pierie. Originally published in JMIR Serious Games (<https://games.jmir.org>), 10.Apr.2026. 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.

Navigating Virtual Reality in Stroke Rehabilitation: Scoping Review of Diverse Intervention Effects

Ruixue Liu, BSc; Beibei Lin, BSc; Yanhuan Zhai, BSc; Wei Wang, PhD

Nursing Department, The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou, China

Corresponding Author:

Wei Wang, PhD

Nursing Department, The First Affiliated Hospital, Zhejiang University School of Medicine, 79 Qingchun Road, Hangzhou, China

Abstract

Background: Virtual reality (VR) technology has been increasingly explored in stroke rehabilitation due to its immersive and interactive features. However, considerable heterogeneity exists in intervention designs, study populations, and outcome measures, limiting the feasibility of conducting a systematic review.

Objective: This scoping review aims to comprehensively map randomized controlled trials (RCTs) investigating the use of VR interventions in stroke rehabilitation, with particular focus on upper limb function, gait and balance, cognitive function, and quality of life.

Methods: Following the Arksey and O'Malley framework and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines, we conducted a scoping review of RCTs published in databases from their inception to January 4, 2025. Seven databases were searched, including PubMed, Embase, Web of Science, CNKI, Wanfang, VIP, and CBM. Studies were included if they met predefined eligibility criteria, including adult patients with stroke receiving VR-based rehabilitation and a randomized controlled trial design. Exclusion criteria included non-Chinese or non-English literature, literature with unavailable full text, studies with duplicate publication or data, and studies that were irrelevant to the research topic or did not incorporate VR technology in their intervention measures. Data extracted included intervention type, sample size, training frequency and duration, outcomes, and study setting. Due to significant heterogeneity across studies, a narrative synthesis approach was used. No formal risk of bias or quality appraisal was conducted.

Results: Fifteen RCTs involving 804 patients with stroke were included. Intervention modalities varied significantly in terms of type, content, frequency, and duration. Nonimmersive VR (NIVR) interventions were more frequently applied in studies targeting upper limb function and gait training, while fully immersive VR (FIVR) was assessed in 2 head-mounted display (HMD)-based trials, whereas most studies used NIVR for upper-limb and gait-related outcomes. Many studies reported positive trends in motor function, cognitive performance, gait balance, and quality of life. However, findings were inconsistent, and not all outcomes reached statistical significance. Mild adverse events, such as fatigue or dizziness, were occasionally reported; however, no serious events occurred.

Conclusions: This scoping review outlines the research status of VR in stroke rehabilitation. VR may offer potential benefits; however, existing studies have limitations, including substantial heterogeneity in intervention protocols, limited long-term follow-up, and baseline imbalances in some trials. In addition, because this review did not include a formal quality or risk-of-bias assessment, the observed effects should be interpreted as preliminary signals rather than definitive evidence of efficacy, and the certainty of the evidence cannot be determined. Future research should standardize outcome measures, improve methodological rigor, and incorporate quality and risk-of-bias evaluation to strengthen the evidence base and support clinical implementation.

(*JMIR Serious Games* 2026;14:e72498) doi:[10.2196/72498](https://doi.org/10.2196/72498)

KEYWORDS

stroke; virtual reality; rehabilitation training; upper limb function; gait balance; quality of life

Introduction

Stroke is a disease with high incidence, high disability rate, and high recurrence rate [1-4], and it is one of the leading causes of death and long-term disability worldwide [5]. According to the World Health Organization (WHO), approximately 15 million people globally suffer from stroke each year [6,7], with about

one-third of patients left with varying degrees of functional impairment, including limited physical activity [8,9], decreased cognitive ability, and significantly reduced quality of life. Stroke rehabilitation is an important issue in clinical medicine [10] and a critical step for patients to return to normal life [11]. Although traditional rehabilitation methods, such as physical therapy and occupational therapy, have achieved significant results in

restoring basic functions, these methods have limitations in stimulating patient initiative and enhancing effectiveness [12]. Moreover, the rehabilitation needs of different patients vary significantly, and the lack of personalized treatment plans further limits the effectiveness of rehabilitation [13].

In recent years, virtual reality (VR) technology has gradually become an emerging tool in rehabilitation medicine due to its immersion, interactivity, and customizability [14]. VR technology can simulate real or virtual environments, providing patients with a safe, efficient, and highly repeatable training platform, thus enhancing patient engagement and motivation in rehabilitation [15,16]. Existing research has shown that VR has significant potential to improve upper limb function, gait balance, and cognitive abilities in patients with stroke [17]. For example, nonimmersive virtual reality (NIVR) systems can improve patients' motor coordination and goal-directed behavior through screen display and motion-sensing interaction; fully immersive virtual reality (FIVR) systems, on the other hand, use head-mounted displays (HMDs) and motion capture technology to provide patients with highly realistic virtual environments [18], enhancing their cognitive abilities and spatial awareness. In this review, FIVR specifically refers to HMD-based systems; interventions delivered via screens or monitors without an HMD were considered NIVR. However, systematic research on the specific efficacy and applicability of different types of VR technology is still lacking. Furthermore, the cost of VR equipment, its technological complexity, and issues with patient compliance also limit its widespread adoption in clinical practice.

The application of VR in stroke rehabilitation has expanded rapidly in recent years, demonstrating promising potential in facilitating functional recovery. However, this evidence base remains fragmented, with several unresolved issues. Significant heterogeneity exists across studies in terms of intervention design, outcome measures, and patient characteristics, limiting comparability and generalizability. Most studies lack multidimensional and comprehensive evaluation frameworks, relying predominantly on single functional outcomes. Furthermore, the specific advantages of FIVR versus NIVR systems at different stages of rehabilitation remain unclear [19], with some findings even appearing contradictory [20]. Long-term effects, individualized adaptation strategies, and potential adverse events associated with VR interventions have also not been systematically explored [21]. These gaps highlight a disconnect between emerging technological applications and their clinical translation.

Given the high heterogeneity in intervention modalities, assessment tools, and study populations, a conventional systematic review remains unfeasible. This study, therefore, uses a scoping review approach to comprehensively map randomized controlled trials (RCTs) published between 2016 and 2025, examining the application of various VR interventions in stroke rehabilitation. It specifically evaluates the impact of immersive and NIVR on upper limb function, gait and balance, cognitive performance, and quality of life across different rehabilitation stages. By clarifying the comparative effectiveness and contextual suitability of different VR modalities, this review aims to address critical gaps in intervention comparison and

mechanistic understanding. Scientifically, it provides a structured framework for future standardized research; clinically, it informs personalized rehabilitation strategies and supports the broader integration of VR technologies to enhance functional outcomes and long-term recovery in patients with stroke.

Methods

Defining the Research Questions

This scoping review was not registered on PROSPERO (International Prospective Register of Systematic Reviews), as scoping reviews are not consistently eligible for inclusion. Nonetheless, all methodological decisions were made a priori and adhered to throughout the review process. This study defines the research questions based on the PCC principle (Population, Concept, and Context) [22]. The methodology was guided by the scoping review framework proposed by Arksey and O'Malley, and the reporting adhered to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [22]. The study population (Population) consists of patients with stroke who receive VR technology interventions, including both patients with ischemic and hemorrhagic stroke aged ≥ 18 years. The research concept (Concept) focuses on VR-based interventions. VR [23] is a medium composed of interactive computer simulations that can sense the participant's location and movement, providing feedback to one or more senses, allowing patients to interact with the virtual environment in a way that mimics real-life interactions. Based on the level of immersion, VR technology can be categorized into NIVR, semi-immersive, and FIVR types. NIVR [24] is presented in 2 dimensions and interacts through a computer display or gaming system; semi-immersive VR [25] creates 3D images using stereoscopic projection or fixed-angle displays; FIVR [26] allows users to interact in real-time through an HMD, typically combined with motion tracking, to achieve a high sense of presence in a virtual environment. Operationally, in this review, we classified an intervention as FIVR only when an HMD was explicitly used; screen-based systems (including motion-capture or force-platform games viewed on a monitor) were classified as NIVR. The research context (Context) is patients with stroke receiving VR interventions in hospital or home settings.

To systematically summarize the application and effects of VR in stroke rehabilitation, this study addresses the following questions: (1) What types of VR technologies are available at the time of this writing? (2) What are VR technology's main contents and applications in stroke rehabilitation? (3) What are the rehabilitation effects, safety, and limitations of different types of VR interventions?

Literature Search

The inclusion and exclusion criteria are presented in Table 1. To ensure a systematic and reproducible search process, we performed comprehensive searches across 7 databases, including PubMed, Embase, Web of Science, CNKI, VIP Chinese Science and Technology Journal Database, Wanfang Medical Journal Database, and China Biomedical Literature Database (SinoMed). The search period covered from database inception to January 4, 2025. Search strategies incorporated both controlled

vocabulary (eg, MeSH [Medical Subject Headings]) and free-text terms. The final search strategies were peer-reviewed by two information specialists to ensure rigor and completeness.

Detailed, reproducible search strings for each database are provided in [Multimedia Appendix 1](#).

Table . Inclusion and exclusion criteria for randomized controlled trials of virtual reality–based rehabilitation in adult patients with stroke.

Criteria category	Inclusion criteria	Exclusion criteria
Population	Adult patients with stroke (≥ 18 years), diagnosed with ischemic or hemorrhagic stroke	Patients with pediatric (< 18 years); patients with nonstroke neurological or musculoskeletal disorders
Intervention	Virtual reality (VR)–based rehabilitation interventions, including nonimmersive (NIVR), semi-immersive, or fully immersive (FIVR) systems	Non-VR rehabilitation methods, such as traditional physical or occupational therapy without VR integration
Comparator	Studies with or without control or comparator groups (eg, conventional rehabilitation, sham VR, or no treatment)	^a
Outcomes	Reported at least one rehabilitation-related outcome, including motor function, gait, balance, cognitive function, quality of life, or safety of the intervention	Studies without reported outcomes or unrelated to rehabilitation efficacy
Study design	Randomized controlled trials (RCTs)	Non-RCT designs such as case reports, cross-sectional studies, reviews, protocols, editorials
Language	Published in English or Chinese	Literature published in other languages without available translation
Publication type	Peer-reviewed full-text journal articles	Abstract-only publications, dissertations, conference papers, gray literature
Full-text availability	Full text available	Full text unavailable or inaccessible
Duplication	Original studies with unique data	Duplicate publications or overlapping datasets already included

^aNot applicable.

Table 1 outlines the predefined inclusion and exclusion criteria applied in this scoping review of RCTs assessing VR-based rehabilitation interventions for adult patients with stroke. The review included studies published in English or Chinese between database inception and January 4, 2025, identified through comprehensive searches in 7 databases (PubMed, Embase, Web of Science, CNKI, VIP, Wanfang, and SinoMed). Eligible studies enrolled adult patients (aged 18 years or older) diagnosed with ischemic or hemorrhagic stroke and implemented VR-based interventions, including nonimmersive, semi-immersive, or fully immersive systems, either as standalone therapy or in combination with conventional rehabilitation. Studies were required to report at least one rehabilitation-related outcome such as motor function, gait, balance, cognitive function, quality of life, or safety. Nonrandomized designs, nonstroke populations, and studies without available full text or outcome data were excluded.

Literature Screening and Data Extraction

Duplicate records were removed using EndNote X9 (Clarivate). Two types of duplication were addressed at different screening stages. The first involved technical duplicates identified by EndNote X9 based on identical bibliographic details (eg, title, author, and DOI). The second, labeled as “duplicate content” during full-text screening, referred to studies with substantially overlapping content despite differences in bibliographic metadata, for example, publications in different languages or multiple reports from the same research dataset. Such redundant

publications were excluded following careful comparison to prevent data duplication bias. Two reviewers (RL and BL) independently screened studies based on predefined inclusion and exclusion criteria, with reasons for exclusion documented at each stage. Discrepancies were resolved through discussion or consultation with a third reviewer. The selection process was recorded using a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)-compliant flow diagram ([Checklist 1](#)). Data extraction was performed using a standardized form with 5 - 10 predefined variables, which were adjusted as needed to align with the study objectives. Extracted variables included study location, stroke type, sample size, intervention characteristics (type, content, setting, frequency, and duration), and outcome measures (motor function, gait and balance, cognitive function, health-related quality of life, and safety). Extraction was conducted independently by 2 reviewers (RL and BL) and cross-checked for consistency. As a scoping review, this study aimed to map this evidence rather than assess methodological quality or synthesize effect sizes; thus, no risk of bias or quality assessment was performed.

Heterogeneity Handling

Due to significant heterogeneity in intervention types, equipment usage, and outcome measure assessments across studies, this study did not perform a meta-analysis but used a descriptive approach to summarize the results. The potential impact of heterogeneity on the study outcomes is further discussed in the discussion section.

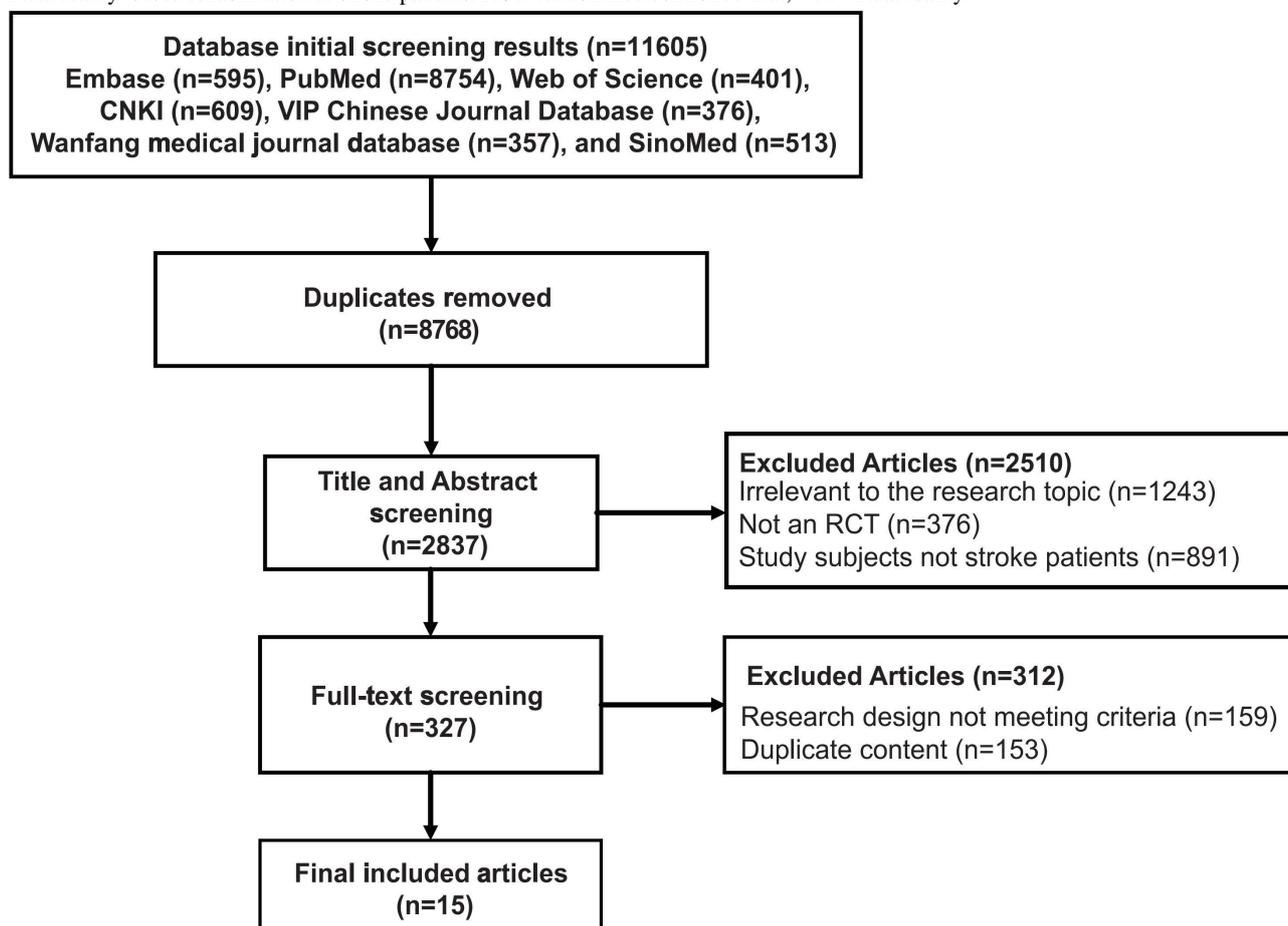
Results

Literature Screening Process and Results

A total of 11,605 records were initially identified. After removing 8768 duplicates, 2837 records remained for title and abstract screening. At this stage, 2510 records were excluded

for being unrelated to the study topic, leaving 327 articles for full-text assessment. Of these, 312 were excluded, 159 due to inappropriate study design and 153 due to content duplication. Ultimately, 15 studies were included in the review (13 English articles and 2 Chinese articles) [27-41]. The specific screening process is documented in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of literature screening for the scoping review of virtual reality-based rehabilitation in stroke patients. RCT: randomized controlled trial; VR: virtual reality.



Basic Characteristics of the Included Studies

A total of 15 RCTs [27-41] were included in this study, involving 804 patients with stroke aged 45-85 years. The studies were published between 2016 and 2024. The included studies were classified and summarized based on intervention types, patient stages, and research objectives to present the study characteristics (Multimedia Appendix 2) [27-41]. Among these, 5 studies [27,32-34,41] compared VR training alone with conventional rehabilitation training; 10 studies [28-31,35-40] compared conventional rehabilitation training combined with

FIVR or NIVR with conventional rehabilitation training. Nine studies [28,29,31-33,35,36,39,40] focused on upper limb rehabilitation; 4 studies [27,33,34,38] assessed cognitive or psychological outcomes. Among them, 2 studies [27,34] focused on cognitive function and anxiety improvement, 3 studies [37,38,41] focused on lower limb gait and balance improvement, and 1 study [30] focused on safety and feasibility assessment. Three studies [30,34,36] reported adverse events such as fatigue, transient dizziness, and shoulder pain, all alleviated after rest, with no severe adverse events reported (see Table 2).

Table . Forms and effects of virtual reality interventions in stroke rehabilitation research.

VR ^a form	Core feature	Applied research number	Main application target
Non-fully immersive (screen-based/non-HMD ^b)	Use 2D screen, game controller	13	Improved upper limb function and gait balance
Full immersion (HMD-based)	Simulate a virtual environment using an HMD	2	Cognitive improvement and homeostasis improvement

^aVR: virtual reality.

^bHMD: head-mounted display.

Table 2 summarizes the forms and functional applications of VR interventions reported in stroke rehabilitation research. Data were synthesized from 15 randomized controlled trials conducted between 2016 and 2024 across 9 countries, including China, Spain, Canada, Turkey, and Australia. Nonimmersive VR systems, typically based on 2D screen interaction, were primarily used for upper-limb motor and gait-balance training. Fully immersive VR systems using head-mounted displays were evaluated in a limited number of trials (n=2); therefore, evidence is insufficient to draw firm conclusions regarding modality-specific application patterns (eg, cognitive or integrative rehabilitation). The evidence indicates that different levels of immersion may be associated with distinct therapeutic goals and patient suitability, and modality-specific trends should be interpreted cautiously given the small number of HMD-based FIVR trials.

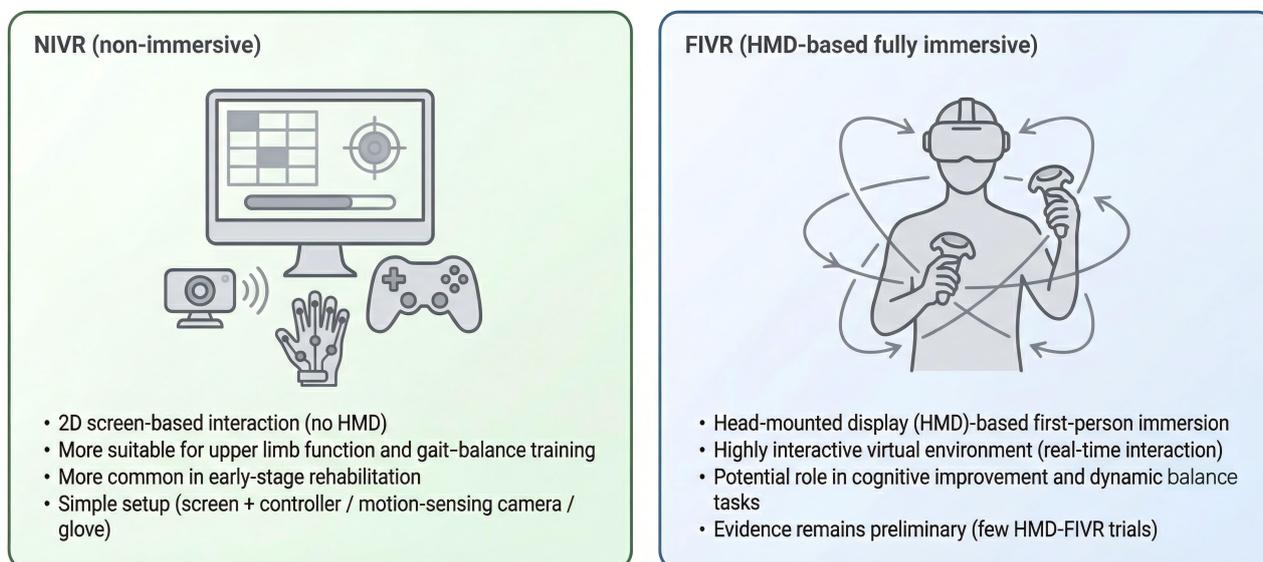
Overall, the majority of included studies (n=9 [28,29,31-33,35,36,39,40]) focused on upper limb rehabilitation,

suggesting the potential of VR technology to support fine motor and strength training. Several studies reported improvements in gait and balance, indicating possible benefits of VR-assisted rehabilitation for dynamic postural control. However, substantial heterogeneity remains across study designs and outcomes, underscoring the need for further high-quality trials. A small number of studies investigated cognitive and psychological outcomes, though this evidence in these domains remains limited.

Differences in the Intervention Effects and Applicability of NIVR and FIVR

Among the 15 RCTs included in this study, two types of VR interventions were mainly explored: NIVR and FIVR. Each type played a different role in the rehabilitation of patients with stroke. The applications of NIVR and FIVR in rehabilitation varied, showing different effects based on the research goals and intervention stages (**Figure 2**).

Figure 2. Comparison of nonimmersive and fully immersive virtual reality in stroke rehabilitation. FIVR: fully immersive virtual reality.



NIVR intervention was applied in 13 studies [27-33,36-41] primarily targeting upper limb function and gait-balance training. These interventions commonly used 2D screens, customized game controllers, or training gloves. Several studies reported improvements in upper limb flexibility, strength, and gait stability, indicating potential utility across rehabilitation stages, although the timing of application varied across trials.

In contrast, FIVR intervention was applied in 2 studies [34,35], in which cognitive and/or integrative functional outcomes were

assessed, often within highly interactive virtual environments. Findings indicated potential benefits in attention, visuomotor coordination, and dynamic postural control; however, given the limited number of HMD-based FIVR trials, evidence remains insufficient to draw firm conclusions regarding modality-specific application patterns or comparative advantage (**Table 2**).

The study subjects included acute, subacute, and patients with chronic stroke, including both ischemic and hemorrhagic stroke types. One study [41] focused on acute patients with stroke; the

study is a force platform-screen interactive VR training, 10 studies [28,30,32,34-40] focused on subacute patients with stroke, and 4 studies [27,31,33,37] focused on patients with chronic stroke. Only one study [33] occurred at home, while the rest were conducted in hospitals. The included studies came from various countries, including China [28,35,39,41], Spain [29,31], Turkey [30,37], Canada [32], the United Kingdom [34], France [36], Italy [38], Norway [39], and Australia [33], ensuring a certain level of representativeness.

The application range and effects of the 2 VR forms in rehabilitation demonstrate their respective advantages and characteristics. NIVR, due to its simple equipment and ease of operation, is widely used to recover motor function and gait balance. In contrast, FIVR offers a more immersive and dynamic cognitive and balance training experience through its high level of immersion and complex virtual environments. Based on this evidence, each form of VR intervention exhibits specific strengths depending on rehabilitation goals and stages. Selection of the appropriate modality should be tailored to individual patient needs, and future studies are needed to further clarify their comparative effectiveness and clinical applicability.

Training Platforms and Content

Three studies [30-32] used commercial games based on the Nintendo Wii (Nintendo) and Xbox Kinect (Microsoft Corp) for rehabilitation training. Games based on Nintendo Wii included card games, bingo, jigsaw puzzles, and ball games. For upper limb rehabilitation, the Wii Sports package (bowling, golf, and tennis) was used, while for lower limb rehabilitation, the Wii Fit balance training package (header ball, skiing obstacle course, tilt platform, wire tension, downstream, and ice fishing) was used. Games based on the Kinect Sports software package included bowling and “Whac-A-Mole.”

Eleven studies [27-29,33-40] used customized rehabilitation games. These included (1) functional software and hand-tutor gloves that simulate everyday tasks that patients with strokes must complete. During practice, visual and auditory feedback is provided for success or failure, scores are displayed, and therapists can adjust the sensitivity of movements to prevent frustration and reduce motivation loss. (2) The cognitive rehabilitation program (VIRTUE program) was designed after close consultation between clinical experts and patient representatives. Patients, with the help of a therapist, use VIRTUE to complete tasks in various scenarios, such as in the bedroom (making the bed and choosing clothes); the bathroom

(brushing teeth and showering); the kitchen (toasting bread, preparing tea, washing dishes, cooking pasta, and using a coffee machine); the café (choosing a set meal and paying the bill); a restaurant (interacting with the waiter, ordering, or paying); and a garden (watering plants). The primary requirement is for patients to perform all or part of the activities of daily living in appropriate environments, with necessary materials such as food, kitchen utensils, and coins easily accessible while immersed in the environment. (3) Rehabilitation training using a posture control system (BioFlex-FP), including tasks like “Big Fish Eats Little Fish”, this task focuses on training weight shift in the front-back and left-right directions while also training muscle strength and endurance; “Phoenix Flying”, this task primarily targets the patient’s static balance; “Mushroom Picking”, this task mainly trains balance ability for weight shifting in all directions, stepping, and shifting weight vertically. (4) In a virtual kitchen, patients control a pan handle to cook dumplings and noodles. In a virtual fencing hall, they control a sword to burst balloons; in a virtual boxing ring, they control a large fist to hit a dummy; on a virtual basketball court, they collect eggs in a virtual basket using a controller; and in a virtual office, they organize the desk, moving items to designated positions, and perform specific shoulder joint flexion angle (30°, 60°, and 90°) training. Two studies [37,38] used 2D screen-based running and gait training, such as using the Lokomat gait orthosis to guide subjects in leg movements within the sagittal plane of the hip and knee joints, providing haptic feedback.

Frequency and Duration of VR Interventions

In the 15 studies included, the frequency and duration of VR interventions varied depending on the intervention content (Table 3). Ten studies [27,29,30,32,34-36,38-40] required 5 VR training sessions per week, 4 studies [28,31,33,37] required 2-4 sessions per week, and one study [41] required 6 sessions per week. The training duration was 1 hour per session in 5 studies [30,32,35,36,39], 20 - 45 minutes per session in 6 studies [27,33,37,38,40,41], 2.5 - 3.5 hours per session in 3 studies [28,29,31], and one study [34] used personalized session durations. Additionally, the total intervention duration also varied. One study [28] had a total duration of 56 hours, one study [29] had a total duration of 30 - 40 hours, 4 studies [30,31,36,38] had total durations of 20 - 30 hours, 6 studies [27,32,33,35,39,40] had total durations of 10 - 20 hours, 2 studies [37,41] had a total duration of 9 hours, and one study [34] had personalized intervention duration.

Table . Distribution of intervention frequency and duration in relevant rehabilitation studies.

Intervention frequency	Number of documents	Single duration	Number of documents	Total time	Number of documents
2 - 4 times a week	4	20 - 45 minutes/time	6	9 hours	2
5 times a week	10	1 hour/time	5	10 - 20 hours	6
6 times a week	1	2.5 - 3.5 hour/time	3	20 - 30 hours	4
Personalized frequency	0	Personalized duration	1	30 - 40 hours	1
^a	—	—	—	56 hours	1
—	—	—	—	Personalized total time	1

^aNot applicable.

Table 3 provides a systematic summary of dosage characteristics of VR rehabilitation interventions across 15 RCTs, including training frequency, single-session duration, and total intervention time. The included studies were published between 2016 and 2024 and were conducted across 9 countries, including China, Spain, Canada, Turkey, and Australia, covering a broad range of clinical and rehabilitation settings. The VR protocols encompassed both nonimmersive and fully immersive systems and were mainly compared with conventional rehabilitation or implemented as an adjunct to standard therapy. Considerable variability was observed in intervention dosage: training frequency ranged from 2 to 6 sessions per week; single-session duration varied from 20 minutes to 3.5 hours, depending on the VR system type, training objectives, and patients' functional status; and total intervention time ranged from 9 to 56 hours, with one study reporting a personalized total intervention duration. Overall, this table offers an independent and comprehensive overview of VR dosage designs in contemporary poststroke RCTs and highlights substantial differences in implementation strategies, training intensity, and clinical applicability across settings and patient populations.

The results suggest that the VR intervention frequency and duration design should be adjusted based on the patient's rehabilitation stage and specific goals. A common and effective approach is to conduct 3-5 weekly sessions, each lasting 20 minutes to 1 hour in duration. High-frequency training is suitable for patients in the acute phase, while personalized plans offer potential solutions for more complex cases.

Effects of VR Interventions in Stroke Rehabilitation and Heterogeneity Analysis

This study evaluated 15 studies that assessed the effects of VR interventions on various aspects of stroke rehabilitation, including motor function, health-related quality of life, gait and balance, cognitive and psychological status, and safety and feasibility. Notably, although every included study reported at least one statistically significant finding, several outcome measures within individual studies did not reach statistical significance.

Regarding motor function, 9 studies [27-32,34-36] assessed improvements in upper and lower limb function. In these studies, screen-based NIVR was frequently used to support motor training through interactive feedback. Because only a small number of trials used HMD-based FIVR, this evidence is insufficient to conclude modality-specific preferences (eg, upper- vs lower-limb targets) or comparative advantages.

Regarding health-related quality of life, 9 studies [27,28,31,32,34-36,38,39] investigated the effects of VR interventions on patients' quality of life, with the primary measurement tools being the short form health survey 36 (SF-36) and stroke impact scale (SIS). Some studies found that VR interventions, whether applied alone or in combination with conventional rehabilitation, may be associated with improvements in patients' quality of life, particularly showing positive trends in daily functional independence and mental well-being.

The improvement of gait and balance was the focus of 3 studies [31,37,38], using tools such as the Berg Balance Scale and Functional Gait Assessment (FGA). These predominantly screen-based interventions showed mixed but potentially beneficial effects on postural control. However, substantial heterogeneity exists across studies. Regarding cognitive and psychological status, 2 studies [27,34] explored the effects of VR interventions on cognitive function and anxiety relief, using the Montreal Cognitive Assessment (MoCA) and the Hospital Anxiety and Depression Scale (HADS) as measurement tools. Some studies suggested that VR interventions may be associated with changes in cognitive attention and psychological status, particularly during complex task training delivered in interactive virtual settings. While certain findings indicate positive trends, the overall evidence remains limited, and the specific contribution of HMD-based immersion cannot be determined given the small number of HMD-based FIVR trials.

Regarding safety and feasibility, 3 studies [30,34,36] reported mild adverse events, including fatigue, transient dizziness, and shoulder pain, all of which were alleviated with rest, and no serious adverse events occurred. The included studies indicate that VR interventions in patients with stroke appear to be generally safe and feasible; however, the limited sample sizes and insufficient reporting details warrant further investigation to confirm these findings. However, there was significant heterogeneity in the interventions, outcome measurement tools, and patient characteristics. Specifically, the target functions of NIVR and FIVR differ, and the use of measurement tools was inconsistent across studies. This heterogeneity may have contributed to inconsistencies in the outcome measures. Future research should address this issue by standardizing intervention protocols and unifying measurement tools.

Discussion

Principal Findings

In recent years, VR technology has emerged as an innovative approach to rehabilitation, gaining widespread attention in the medical field [16]. Previous studies have reported that VR can provide an immersive and highly interactive training environment for stroke rehabilitation [42], showing potential value in improving motor function, enhancing cognitive abilities, and increasing quality-of-life-related outcomes [42-45]. Following a stroke, patients often experience limb dysfunction, gait and balance impairments, and cognitive deficits. Traditional rehabilitation methods frequently lack sufficient diversity and personalized intervention.

This review highlights that as a novel adjunct therapy, VR's unique virtual interaction and real-time feedback mechanisms may play a crucial role in addressing these challenges. However, although numerous studies have reported positive rehabilitation outcomes with VR intervention, the effectiveness of different types of VR interventions and their specific applications remains insufficiently validated. This represents a key challenge in the field of stroke rehabilitation. Existing studies have shown that NIVR and FIVR interventions have distinct effects and scopes of application in stroke rehabilitation. NIVR, characterized by its simplicity and ease of use, is more commonly applied in the

early stages of rehabilitation, particularly in upper limb recovery and gait-balance training. This aligns with its use of basic interactive and feedback mechanisms, often delivered through gamified tasks to enhance patient motivation and engagement. In contrast, HMD-based FIVR has been investigated in only a small number of trials ($n=2$) to explore potential effects on cognitive function and dynamic balance. Although highly interactive virtual environments may enhance immersion and user engagement, the evidence remains preliminary, and it is not possible to draw firm conclusions regarding modality-specific advantages or preferential application domains (eg, lower limb function or cognitive attention).

Comparison to Prior Work

The studies included in this review demonstrate substantial heterogeneity in this VR intervention research. Variations in intervention protocols, assessment tools, and individual patient characteristics contribute to inconsistencies in outcome measures. For example, the target populations and training objectives differ between NIVR and FIVR applications. The selection of evaluation instruments, such as functional gait assessments and cognitive measurement tools, is not standardized across studies, which may account for differences in reported outcomes. Additionally, variability in intervention frequency and duration may impact overall effectiveness. Most studies adopted protocols involving three to five sessions per week, with each session lasting 20 minutes to 1 hour. Although some studies have reported positive effects on multiple rehabilitation outcomes, particularly in patients with acute-phase conditions and those with complex conditions, research on personalized intervention strategies remains limited and should be prioritized in future investigations.

VR technology represents a novel adjunct in stroke rehabilitation. Existing studies have explored its use in promoting motor function and neural engagement. NIVR interventions have predominantly focused on upper limb function and gait and balance, while FIVR has been associated with cognitive and dynamic postural training. These effects may also contribute to increased training adherence and improvements in quality of life. Most studies did not report serious adverse events, and VR interventions were generally considered safe. Mild adverse effects, such as fatigue, transient dizziness, and shoulder discomfort, were occasionally reported. However, the high cost and technical complexity of FIVR systems may limit their clinical use, especially in resource-limited settings. Reducing equipment costs and improving usability are key areas for future development. The considerable heterogeneity among the included studies, in terms of VR intervention types, assessment tools, and patient profiles, may have affected the consistency of the findings.

Limitations

While VR interventions have shown promise in stroke rehabilitation, several limitations remain. First, significant heterogeneity exists across studies in terms of device types, training content, frequency, and duration, which may impact the consistency of outcomes. Second, most included RCTs were short-term and lacked follow-up, limiting understanding of long-term efficacy. Third, some studies did not match baseline

characteristics well, potentially underestimating individual variability. Fourth, older adults and individuals with cognitive impairments may experience difficulty interacting with complex VR systems, which could compromise engagement and the intervention's effectiveness. Developing simplified and adaptive interfaces could help mitigate this issue and improve accessibility. Fifth, this review did not conduct a formal assessment of methodological quality or risk of bias for the included RCTs (eg, using the Cochrane RoB tool or the Physiotherapy Evidence Database [PEDro] scale). Therefore, the synthesized effects and between-group differences should be interpreted as preliminary and indicative, reflecting patterns and trends in the evidence rather than definitive efficacy conclusions. Future systematic reviews should incorporate standardized quality and risk-of-bias appraisal to strengthen the reliability and robustness of the evidence.

Future Directions

Future research should further explore the differential effects of VR interventions across various stroke subgroups, particularly between acute and chronic phases. Given the heterogeneity in recovery trajectories, developing personalized intervention strategies is essential. Selecting appropriate types of VR, such as nonimmersive versus fully immersive systems, based on specific patient needs could enhance clinical relevance and therapeutic outcomes. Additionally, integrating VR with conventional rehabilitation modalities, such as physical therapy, electrical stimulation, and robotic training, may offer synergistic benefits.

To improve comparability and rigor across studies, the adoption of standardized assessment tools and outcome measures is recommended. Moreover, the potential of emerging technologies such as brain-computer interfaces, artificial intelligence, and telerehabilitation platforms warrants further exploration, particularly for enhancing the precision and individualization of interventions. Finally, to increase accessibility, especially in resource-limited or home-based settings, future efforts should focus on reducing equipment costs, simplifying user interfaces, and improving system usability.

Conclusion

This scoping review highlights the growing role of VR as an innovative adjunct to conventional stroke rehabilitation. This evidence indicates that VR-based interventions may be associated with improvements in upper limb function, gait and balance, cognitive performance, and health-related quality of life, with a generally favorable safety profile (Figure 3). Across the included trials, NIVR was more commonly applied to upper limb rehabilitation and gait-balance training, particularly in earlier rehabilitation stages. HMD-based FIVR was examined in only a small number of trials; therefore, any modality-specific observations regarding cognitive rehabilitation or dynamic postural tasks should be considered preliminary and interpreted with caution. These modality-specific observations should be interpreted as potential trends rather than confirmed comparative superiority, given the substantial heterogeneity in intervention protocols, outcome measures, and patient characteristics, as well as the limited evidence on long-term outcomes. A

comparative overview of VR intervention types, functional targets, and application characteristics is presented in Table 4.

Figure 3. Virtual reality intervention strategies for different stroke recovery stages.

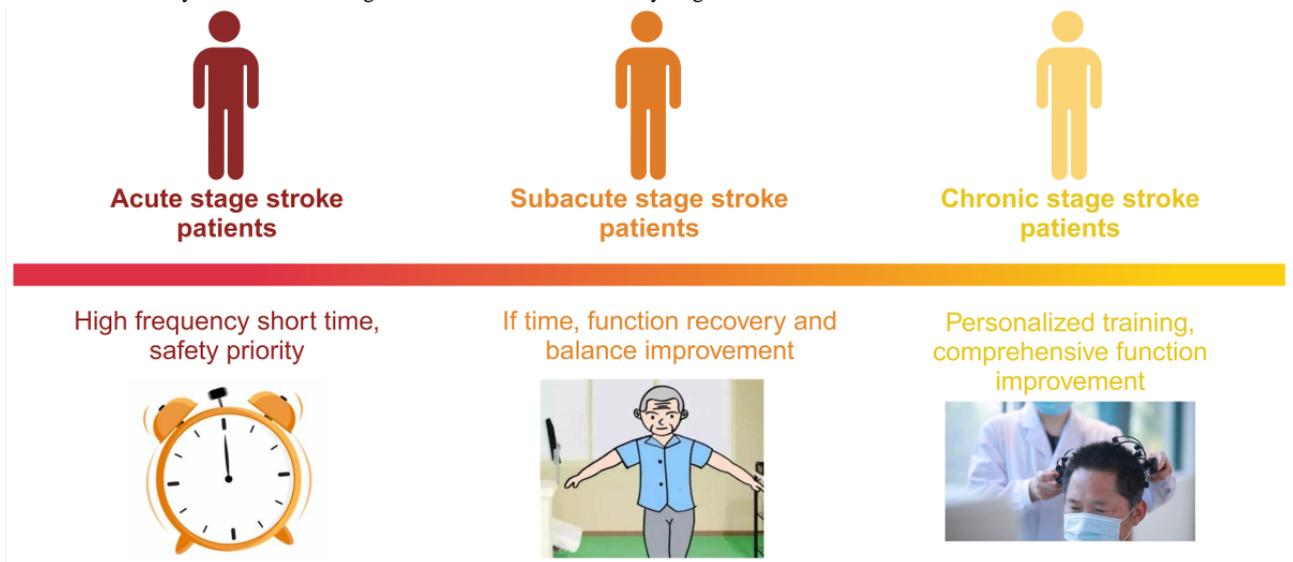


Table . Summary of key findings and insights in virtual reality–based stroke rehabilitation research.

Content category	Main findings and insights
Research background	VR ^a technology has attracted increasing attention in stroke rehabilitation due to its strong immersion and high interactivity.
Differences in intervention types	Nonimmersive VR is more suitable for upper limb function and gait-balance training, whereas HMD ^b -based fully immersive VR has been investigated in only a small number of trials; its potential role in cognitive improvement and dynamic balance tasks remains preliminary.
Effectiveness evaluation	Most studies suggest that VR interventions improve motor function, cognitive ability, and quality of life, with generally high safety and mild side effects.
Heterogeneity in Research	Significant inconsistencies exist in VR intervention protocols, assessment tools, and individual characteristics, limiting the comparability of results.
Limitations	Lack of long-term follow-up, standardized assessment tools, and high-quality RCTs ^c ; some populations show poor adaptability to VR.
Future Directions	Future research should focus on personalized intervention strategies, integration with AI ^d and BCI ^e technologies, and enhancing accessibility and home-based applications.

^aVR: virtual reality.

^bHMD: head-mounted display.

^cRCT: randomized controlled trial.

^dAI: artificial intelligence.

^eBC: brain-computer Interface.

Acknowledgments

During the preparation of this manuscript, the authors used the generative artificial intelligence tool ChatGPT (OpenAI) to assist with language editing and refinement of the initial draft. All study design, data analysis, and interpretation of results were independently conducted by the authors. The AI tool was not used to generate any research data, results, or conclusions.

Funding

The authors declare no financial support was received for this work.

Data Availability

The datasets used or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

RL conceived and designed the study, performed data curation and formal analysis, and drafted the initial manuscript. BL contributed to literature search, data extraction, and methodology, and assisted in manuscript review and editing. YZ participated in validation, visualization, and critical revision of the manuscript. WW supervised the overall project, provided guidance throughout the research process, and served as the corresponding author. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategies.

[[DOCX File, 26 KB - games_v14i1e72498_app1.docx](#)]

Multimedia Appendix 2

Characteristics of the included studies (n=15).

[[DOCX File, 24 KB - games_v14i1e72498_app2.docx](#)]

Checklist 1

PRISMA checklist.

[[PDF File, 26 KB - games_v14i1e72498_app3.pdf](#)]

References

1. Steinmetz JD, Seeher KM, Schiess N, et al. Global, regional, and national burden of disorders affecting the nervous system, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet Neurol* 2024 Apr;23(4):344–381. [doi: [10.1016/S1474-4422\(24\)00038-3](https://doi.org/10.1016/S1474-4422(24)00038-3)]
2. Ferrari AJ, Santomauro DF, Aali A. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. *Lancet* 2024 May 18;403(10440):2133–2161. [doi: [10.1016/S0140-6736\(24\)00757-8](https://doi.org/10.1016/S0140-6736(24)00757-8)] [Medline: [38642570](https://pubmed.ncbi.nlm.nih.gov/38642570/)]
3. Govori V, Budinčević H, Morović S, Đerke F, Demarin V. Updated perspectives on lifestyle interventions as secondary stroke prevention measures: a narrative review. *Med Bogota Colomb* 2024 Mar 19;60(3):504. [doi: [10.3390/medicina60030504](https://doi.org/10.3390/medicina60030504)]
4. Ganesh A, Luengo-Fernandez R, Wharton RM, et al. Time course of evolution of disability and cause-specific mortality after ischemic stroke: implications for trial design. *J Am Heart Assoc* 2017 Jun 11;6(6):e005788. [doi: [10.1161/JAHA.117.005788](https://doi.org/10.1161/JAHA.117.005788)] [Medline: [28603141](https://pubmed.ncbi.nlm.nih.gov/28603141/)]
5. Johnson CO, Nguyen M, Roth GA, et al. 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. [doi: [10.1016/S1474-4422\(19\)30034-1](https://doi.org/10.1016/S1474-4422(19)30034-1)]
6. Berehan HB, Dessie ZG, Tesfaw LM. Exploring functional abilities and competing risks among stroke patients: a longitudinal and survival analysis study at Felege Hiwot Referral Hospital, Ethiopia. *BMJ Open* 2024 May 2;14(5):e073384. [doi: [10.1136/bmjopen-2023-073384](https://doi.org/10.1136/bmjopen-2023-073384)] [Medline: [38697761](https://pubmed.ncbi.nlm.nih.gov/38697761/)]
7. Luo X, Zhu S, Song Z. Quantifying the income-increasing effect of digital agriculture: take the new agricultural tools of smartphone as an example. *Int J Environ Res Public Health* 2023 Feb 10;20(4):3127. [doi: [10.3390/ijerph20043127](https://doi.org/10.3390/ijerph20043127)] [Medline: [36833825](https://pubmed.ncbi.nlm.nih.gov/36833825/)]
8. Wu S, Wu B, Liu M, et al. Stroke in China: advances and challenges in epidemiology, prevention, and management. *Lancet Neurol* 2019 Apr;18(4):394–405. [doi: [10.1016/S1474-4422\(18\)30500-3](https://doi.org/10.1016/S1474-4422(18)30500-3)] [Medline: [30878104](https://pubmed.ncbi.nlm.nih.gov/30878104/)]
9. Meyer J, Gurrin C, Price B, Kay J, Jain R. Editorial: long-term self-tracking for life-long health and well-being. *Front Digit Health* 2021;3:827586. [doi: [10.3389/fgdth.2021.827586](https://doi.org/10.3389/fgdth.2021.827586)] [Medline: [35005699](https://pubmed.ncbi.nlm.nih.gov/35005699/)]
10. Konishi T, Inokuchi H, Sasabuchi Y, et al. Association between care-need level after discharge and long-term outcomes in 7491 patients requiring rehabilitation for stroke. *JMA J* 2024 Jan 15;7(1):52–59. [doi: [10.31662/jmaj.2023-0094](https://doi.org/10.31662/jmaj.2023-0094)] [Medline: [38314431](https://pubmed.ncbi.nlm.nih.gov/38314431/)]

11. de Goes ACG, Souza KAS, Tunes G, et al. Predictive value of functional disability scales among stroke survivors: a long-term mortality evaluation in a Brazilian stroke cohort. *J Stroke Cerebrovasc Dis* 2024 Jan;33(1):107487. [doi: [10.1016/j.jstrokecerebrovasdis.2023.107487](https://doi.org/10.1016/j.jstrokecerebrovasdis.2023.107487)] [Medline: [37980846](https://pubmed.ncbi.nlm.nih.gov/37980846/)]
12. Salatino A, Burin D. Editorial: virtual, mixed and augmented reality in cognitive neuroscience and neuropsychology, volume II. *Front Psychol* 2024;15:1444412. [doi: [10.3389/fpsyg.2024.1444412](https://doi.org/10.3389/fpsyg.2024.1444412)] [Medline: [38966743](https://pubmed.ncbi.nlm.nih.gov/38966743/)]
13. Muheim J, Hotz I, Kübler F, et al. ExerG - an exergame-based training device for the rehabilitation of older adults: a functional model usability study. *BMC Geriatr* 2024 Dec 21;24(1):1029. [doi: [10.1186/s12877-024-05617-7](https://doi.org/10.1186/s12877-024-05617-7)] [Medline: [39709349](https://pubmed.ncbi.nlm.nih.gov/39709349/)]
14. Sánchez-Gil JJ, Sáez-Manzano A, López-Luque R, Ochoa-Sepúlveda JJ, Cañete-Carmona E. Gamified devices for stroke rehabilitation: a systematic review. *Comput Methods Programs Biomed* 2025 Jan;258:108476. [doi: [10.1016/j.cmpb.2024.108476](https://doi.org/10.1016/j.cmpb.2024.108476)] [Medline: [39520875](https://pubmed.ncbi.nlm.nih.gov/39520875/)]
15. Bang M, Kim MA, Kim SS, Kim HS. Cognitive training using virtual reality: an assessment of usability and adverse effects. *Arch Rehabil Res Clin Transl* 2024 Dec;6(4):100378. [doi: [10.1016/j.arrct.2024.100378](https://doi.org/10.1016/j.arrct.2024.100378)] [Medline: [39822192](https://pubmed.ncbi.nlm.nih.gov/39822192/)]
16. Barger S, Baggio M, Guida S, Castellini G, Gianola S. Evolving trends of systematic reviews on virtual reality for stroke rehabilitation. *Arch Physiother* 2024;14:182-188. [doi: [10.33393/aop.2024.3155](https://doi.org/10.33393/aop.2024.3155)] [Medline: [39822392](https://pubmed.ncbi.nlm.nih.gov/39822392/)]
17. Palidis DJ, Gardiner Z, Stephenson A, Zhang K, Boruff J, Fellows LK. The use of extrinsic performance feedback and reward to enhance upper limb motor behavior and recovery post-stroke: a scoping review. *Neurorehabil Neural Repair* 2025 Feb;39(2):157-173. [doi: [10.1177/15459683241298262](https://doi.org/10.1177/15459683241298262)] [Medline: [39659261](https://pubmed.ncbi.nlm.nih.gov/39659261/)]
18. Lo SHS, Chau JPC, Choi KC, et al. Stroke survivor and caregiver experiences of virtual reality gaming to promote social participation: a qualitative study. *PLoS ONE* 2024;19(12):e0315826. [doi: [10.1371/journal.pone.0315826](https://doi.org/10.1371/journal.pone.0315826)] [Medline: [39693334](https://pubmed.ncbi.nlm.nih.gov/39693334/)]
19. Salatino A, Zavattaro C, Gammeri R, et al. Virtual reality rehabilitation for unilateral spatial neglect: a systematic review of immersive, semi-immersive and non-immersive techniques. *Neurosci Biobehav Rev* 2023 Sep;152:105248. [doi: [10.1016/j.neubiorev.2023.105248](https://doi.org/10.1016/j.neubiorev.2023.105248)] [Medline: [37247829](https://pubmed.ncbi.nlm.nih.gov/37247829/)]
20. Henderson A, Korner-Bitensky N, Levin M. Virtual reality in stroke rehabilitation: a systematic review of its effectiveness for upper limb motor recovery. *Top Stroke Rehabil* 2007;14(2):52-61. [doi: [10.1310/tsr1402-52](https://doi.org/10.1310/tsr1402-52)] [Medline: [17517575](https://pubmed.ncbi.nlm.nih.gov/17517575/)]
21. Hao J, He Z, Yu X, Remis A. Comparison of immersive and non-immersive virtual reality for upper extremity functional recovery in patients with stroke: a systematic review and network meta-analysis. *Neurol Sci* 2023 Aug;44(8):2679-2697. [doi: [10.1007/s10072-023-06742-8](https://doi.org/10.1007/s10072-023-06742-8)] [Medline: [36959332](https://pubmed.ncbi.nlm.nih.gov/36959332/)]
22. Lockwood C, Dos Santos KB, Pap R. Practical guidance for knowledge synthesis: scoping review methods. *Asian Nurs Res (Korean Soc Nurs Sci)* 2019 Dec;13(5):287-294. [doi: [10.1016/j.anr.2019.11.002](https://doi.org/10.1016/j.anr.2019.11.002)] [Medline: [31756513](https://pubmed.ncbi.nlm.nih.gov/31756513/)]
23. Bamodu O, Ye XM. Virtual reality and virtual reality system components. *AMR* 2013;765-767:1169-1172 [FREE Full text] [doi: [10.4028/www.scientific.net/AMR.765-767.1169](https://doi.org/10.4028/www.scientific.net/AMR.765-767.1169)]
24. Ventura S, Brivio E, Riva G, Baños RM. Immersive versus non-immersive experience: exploring the feasibility of memory assessment through 360° technology. *Front Psychol* 2019;10:2509. [doi: [10.3389/fpsyg.2019.02509](https://doi.org/10.3389/fpsyg.2019.02509)] [Medline: [31798492](https://pubmed.ncbi.nlm.nih.gov/31798492/)]
25. Moro SB, Carrieri M, Avola D, et al. A novel semi-immersive virtual reality visuo-motor task activates ventrolateral prefrontal cortex: a functional near-infrared spectroscopy study. *J Neural Eng* 2016 Jun;13(3):036002. [doi: [10.1088/1741-2560/13/3/036002](https://doi.org/10.1088/1741-2560/13/3/036002)] [Medline: [27001948](https://pubmed.ncbi.nlm.nih.gov/27001948/)]
26. Lim DY, Hwang DM, Cho KH, Moon CW, Ahn SY. A fully immersive virtual reality method for upper limb rehabilitation in spinal cord injury. *Ann Rehabil Med* 2020 Aug;44(4):311-319. [doi: [10.5535/arm.19181](https://doi.org/10.5535/arm.19181)] [Medline: [32721988](https://pubmed.ncbi.nlm.nih.gov/32721988/)]
27. Maier M, Ballester BR, Leiva Bañuelos N, Duarte Oller E, Verschure P. Adaptive conjunctive cognitive training (ACCT) in virtual reality for chronic stroke patients: a randomized controlled pilot trial. *J Neuroeng Rehabil* 2020 Mar 6;17(1):42. [doi: [10.1186/s12984-020-0652-3](https://doi.org/10.1186/s12984-020-0652-3)] [Medline: [32143674](https://pubmed.ncbi.nlm.nih.gov/32143674/)]
28. Lam SSL, Liu TW, Ng SSM, Lai CWK, Woo J. Bilateral movement-based computer games improve sensorimotor functions in subacute stroke survivors. *J Rehabil Med* 2022 Jul 25;54:jrm00307. [doi: [10.2340/jrm.v54.913](https://doi.org/10.2340/jrm.v54.913)] [Medline: [35730901](https://pubmed.ncbi.nlm.nih.gov/35730901/)]
29. Rodríguez-Hernández M, Polonio-López B, Corregidor-Sánchez AI, Martín-Conty JL, Mohedano-Moriano A, Criado-Álvarez JJ. Can specific virtual reality combined with conventional rehabilitation improve poststroke hand motor function? A randomized clinical trial. *J Neuroeng Rehabil* 2023 Apr 4;20(1):38. [doi: [10.1186/s12984-023-01170-3](https://doi.org/10.1186/s12984-023-01170-3)] [Medline: [37016408](https://pubmed.ncbi.nlm.nih.gov/37016408/)]
30. Türkbey TA, Kutlay S, Gök H. Clinical feasibility of Xbox Kinect™ training for stroke rehabilitation: a single-blind randomized controlled pilot study. *J Rehabil Med* 2017 Jan 19;49(1):22-29. [doi: [10.2340/16501977-2183](https://doi.org/10.2340/16501977-2183)] [Medline: [27973678](https://pubmed.ncbi.nlm.nih.gov/27973678/)]
31. Marques-Sule E, Arnal-Gómez A, Buitrago-Jiménez G, Suso-Martí L, Cuenca-Martínez F, Espí-López GV. Effectiveness of nintendo wii and physical therapy in functionality, balance, and daily activities in chronic stroke patients. *J Am Med Dir Assoc* 2021 May;22(5):1073-1080. [doi: [10.1016/j.jamda.2021.01.076](https://doi.org/10.1016/j.jamda.2021.01.076)] [Medline: [33639116](https://pubmed.ncbi.nlm.nih.gov/33639116/)]
32. Saposnik G, Cohen LG, Mamdani M, et al. Efficacy and safety of non-immersive virtual reality exercising in stroke rehabilitation (EVREST): a randomised, multicentre, single-blind, controlled trial. *Lancet Neurol* 2016 Sep;15(10):1019-1027. [doi: [10.1016/S1474-4422\(16\)30121-1](https://doi.org/10.1016/S1474-4422(16)30121-1)] [Medline: [27365261](https://pubmed.ncbi.nlm.nih.gov/27365261/)]

33. Wilson PH, Rogers JM, Vogel K, Steenbergen B, McGuckian TB, Duckworth J. Home-based (virtual) rehabilitation improves motor and cognitive function for stroke patients: a randomized controlled trial of the elements (EDNA-22) system. *J Neuroeng Rehabil* 2021 Nov 25;18(1):165. [doi: [10.1186/s12984-021-00956-7](https://doi.org/10.1186/s12984-021-00956-7)] [Medline: [34823545](https://pubmed.ncbi.nlm.nih.gov/34823545/)]
34. Chatterjee K, Buchanan A, Cottrell K, Hughes S, Day TW, John NW. Immersive virtual reality for the cognitive rehabilitation of stroke survivors. *IEEE Trans Neural Syst Rehabil Eng* 2022;30:719-728. [doi: [10.1109/TNSRE.2022.3158731](https://doi.org/10.1109/TNSRE.2022.3158731)] [Medline: [35271448](https://pubmed.ncbi.nlm.nih.gov/35271448/)]
35. Huang Q, Jiang X, Jin Y, et al. Immersive virtual reality-based rehabilitation for subacute stroke: a randomized controlled trial. *J Neurol* 2024 Mar;271(3):1256-1266. [doi: [10.1007/s00415-023-12060-y](https://doi.org/10.1007/s00415-023-12060-y)] [Medline: [37947856](https://pubmed.ncbi.nlm.nih.gov/37947856/)]
36. Laffont I, Froger J, Jourdan C, et al. Rehabilitation of the upper arm early after stroke: video games versus conventional rehabilitation. A randomized controlled trial. *Ann Phys Rehabil Med* 2020 May;63(3):173-180. [doi: [10.1016/j.rehab.2019.10.009](https://doi.org/10.1016/j.rehab.2019.10.009)] [Medline: [31830535](https://pubmed.ncbi.nlm.nih.gov/31830535/)]
37. Kayabinar B, Alemdaroğlu-Gürbüz İ, Yılmaz Ö. The effects of virtual reality augmented robot-assisted gait training on dual-task performance and functional measures in chronic stroke: a randomized controlled single-blind trial. *Eur J Phys Rehabil Med* 2021 Apr;57(2):227-237. [doi: [10.23736/S1973-9087.21.06441-8](https://doi.org/10.23736/S1973-9087.21.06441-8)] [Medline: [33541040](https://pubmed.ncbi.nlm.nih.gov/33541040/)]
38. Calabrò RS, Naro A, Russo M, et al. The role of virtual reality in improving motor performance as revealed by EEG: a randomized clinical trial. *J Neuroeng Rehabil* 2017 Jun 7;14(1):53. [doi: [10.1186/s12984-017-0268-4](https://doi.org/10.1186/s12984-017-0268-4)] [Medline: [28592282](https://pubmed.ncbi.nlm.nih.gov/28592282/)]
39. Brunner I, Skouen JS, Hofstad H, et al. Virtual reality training for upper extremity in subacute stroke (VIRTUES): a multicenter RCT. *Neurology (ECronicon)* 2017 Dec 12;89(24):2413-2421. [doi: [10.1212/WNL.0000000000004744](https://doi.org/10.1212/WNL.0000000000004744)] [Medline: [29142090](https://pubmed.ncbi.nlm.nih.gov/29142090/)]
40. Li H, Lao F, Gu X. Effect of virtual reality-based shoulder flexion training at specific angles on upper limb function recovery in stroke patients with hemiplegia. *Chin J Phys Med Rehabil* 2023;45(2):134-136. [doi: [10.3760/cma.j.issn.0254-1424.2023.02.009](https://doi.org/10.3760/cma.j.issn.0254-1424.2023.02.009)]
41. Liang S, Wu Y, Zhang R. Effect of virtual reality training on balance and gait in patients with cerebellar infarction and balance disorder. *Chin J Rehabil Med* 2020;35(6):700-704. [doi: [10.3969/j.issn.1001-1242.2020.06.010](https://doi.org/10.3969/j.issn.1001-1242.2020.06.010)]
42. Cucinella SL, de Winter JCF, Grauwmeijer E, Evers M, Marchal-Crespo L. Towards personalized immersive virtual reality neurorehabilitation: a human-centered design. *J Neuroeng Rehabil* 2025 Jan 20;22(1):7. [doi: [10.1186/s12984-024-01489-5](https://doi.org/10.1186/s12984-024-01489-5)] [Medline: [39833912](https://pubmed.ncbi.nlm.nih.gov/39833912/)]
43. Munoz-Novoa M, Kristoffersen MB, Sunnerhagen KS, Naber A, Ortiz-Catalan M, Alt Murphy M. Myoelectric pattern recognition with virtual reality and serious gaming improves upper limb function in chronic stroke: a single case experimental design study. *J Neuroeng Rehabil* 2025 Jan 17;22(1):6. [doi: [10.1186/s12984-025-01541-y](https://doi.org/10.1186/s12984-025-01541-y)] [Medline: [39825410](https://pubmed.ncbi.nlm.nih.gov/39825410/)]
44. Shannon MM, Lipson-Smith R, Elf M, et al. Exploration of hospital room design on emotions, activity, and social connections after stroke: a qualitative study of patient experiences using virtual reality. *Disabil Rehabil* 2025 Aug;47(17):4496-4505. [doi: [10.1080/09638288.2025.2451771](https://doi.org/10.1080/09638288.2025.2451771)] [Medline: [39840937](https://pubmed.ncbi.nlm.nih.gov/39840937/)]
45. Contrada M, Pignolo L, Vatrano M, et al. Application of multidomain cognitive training in a tele-neurorehabilitation setting for treatment of post-stroke cognitive disorders. *Brain Sci* 2024 Dec 26;15(1):11. [doi: [10.3390/brainsci15010011](https://doi.org/10.3390/brainsci15010011)] [Medline: [39851379](https://pubmed.ncbi.nlm.nih.gov/39851379/)]

Abbreviations

FGA: Functional Gait Assessment

FIVR: fully immersive virtual reality

HADS: Hospital Anxiety and Depression Scale

HMD: head-mounted display

MeSH: Medical Subject Headings

MoCA: Montreal Cognitive Assessment

NIVR: nonimmersive virtual reality

PCC: Population, Concept, and Context

PRISMA : Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

PROSPERO: International Prospective Register of Systematic Reviews

RCT: randomized controlled trial

SF-36: Short Form Health Survey-36 item

SIS: Stroke Impact Scale

VR: virtual reality

WHO: World Health Organization

Edited by A Coristine; submitted 20.Feb.2025; peer-reviewed by J Hao, T Vagg; revised version received 23.Jan.2026; accepted 23.Jan.2026; published 08.Apr.2026.

Please cite as:

Liu R, Lin B, Zhai Y, Wang W

Navigating Virtual Reality in Stroke Rehabilitation: Scoping Review of Diverse Intervention Effects

JMIR Serious Games 2026;14:e72498

URL: <https://games.jmir.org/2026/1/e72498>

doi: [10.2196/72498](https://doi.org/10.2196/72498)

© Ruixue Liu, Beibei Lin, Yanhuan Zhai, Wei Wang. Originally published in JMIR Serious Games (<https://games.jmir.org>), 8.Apr.2026. 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.

Effectiveness of Gamified Swallowing Exercises in Adults With Dysphagia: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Jijia Liu^{1,2}, BSN; Meijun Ou¹, BSN, MSN; Mengyao Yan^{1,2}, BSN, MSN; Xianghua Xu¹, BSN, MSN, PhD; Li Qin³, BSN

¹The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University/Hunan Cancer Hospital, 283 Tongzipo Road, Yuelu District, Changsha, China

²Xiangya School of Nursing, Central South University, Changsha, China

³The 921st Hospital of Chinese People's Liberation Army (Second Affiliated Hospital of Hunan Normal University), Changsha, China

Corresponding Author:

Meijun Ou, BSN, MSN

The Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University/Hunan Cancer Hospital, 283 Tongzipo Road, Yuelu District, Changsha, China

Abstract

Background: Dysphagia is a prevalent health issue affecting quality of life. Gamified swallowing exercises have the potential to enhance swallowing function and adherence in adults with dysphagia. Nevertheless, such evidence has not yet been systematically synthesized.

Objective: This study aimed to systematically evaluate the effects of gamified swallowing exercises and promote their clinical application in swallowing rehabilitation.

Methods: Eleven electronic databases were searched from inception to June 25, 2025. Randomized controlled trials (RCTs) that explored gamified swallowing exercises in adults with dysphagia, regardless of etiology, and reported on swallowing function, adherence, nutritional status, or quality of life were included. The Cochrane Risk of Bias tool 2.0 was applied to assess the methodological quality. We employed the Hartung-Knapp-Sidik-Jonkman method for random-effects model analysis to calculate effect sizes and 95% CIs. Subgroup analysis was conducted to explore potential causes of heterogeneity, and the overall quality of evidence was evaluated through GRADEpro GDT software.

Results: Of 2400 initially identified records, 6 RCTs met the inclusion criteria, with 2 rated as low risk, 3 as some concerns, and 1 as high risk. These studies, conducted in China, South Korea, and Turkey between 2019 and 2025, enrolled a total of 330 poststroke patients with dysphagia. Interventions involved 1 to 3 games targeting the lips, tongue, and pharynx, with additional equipment. This meta-analysis indicated that gamified swallowing exercises improved swallowing function, as evidenced by enhanced swallowing performance (mean difference [MD]=1.1; 95% CI 0.9 to 1.3; $I^2=0$; prediction interval [PI] 0.3 to 1.6), reduced dysphagia severity (standardized mean difference [SMD]=0.4; 95% CI 0.3 to 0.5; $I^2=0$; PI 0.3 to 0.5), and increased adherence (MD=2.4; 95% CI 1.8 to 2.9; $I^2=0$; PI 1.2 to 2.7). However, no significant effects were observed on dysphagia screening (SMD=-1.8, 95% CI -5.4 to 1.8; $I^2=96.5%$; PI -8.6 to 4.9) or quality of life (SMD=-2.30, 95% CI -9.6 to 5; $I^2=97.5%$; PI -16.1 to 14.1). Overall, the quality of the included studies ranged from moderate to very low, which limits confidence in the pooled estimates.

Conclusions: This study provides the first quantitative synthesis of the effects of gamified swallowing exercises. The results demonstrate that gamified swallowing exercises improve swallowing function and enhance adherence, suggesting that gamified swallowing exercises are a promising method for swallowing rehabilitation. However, the overall risk of bias across included studies and the suboptimal evidence quality reduce the certainty of the current evidence. Therefore, the findings should be interpreted with caution. The small number of high-quality RCTs, reliance on additional equipment, and limited standardization of gamified interventions further limit clinical generalizability. Future research should focus on large-scale, robustly designed RCTs, while designing etiology-specific games and developing more accessible rehabilitation equipment, such as smartphones, could enhance the overall effects and facilitate the widespread adoption of these interventions in remote and home care settings.

Trial Registration: PROSPERO CRD42024617169; <https://www.crd.york.ac.uk/PROSPERO/view/CRD42024617169>

(*JMIR Serious Games* 2026;14:e82017) doi:[10.2196/82017](https://doi.org/10.2196/82017)

KEYWORDS

deglutition disorders; gamified; swallowing exercise; systematic review; meta-analysis

Introduction

Dysphagia is a swallowing disorder where foods or liquids fail to pass normally through the throat, often resulting in penetration or aspiration of food contents into the airway. In a survey of 4000 participants in the general population, more than 36% of respondents indicated dysphagia [1]. Furthermore, a systematic review reported a global dysphagia prevalence of 43.8% across diverse populations, noting an upward trend over time [2]. Dysphagia is common in high-risk groups, such as older adults and people with stroke, head and neck cancer, and neurodegenerative disease. An epidemiological survey conducted in China involving 5943 individuals indicated that the prevalence of dysphagia among healthy older adults, patients with stroke, head and neck cancer, and neurodegenerative disease was 19.2%, 51.1%, 34.4%, and 48.3%, respectively [3].

Dysphagia is associated with several complications, including an increased risk of malnutrition, aspiration, and pneumonia, leading to elevated medical expenses, prolonged hospitalizations, and greater mortality rates [3-5]. In addition, dysphagia has a direct negative effect on cognitive and psychosocial function. Several studies indicate that swallowing disorders are directly correlated with mild cognitive impairment, depression, anxiety, social isolation, and a reduced quality of life [6-9].

It is well-documented that swallowing exercises can improve the swallowing function, nutritional status, and quality of life of participants with swallowing disorders [10-12]. Generally, swallowing exercises require long-term persistence to achieve the expected results [13,14]. In addition, conventional swallowing exercises are mechanically repetitive and do not provide real-time feedback, resulting in a gradual decline in adherence [15,16]. Therefore, improving patient adherence is a key priority in swallowing rehabilitation.

Gamification involves applying game attributes in a nongaming context. In clinical rehabilitation, gamification utilizes game-like elements—such as point scoring, challenges, real-time feedback, and progression levels—to promote patient motivation, involvement, and adherence to therapy. In the context of dysphagia rehabilitation, these elements are embedded within exercises to make repetitive tasks more interactive and rewarding, potentially improving neuroplasticity and functional outcomes. For example, gamified swallowing exercises, coordinated with surface electromyography biofeedback (sEMG-BF), exhibited a greater improvement in the swallowing function of stroke patients than routine swallowing exercises or sEMG-BF alone [17]. Furthermore, the Park study demonstrated that gamified swallowing exercises could enhance training motivation, interest, enjoyment, and adherence in stroke patients [18].

Some researchers have applied gamified swallowing exercises in patients with swallowing disorders; however, the intervention effect of gamified swallowing exercises remains uncertain. For example, some studies have reported that gamified swallowing

training can improve swallowing function to a greater extent than conventional swallowing training. In contrast, others have shown that its effects are similar to those of conventional swallowing training [17-19]. Although gamified swallowing exercises have gained attention, their effectiveness remains inconsistently reported, particularly regarding participant adherence, nutritional outcomes, and swallowing function. This inconsistency highlights a critical gap in the evidence base that this review seeks to address. Therefore, in this study, we aimed to evaluate how gamified swallowing exercises affect swallowing function, adherence, nutritional status, and quality of life in patients with dysphagia and advocate for their incorporation into clinical practice in swallowing rehabilitation.

Methods

This systematic review and meta-analysis adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement [20], as seen in . This study was registered in PROSPERO (International Prospective Register of Systematic Reviews) (CRD42024617169), and the protocol has been published [21]. Any deviations are detailed in [Multimedia Appendix 1](#).

Search Strategy

Seven English-language databases (PubMed, Web of Science, Embase, CINAHL, Cochrane Library, JBI, and Scopus) and 3 Chinese-language databases (China National Knowledge Infrastructure, SinoMed, and Wanfang) were systematically searched from inception to June 25, 2025. We conducted a repeated search on February 8, 2026, and found no more eligible literature. Google Scholar was retrieved as a gray literature to minimize publication bias. We retrieved the CENTRAL (Cochrane Central Register of Controlled Trials) database through the Cochrane Library, including resources from trial registries such as ClinicalTrials.gov and ChiCTR. No published search filters were used. The search terms “dysphagia,” “game,” and “swallowing exercise” were combined in each database using both free-text words and Medical Subject Heading terms, if available. We originally developed a scientific and rigorous search strategy based on our research question, under the guidance of a professor specializing in medical information retrieval. No previously published search strategies were adapted or reused in this review. The snowball method was used to retrieve references from eligible studies to identify potential literature and avoid missing studies. No language or other restrictions were applied to the search. The search strategy adhered to the PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Search Extension) checklist to ensure transparency [22] as seen in [Checklist 2](#) . [Multimedia Appendix 2](#) provides the specific search strategies for each database.

Eligibility Criteria

The PICOS (population, intervention, comparison, outcomes, and study design) framework established the inclusion and

exclusion criteria for the systematic review of randomized controlled trials (RCTs). The inclusion criteria were as follows: (1) population: patients with dysphagia aged 18 years or older, irrespective of the origin of the swallowing impairment; (2) intervention of interest: gamified swallowing exercises, defined as swallowing exercises carried out in the form of games, including video games, virtual reality games, etc; (3) comparison: individuals in the comparison group were provided with conventional swallowing exercises or nongamified exercises; (4) outcomes: swallowing function (including swallowing performance, dysphagia severity, and dysphagia screening), adherence, nutritional status, or quality of life; and (5) study design: RCTs (including crossover, cluster, and pilot studies). The exclusion criteria were as follows: (1) editorial articles, conference abstracts, study protocols, comments, and letters and (2) studies with essential data that remained unavailable if data conversion was not possible and the corresponding authors did not respond to our requests.

Study Selection and Data Extraction

All retrieved studies were imported into EndNote 21 software for automatic and manual deduplication. After the removal of duplicates, 2 independent reviewers (JL and MY) screened the titles and abstracts before proceeding to assess the full texts against the established eligibility criteria. Disagreements during the study selection process were addressed through discussions within the research team.

Two reviewers (JL and MY) performed data extraction, with accuracy verification by a third researcher (MO). Information on the eligible studies was extracted, including study characteristics (authors, publication year, and country), participant characteristics (age and sample size), interventions (components, dose, timing, frequency, and length), controls, and outcomes. For studies with more than 1 intervention group where gamification was used in only 1 of the intervention groups, a comparison was conducted between the gamified group and the other 2 groups. If the authors did not report the mean and SD, we contacted the corresponding authors via email to request raw datasets. If unavailable, we converted them to the mean and SD based on the median and IQR [23,24]. Otherwise, the data were classified as missing data.

Quality Appraisal

The revised Cochrane Risk of Bias Tool, version 2.0, was used to assess the methodological quality of the RCTs [25]. Five domains were assessed: randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result. The evaluation results were divided into low-risk, some concerns, and high-risk groups. Additionally, the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) framework was used to grade the certainty of evidence for all outcomes [26]. The certainty was divided into high, moderate, low, and very low based on the risk of bias, imprecision, inconsistency, indirectness, and publication bias [27]. The overall quality of evidence was evaluated using GRADEpro GDT software. Two reviewers (JL and MY) independently assessed the quality of the included studies. Disagreements were

addressed through discussion and/or consultation with a third reviewer (MO) to achieve consensus.

Data Statistics

Data were analyzed using R software (version 4.4.2). According to the Cochrane Handbook for Systematic Reviews of Interventions [28], for the 3-arm RCTs, we combined the 2 nongamified groups into 1 group, and pooled calculations were performed using the following formula:

$$N=N1+N2$$

$$M=N1M1+N2M2N1+N2$$

$$SD=(N1-1)SD1^2+(N2-1)SD2^2+N1N2(N1+N2)M1^2+M2^2-2M1M2N1+N2-1$$

A meta-analysis was conducted on outcomes measured in 2 or more studies. For continuous variables measured using the same scale, we estimated pooled effect sizes using the pooled mean difference (MD); otherwise, the pooled standardized mean difference (SMD) was computed as the pooled effect size. For studies that used multiple instruments to assess the same outcome, we established a priori “scale prioritization” rules based on the psychometric properties of each instrument (such as Gugging Swallowing Screen > Standardized Swallowing Assessment > Water Swallowing Test) [29], and extracted data accordingly for effect size synthesis. We employed the Hartung-Knapp-Sidik-Jonkman method for random-effects model analysis to calculate the pooled estimate and its corresponding 95% CI, addressing the issue of limited study inclusion and high heterogeneity between studies [30]. The heterogeneity of outcomes was tested using Cochran Q (χ^2 test) and I^2 statistic [31]. In detail, $P<.05$ was considered to indicate heterogeneity.

To explore potential sources of heterogeneity, subgroup analyses were conducted based on different assessment tools, kinds of games, different additional equipment, and disease duration. The Instrument for the Credibility of Effect Modification Analyses (ICEMAN) was applied to evaluate the credibility of subgroup effects [32] when the interaction $P<.1$. The credibility of subgroup effects was categorized as “very low,” “low,” “moderate,” or “high” based on 8 core evaluation items. An interaction $P\leq.005$ was used as the threshold to minimize the likelihood of chance findings. Results for subgroups containing only a few studies were interpreted with caution.

Upon detecting substantial heterogeneity, 2 types of sensitivity analyses were conducted to evaluate the robustness of the findings as follows: (1) a leave-one-out method, where each study was sequentially excluded to assess changes in heterogeneity and result stability and (2) an alternative measurement instrument analysis, in which alternative measurement results were used for studies that assessed the same outcome using 2 different instruments. Funnel plots were used to assess small-study effects [33]. In addition, when the number of included studies was small, we calculated the 95% prediction interval using the Nagashima correction to illustrate the practical implications of heterogeneity and present the expected range of true effects across studies [34].

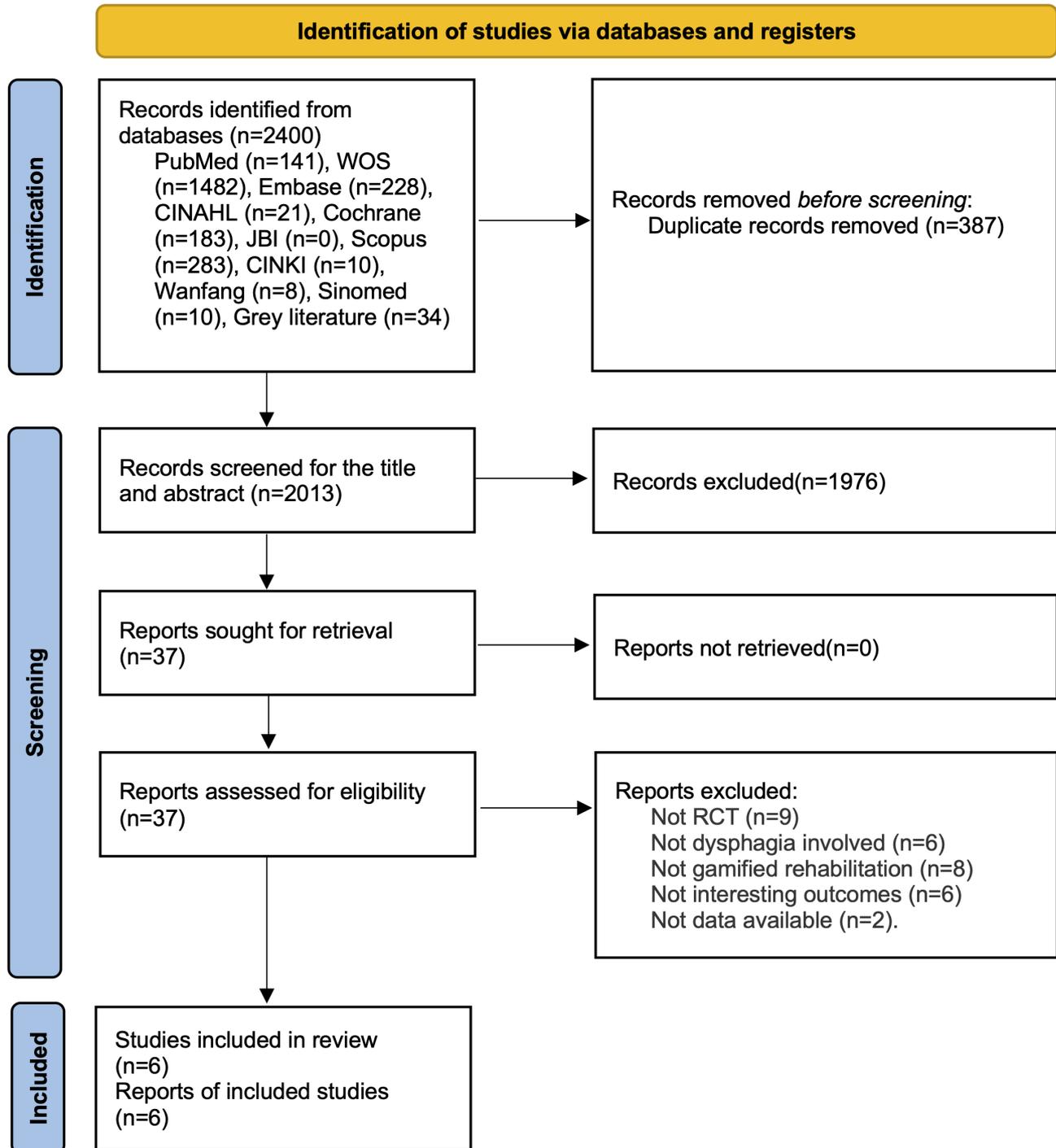
Results

Study Selection

The initial search identified 2400 publications, including 2366 records from 9 databases (the JBI database yielded no publications) and 34 from the gray literature database. After removing 387 duplicate articles, the titles and abstracts were

scanned, and 1976 articles were excluded owing to irrelevant topics. After reviewing the full text, 31 articles were excluded because of their irrelevant study design, participants, interventions, outcomes, or unavailability of data. Ultimately, 6 studies [17-19,35-37] were included in this review. A flow diagram is presented in Figure 1, and the details of excluded studies are shown in Multimedia Appendix 3.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for study selection. RCT: randomized controlled trials.



Study Characteristics

Table 1 summarizes the characteristics of the 6 studies included in this review. All studies were published between 2019 and

2025. Four studies [17,19,36,37] were conducted in China, 1 [18] in Korea, and 1 [35] in Turkey. Five studies [18,19,35-37] were 2-arm RCTs, and 1 [17] was a 3-arm RCT. In total, the studies involved 330 participants with poststroke dysphagia.

The disease duration ranged from 8 days (approximately 0.27 mo) to 5.18 months.

Table . Characteristics of included studies.

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
Zhang (2025b) [36], China	Sample size: 42/42, age: 64.98 (SD 9.66)/66.43 (SD 13.12) y, male: 24/29, female: 18/13	Hemorrhagic stroke: 7/8, ischemic stroke: 35/34, disease duration: 1.67 (SD 1.48)/1.76 (SD 1.53) mo	AI ^b -based video game intervention: 1. Lip exercise (“Collecting Carrots”): participants control the in-game character to move left and right by puffing out left or right cheeks (2 - 3 s each time, 15 repetitions per cheek). When the game character reaches the target, carrots automatically drop into the backpack.) 2. Tongue exercise (“Out of the Maze”): participants guide the game character through a maze by extending the tongue upwards, downwards, left, or right. Each tongue movement is repeated 15 times.	AI-based video game system and screen with an overhead camera	30 min per day, 5 times a week, for a total of 4 weeks	Usual care treatment, including lip exercises, tongue exercises, and CTAR	GUSS ^d , SSA ^e , FOIS ^f , MNA-SF ^g , SWAL-QOL ^h , and adherence	Pre/post, 4 wk follow-up

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
			<p>The character exits the maze once the routine is completed.)</p> <p>3. CTAR^c exercise (“Leaping Barriers”: participants press a rubber ball with their chin, hold for 2 or 3 s, and then release, to make the bird descend and navigate around 15 obstacles in total.)</p>					
Park (2019) [18], South Korea	Sample size: 20/20, age: 60.95 (SD 11.19)/59.45 (SD 9.34) y, male:13/10, female:7/10 ⁱ	Hemorrhagic stroke: 12/14, ischemic stroke: 8/6, disease duration: 3.60 (SD 1.19)/3.85 (SD 1.18) mo	Game-based CTAR ^c : tucks the chin downward against a resistance bar to reach the target value displayed on the screen; 30 consecutive repetitions.	Tablet PC screen, a Bluetooth connector, or a resilient resistance bar	5 times a week, 4 wk		Head-lift exercise in the supine position FOIS ^f , VDS ^j , and PAS ^k	Pre/post

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
Hou (2024) [17], China	Sample size: 30/60, age: 63.90 (SD 5.22)/63.88 (SD 6) y, male: 23/46, female: 7/14	Hemorrhagic stroke: 2/7, ischemic stroke: 28/53, disease duration: 8.60 (SD 3.50)/8.09 (SD 3.84) d	Game-based Mendelsohn maneuver ("Rabbit mountain-climbing game": patient imagines themselves as a rabbit climbing a mountain. During the climbing process, they are instructed to perform the Mendelsohn maneuver and maintain it for a few seconds. After reaching the top of the mountain, the patient immediately exhales and relaxes, and the rabbit lies down and eats a carrot.) + usual care + tDCS ^l + surface EMG-BF ^m	American VitalStim Plus 5923 - 3 Electrotherapy System, computer; EMG-BF	20 min each time, once daily, 7 - 14 d	Group 1: usual care + tDCS ^l , group 2: usual care + tDCS ^l + sEMG-BF ^m	FOIS ^f , WST ⁿ , amplitude of the submental muscle group based on sEMG ^j , and tongue pressure	Pre/post

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
Zhang (2025a) [19], China	Sample size: 14/12, age: 66.50 (SD 8.25)/66.20 (SD 14.99) y, male: 8/10, female: 6/2	Hemorrhagic stroke: 3/2, ischemic: 11/10, disease duration: 1.91 (SD 0.85)/2.19 (SD 2.21) mo		Face recognition-driven video game system, computer screen, and camera (recognize and record the participant's facial movements)	30 min per day, 5 d per week, for 4 wk	Conventional therapy: exercises are identical to the intervention group, except without the FR-VG system	GUSS ^d , SSA ^e , FOIS ^f , SWAL-QOL ^h , VVST ^o , and adherence	Pre/post

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
			<p>Face recognition-driven video game swallowing training program:</p> <ol style="list-style-type: none"> 1. Lip exercises ("Carrot gathering": participants control an in-game character to move left and right by alternately puffing out their left and right cheeks, maintaining each movement for 2 - 3 s, 15 repetitions for each cheek. As the game character reaches designated spots, carrots automatically drop into the backpack.) 2. Tongue exercises ("Labyrinth Navigation": participants guide a character through a maze by extending the tongue upwards, downwards, left, or 					

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
			right, with 15 repetitions for each direction required to complete the maze.) 3. Jaw exercises ("Soaring Bird": participants control a bird descending and navigating around 15 obstacles by lowering their jaw to press a rubber ball with their chin for 2 - 3 s, and then releasing.)					
Kang (2024) [37], China	Sample size: 30/30, age: 61.57 (SD 10.79)/63.40 (SD 10.26) y, male: 23/21, female: 7/9	Hemorrhagic stroke: 16/17, ischemic stroke: 14/13, disease duration: 32.03 (SD 8.72)/32.07 (SD 11.30) d	Virtual reality-based swallowing game: participant wears a head-mounted VR device, and when swallowing forcefully, the electromyographic signals of the sublingual muscles are transmitted to the VR device, triggering a VR scenario game. When the electromyographic signals exceed the set threshold, the virtual game character will successfully ingest food and earn points.	Pico 4 All-in-one VR Headset, computer; JP-001 sEMG Signals Acquisition System	30 min per day, 5 d per week, for 3 wk	Conventional therapy	SSA ^e , FOIS ^f , WST ⁿ , and surface electromyographic value	Pre/post

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
Alyanak (2025) [35], Turkey	Sample size: 16/17, age: 63.38 (SD 11.38)/60.59 (SD 12.54) y, male: 11/11, female: 5/6	Hemorrhagic stroke: 1/2, ischemic stroke: 15/15; disease duration: 5 (SD 2.84)/5.18 (SD 2.02) mo		Vitalstim Plus device, the Vitalstim Software, computer; EMG-BF	30 min per day, 5 d per week, 3 wk	Effortful swallow + Mendelsohn maneuver with only verbal feedback	FOIS ^f , PAS ^k , DOSS ^p , and DHI ^q	Pre/post

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
			<p>Game-based EMG-BF^m therapy</p> <ol style="list-style-type: none"> 1. Effortful swallow (“Rose game”): patients press their tongue hard against the palate and swallow with all their strength; a rose on the computer screen would fade when muscle activity exceeds a threshold, and then, the muscles relax, and the rose blooms again.) 2. Mendelsohn maneuver (“Rabbit game”): patients elevate the larynx to reach the highest point for 2 - 3 s (Mendelsohn maneuver) to make the rabbit climb a hill and reach the top with the carrot, and then relax for 2 s, the rab- 					

First author (y), country	Sample details (IG/CG ^a)	Diagnosis (IG/CG)	Intervention components (game name, details)	Equipment	Dose, frequency, and length of intervention	Control group	Measurements	Time points
			bit would descend the hill and eat the carrot.) Each game is 15 min.					

^aIG/CG: intervention group/control group.

^bAI: artificial intelligence.

^cCTAR: chin tuck against resistance.

^dGUSS: Gugging Swallowing Screen.

^eSSA: standardized swallowing assessment.

^fFOIS: Functional Oral Intake Scale.

^gMNA-SF: Mini Nutritional Assessment Short Form.

^hSWAL-QOL: Swallowing Quality-of-Life Questionnaire.

ⁱThe study initially presented characteristics for 20 participants per group. However, the analysis was ultimately conducted on 37 participants (19 in the intervention group and 18 in the control group). We attempted to clarify this discrepancy with the corresponding author but received no response. Therefore, we proceeded to use the data from the 37 participants in our meta-analysis.

^jVDS: Videofluoroscopic Dysphagia Scale.

^kPAS: Penetration-Aspiration Scale.

^ltDCS: transcranial direct current stimulation.

^msEMG-BF: surface electromyography biofeedback.

ⁿWST: Water Swallowing Test.

^oVVST: Volume-Viscosity Swallow Test.

^pDOSS: Dysphagia Outcome and Severity Scale.

^qDHI: Dysphagia Handicap Index.

Intervention Characteristics

Table 1 summarizes the intervention characteristics of the 6 studies included in this review. The content and form of the gamified swallowing training varied. Three studies [17,18,37] included only 1 game module, and 3 studies [19,35,36] designed various games targeting specific swallowing movements. In terms of exercise content, 6 studies included tongue exercises (n=2), chin tuck against resistance exercises (n=2), lip movements (n=2), effortful swallowing (n=2), the Mendelsohn maneuver (n=2), and jaw exercises (n=1) [17-19,35-37].

All 6 studies required additional equipment (eg, surface electromyography, resistance bars, and computers equipped with face recognition-driven technology or artificial intelligence technology for intervention delivery) to conduct the interventions.

The gamified swallowing exercises in all studies were conducted for a period of 1 to 4 weeks [17-19,35-37], with only 1 study conducting a 4-week follow-up [36].

Control Characteristics

Table 1 summarizes the characteristics of the control groups of the 6 included studies. Swallowing exercise protocols were identical between the control and intervention groups, except that gaming components were excluded in the controls. The

usual care or conventional therapy typically encompassed oral motor exercises, sensory stimulation techniques, compensatory strategies, and supportive care (eg, vital sign monitoring, routine airway humidification, and nutritional management) [17-19,35-37].

Outcomes

Swallowing function was the primary outcome in 6 studies [17-19,35-37]. Swallowing function was categorized into swallowing performance, dysphagia severity, and dysphagia screening based on scale items. Swallowing performance was primarily assessed through the Functional Oral Intake Scale (n=6) [17-19,35-37], which measures the patients' ability to eat orally. The Dysphagia Outcome and Severity Scale (n=1) [35] and videofluoroscopic Dysphagia Scale (n=1) [18] were used to measure dysphagia severity. The measurement tools of dysphagia screening included the Gugging Swallowing Screen (GUSS; n=2) [19,36], the Standardized Swallowing Assessment (SSA; n=3) [19,36,37], and the Water Swallowing Test (WST; n=2) [17,37]. Physiological measures included submental muscle amplitude, tongue pressure, and surface electromyographic value [17,37].

Risk of Bias

Figure 2 summarizes the risk of bias for the 6 included studies. Two reviewers independently assessed the methodological

quality, achieving 95% agreement ($\kappa=0.91$). Two studies had a low risk of bias, 3 raised some concerns, and 1 was high risk. Three studies were considered to have some concerns because they did not report allocation concealment. One study had a high risk of bias because the outcomes were assessed without

blinding, and the results, particularly those based on subjective measures, were therefore susceptible to influence by the outcome assessors. For a detailed assessment, please refer to [Multimedia Appendix 4](#).

Figure 2. Risk-of-bias assessment of all included studies [17-19,35-37].



Meta-Analysis

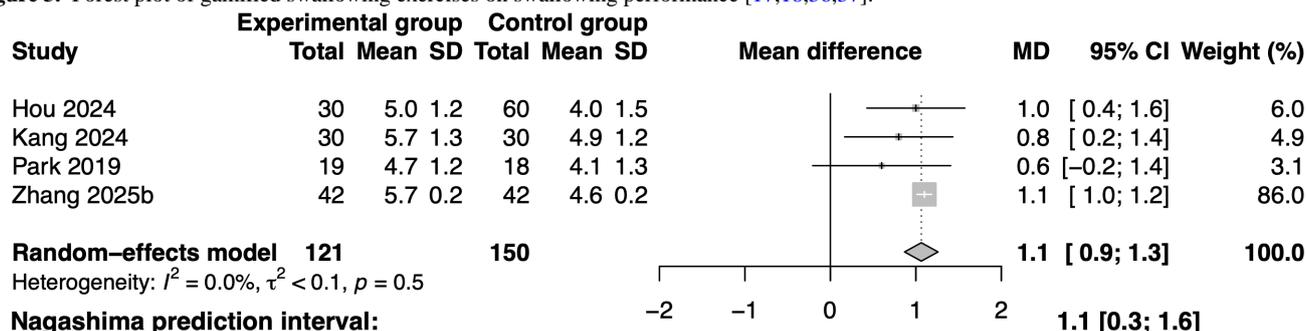
Effectiveness of Gamified Swallowing Exercises on Swallowing Function

Effectiveness of Gamified Swallowing Exercises on Swallowing Performance

Four studies involving 271 participants reported data on swallowing performance. The meta-analysis revealed a

significant positive effect of gamified swallowing exercises on swallowing performance (MD=1.1; 95% CI 0.9 - 1.3; $I^2=0$; Figure 3) compared with the control group.

Figure 3. Forest plot of gamified swallowing exercises on swallowing performance [17,18,36,37].

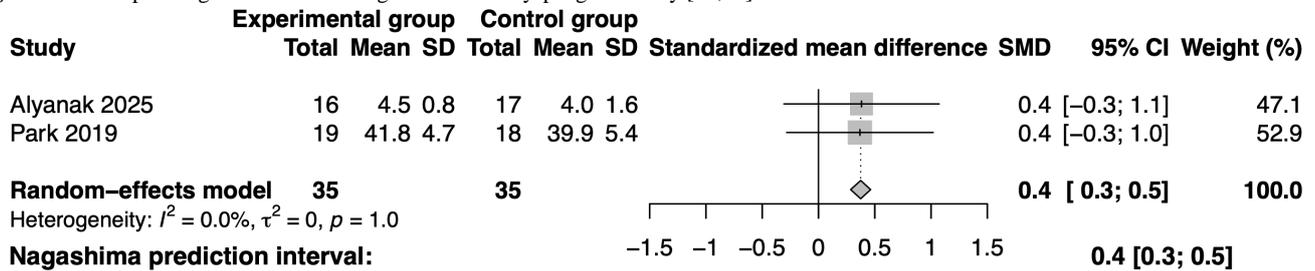


Effectiveness of Gamified Swallowing Exercises on Dysphagia Severity

A total of 2 studies involving 70 participants assessed dysphagia

severity. Compared with the control group, the level of dysphagia severity in the gamified swallowing exercises group was significantly reduced (SMD=0.4; 95% CI 0.3 - 0.5; $I^2=0$; Figure 4).

Figure 4. Forest plot of gamified swallowing exercises on dysphagia severity [18,35]. SMD: standardized mean difference.

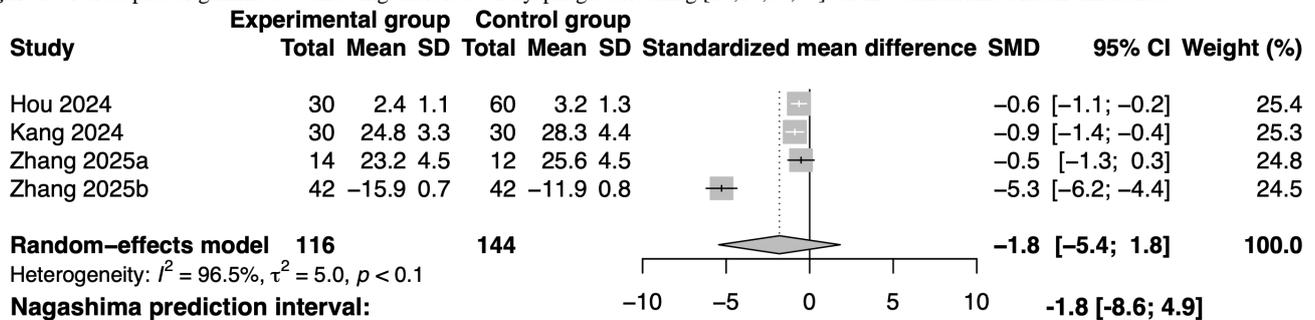


Effectiveness of Gamified Swallowing Exercises on Dysphagia Screening

The effectiveness of the gamified swallowing exercises on dysphagia screening was analyzed across 4 studies, including

260 participants. The meta-analysis demonstrated that the gamified swallowing exercises had no significant effect on dysphagia screening compared with the control group (SMD=-1.8; 95% CI -5.4 to 1.8; $I^2=96.5\%$; Figure 5).

Figure 5. Forest plot of gamified swallowing exercises on dysphagia screening [17,19,36,37]. SMD: standardized mean difference.

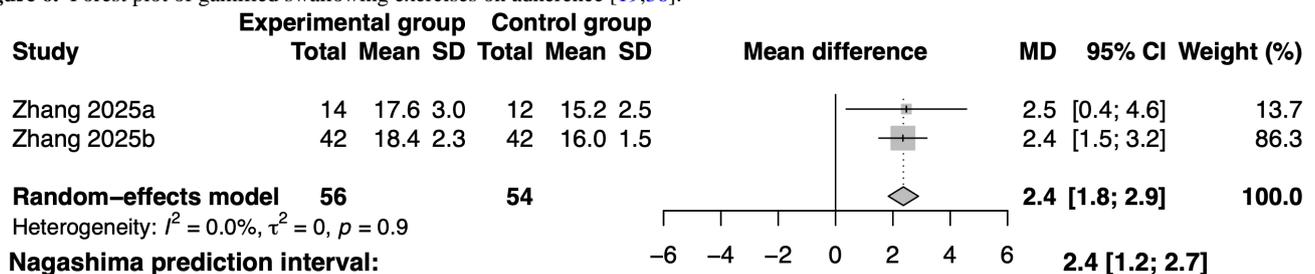


Effectiveness of Gamified Swallowing Exercises on Adherence

Regarding adherence, 2 studies involving 110 participants demonstrated that, compared with the control group, gamified

swallowing exercises were associated with increased adherence among adults with dysphagia (MD=2.4; 95% CI 1.8 - 2.9; $I^2=0$; Figure 6).

Figure 6. Forest plot of gamified swallowing exercises on adherence [19,36].

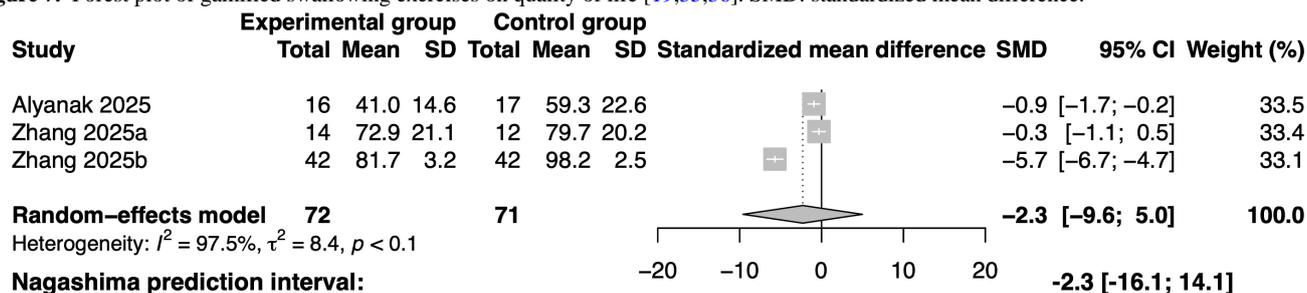


Effectiveness of Gamified Swallowing Exercises on Quality of Life

Quality of life was reported in 3 studies involving 143 participants. The results of the pooled analysis indicated that

gamified swallowing training did not significantly improve the quality of life in adults with dysphagia compared with the control group (SMD=-2.3; 95% CI -9.6 to 5; $I^2=97.5\%$; Figure 7).

Figure 7. Forest plot of gamified swallowing exercises on quality of life [19,35,36]. SMD: standardized mean difference.



Subgroup Analysis and Sensitivity Analysis

Given the preliminary findings of the meta-analysis, dysphagia screening was selected as the primary outcome for subgroup analysis. Subgroup analysis was conducted according to different measurement tools, kinds of games, additional equipment, and disease duration to explore the sources of heterogeneity. A significant between-subgroup difference was observed for the measurement tools (test for subgroup differences: $\chi^2=86.5$; 95% CI -5.4 to 1.8; $P<.001$). In contrast, no significant between-subgroup differences were found for kinds of games or additional equipment ($\chi^2=0.8$; 95% CI -5.4 to 1.8; $P=.40$) or disease duration ($\chi^2=1$; 95% CI -5.4 to 1.8; $P=.30$). The subgroup analysis showed that for patients with dysphagia after stroke, when using the SSA ($k=2$) to assess dysphagia screening, using a single game or using sEMG as the exercise equipment in dysphagia screening, the heterogeneity between studies significantly decreased ($I^2=0\%$), suggesting that such studies were relatively consistent in terms of design or implementation. However, the CI for the combined effect included the null value, indicating that the intervention effect was not statistically significant in this subgroup. Other subgroups, such as multiple games, other additional equipment (such as facial recognition devices), and disease duration ≥ 1 month, still had high heterogeneity ($>95\%$). The remaining 3 subgroups (WST, GUSS, and disease duration <1 mo) had only 1 study, resulting in overall very low credibility of the findings, which should be interpreted with caution. Since the P value for the interaction effect in the kinds of games and additional equipment subgroup was $>.1$, it did not meet the prerequisites for applying the ICEMAN tool; therefore, no further evaluation was conducted. Detailed results of the subgroup analysis are illustrated in [Multimedia Appendix 5](#).

To assess the impact of measurement tools on the robustness of the results, we performed an alternative measurement tool analysis, replacing the dysphagia screening scales used in Zhang study [36] and Kang study [37] with another scale. Specifically, we substituted GUSS with SSA in the Zhang study [36] and SSA with WST in the Kang study [37]. The direction of the pooled effect estimates remained unchanged after these substitutions. The leave-one-out sensitivity analysis indicated that the result of the swallowing performance was robust. For dysphagia screening, after excluding the Zhang study [36], the effect size of the remaining results changed (from ineffective to effective), and the heterogeneity decreased to 0, suggesting

that the robustness for dysphagia screening was limited. For quality of life, if any 1 study was excluded, the effect size of the remaining results remained unchanged; however, after excluding the Zhang study [36], the heterogeneity significantly decreased (from 97.5% to 21.8%). [Multimedia Appendix 6](#) presents the details of the sensitivity analysis.

However, due to the limited number of included studies, the estimation of the 95% prediction interval may be unstable, potentially leading to an overestimation of the true effects [38]. Therefore, the results of the prediction intervals should be interpreted with caution.

Small-Study Effect Analysis

Due to the limited number of studies included in each funnel plot ($n<8$), the observed symmetry (or asymmetry) lacked reliability. Consequently, no meaningful assessment of funnel plot symmetry could be performed. The relevant funnel plots are provided in [Multimedia Appendix 7](#).

Nutritional Status

Among the included studies, only 1 [36] examined the effectiveness of gamified swallowing exercises on nutritional status, which precluded the ability to conduct a meta-analysis on nutritional status. The intervention group exhibited significantly higher nutritional scores compared with the control group (MD 1, SD 0.5, 95% CI 0.1-1.9; $P=.03$). However, the 4-week follow-up revealed no statistical significance in nutritional status between groups. These findings suggest that individuals with dysphagia who received gamified swallowing exercises achieved significantly better short-term nutritional status than those who received conventional swallowing exercises; however, further investigation is required to determine the long-term effects.

Certainty of Evidence

Among the included outcomes, swallowing performance demonstrated moderate certainty of evidence; dysphagia severity and adherence had low certainty of evidence; and dysphagia screening, nutritional status, and quality of life showed very low certainty of evidence.

The certainty was downgraded primarily due to a high risk of bias, wide CIs, limited sample size, and statistical heterogeneity. Full GRADE (Grading of Recommendations Assessment, Development, and Evaluation) assessment details are available in [Table 2](#).

Table . The details of the certainty of evidence.

Certainty assessment	No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Twin mix	Local anesthetic solution	Relative effect (95% CI)	Absolute effect (95% CI)	Certainty
Swallowing performance	4	Randomized trials	Serious ^a	Not serious	Not serious	Not serious	none	121	150	— ^b	MD 1.1 higher (0.9 higher to 1.3 higher)	Moderate ^a
Dysphagia severity	2	Randomized trials	Serious ^a	Not serious	Not serious	Serious ^c	None	35	35	—	SMD 0.4 higher (0.3 higher to 0.5 higher)	Low ^{a,c}
Dysphagia screening	4	Randomized trials	Serious ^a	Serious ^d	Not serious	Serious ^e	None	116	114	—	SMD 1.8 lower (5.4 lower to 1.8 higher)	Very low ^{a,c,e}
Adherence	2	Randomized trials	Not serious	Not serious	Not serious	Serious ^f	None	56	54	—	MD 2.4 higher (1.8 higher to 2.9 higher)	Low ^f
Nutritional status	1	Randomized trials	Not serious ^g	Serious	Not serious	Serious ^c	None	42	42	—	MD 1 higher (1.9 lower to 0.1 lower)	Very low ^{c,g}
Quality of life	3	Randomized trials	Serious ^a	Serious ^d	Not serious	Serious ^c	None	72	71	—	SMD 2.3 lower (9.6 lower to 5.0 higher)	Very low ^{a,c,d}

^aDowngraded 1 level for serious risk of bias in the included studies.

^bNot applicable.

^cDowngraded 1 level for serious imprecision due to limited sample size.

^dDowngraded 1 level for serious inconsistency due to statistical heterogeneity.

^eDowngraded 1 level for serious imprecision due to wide CIs.

^fDowngraded 2 levels for serious imprecision due to limited sample size and data from only 2 studies.

^gDowngraded 1 level for serious inconsistency due to data from only 1 study, so it is impossible to judge the consistency.

Discussion

Principal Findings

Gamified swallowing exercise is an emerging and promising approach in the field of swallowing rehabilitation. To the best of our knowledge, this study is the first systematic review and meta-analysis to rigorously and comprehensively evaluate the effectiveness of gamified swallowing training. The findings suggest that gamified swallowing exercises may help improve swallowing performance, reduce dysphagia severity, and increase adherence among individuals with dysphagia. However, due to the limited number of studies and considerable heterogeneity, the effects of gamified swallowing training on dysphagia screening, nutritional status, and quality of life remain uncertain.

The meta-analysis indicated that gamified swallowing exercises can effectively enhance swallowing performance, reduce dysphagia severity, and improve adherence. Low heterogeneity and narrower prediction intervals also support higher consistency of the effects across different settings [39]. This positive outcome may be attributed to the integration of gamification elements into traditional swallowing exercises without altering the core components of the rehabilitation program. Compared to conventional approaches, the inherent mechanisms of gamification can stimulate participants to maintain their motivation, engagement, and interest, facilitating sustained participation in the process [40]. Moreover, immersive game-based interventions address key challenges in rehabilitation, such as monotony, participant discomfort, and resistance, which can foster positive attitudes toward swallowing

exercises and promote long-term adherence to beneficial health behaviors [41].

Gamified swallowing rehabilitation systems typically employ additional equipment, such as EMG [17,35,37] or facial recognition devices [19,36], to monitor participants' swallowing activity, including movements of the tongue, lips, jaw, and throat muscles. They then present the electrical activity of the relevant muscles or allow the participant to control a game character through these movements, providing real-time visual and auditory feedback. This equipment-assisted feedback mechanism helps participants promptly correct their training posture and reduces the workload of nurses and therapists [36]. Notably, the above equipment is specialized and can only be used in hospitals or rehabilitation institutions. Future development may focus on integrating this method with widely accessible mobile technologies, such as smartphones, to enable remote dysphagia management in community or home settings. As an innovative and scalable intervention model, gamified swallowing rehabilitation holds considerable promise for boosting engagement, promoting long-term recovery, and advancing the application of technology-assisted rehabilitation in real-world contexts.

Regarding dysphagia screening, this meta-analysis did not demonstrate a significant effect of gamified swallowing training, and wide prediction intervals indicate that the effect varies differently. During the subgroup analysis of dysphagia screening, the observation of zero heterogeneity in some subgroups points to consistent intervention design; however, the lack of significant therapeutic effect indicates that the intervention's mechanism itself may need refinement. For example, swallowing is a multistage collaborative process, and a single game typically focuses on muscle training in one stage, which may limit the overall effectiveness of swallowing training. The limited number of studies in each subgroup reduced the statistical power, and the observed differences may reflect random error rather than true subgroup effects. These findings are exploratory and should be interpreted with caution, pending further validation through high-quality studies. Substantial statistical heterogeneity was observed within some subgroups (ie, multiple games, disease duration ≥ 1 mo), which may be attributable to the inclusion of a pilot study [19] with a small sample size, resulting in high uncertainty around its effect estimates. Furthermore, one included study [19] was a pilot version of another included trial [19,36]; however, these 2 studies exhibited markedly different standard deviations, with no overlap in confidence intervals, which further contributed to the heterogeneity.

The sensitivity analysis showed that removing the study by Zhang et al [36] would affect the direction of the results for dysphagia screening, indicating that, in the context of a limited number of studies, a single large-sample study might disproportionately impact the combined effect, reflecting the statistical vulnerability of meta-analyses with small samples rather than indicating a true difference in intervention effects.

Furthermore, several studies [17,19,37] had methodological limitations, primarily the lack of allocation concealment and blinding of outcome measurement. The former may introduce

selection bias, whereas the latter can cause judgment bias in outcome assessments relying on subjective evaluations, which may lead to over- or under-estimation of effects and undermine the robustness of the combined results. These methodological flaws should be carefully considered when interpreting the research findings.

Overall, despite exploring multiple dimensions such as measurement tools, game quantity, additional equipment, and disease duration, the source of heterogeneity remains unclear. The findings of this study should be interpreted with caution, and future research should focus on large-scale RCTs while improving methodological quality and reporting standards to validate the intervention effects and clarify the sources of heterogeneity.

This meta-analysis revealed a potential positive impact of gamified swallowing exercises on reducing dysphagia severity. A major limitation, however, is the scarcity of available trials ($n=2$) and the small overall sample size ($n=70$). These constraints likely led to low statistical power, suggesting that the observed positive result should be considered preliminary and requires validation through larger, more robust studies. Furthermore, poststroke dysphagia often results from neurological impairment [42]. Due to the inclusion of participants who have had a stroke for >3 months, a longer intervention period may be required to achieve meaningful improvement [35].

Quality of life is a crucial long-term outcome indicator for assessing improvements in swallowing function [43]. The meta-analysis did not demonstrate a significant effect of gamified swallowing exercises on improving quality of life, and the wide prediction intervals indicate that the effects could vary widely, potentially benefiting or harming quality of life in different settings. Quality of life, a highly subjective and multidimensional outcome, is influenced by a range of factors, including psychological status, disease duration, and level of social support [44]. Furthermore, most included studies were constrained by small sample sizes and limited follow-up, with outcomes primarily assessed immediately post-intervention, which likely fails to capture the long-term efficacy of the swallowing exercises. Consequently, the current evidence is insufficient to draw definitive conclusions regarding the effect on quality of life, and these findings must be interpreted with caution. Further high-quality studies with larger samples and extended follow-up periods are warranted to validate these effects.

Limitations

Several limitations should be noted in this study. First, a total of 11 databases were systematically searched, and the search was updated before submission. Ultimately, only 6 RCTs involving 330 participants were included. The limited number of studies and small sample size may restrict the comprehensiveness and generalizability of the findings. Nevertheless, rigorous and standardized methods were applied to ensure the methodological quality and transparency of this review. Second, only 2 studies strictly implemented randomization and blinding, whereas the methodological quality

of the remaining 4 studies was suboptimal and requires improvement.

Implications for Future Practice and Research

This study indicated that gamified swallowing exercises have a positive impact on swallowing function and adherence in adults with dysphagia, making them a promising rehabilitation approach for clinical practice. When designing and developing games for swallowing training, it is necessary to consider the cause of dysphagia and the accessibility of rehabilitation. Game-based exercises can be precisely designed to align with key rehabilitation goals. Accordingly, oral swallowing disorders may be addressed through games that enhance the strength of the lip, tongue, and jaw muscles, while pharyngeal disorders may be targeted with games designed to strengthen the throat muscles. The 6 included studies all used specialized and complex equipment, such as EMG and facial recognition devices, and, considering the long-term rehabilitation needs of patients, decreasing equipment dependence for gamified swallowing rehabilitation as much as possible is necessary so that individuals with swallowing disorders can engage in gamified swallowing exercises at home.

The current evidence base is characterized by single-center designs and an inherent high risk of performance bias due to the inability to blind participants and therapists. Moreover, the evidence is markedly limited by the absence of long-term outcome data, with only 1 study reporting a 4-week follow-up, which would bring a significant concern given that swallowing rehabilitation is a protracted process. It is therefore imperative

that future research prioritizes large-scale, multicenter RCTs with extended follow-up periods to minimize confounding factors and provide higher-quality evidence on the effectiveness and long-term effects of gamified swallowing exercises, as well as to further explore the underlying effect mechanisms of gamified swallowing training. Furthermore, research should also be expanded to develop etiology-specific protocols, moving beyond the current narrow focus on poststroke dysphagia.

Conclusion

This is the first systematic review and meta-analysis to evaluate and synthesize the effects of gamified swallowing exercises among adults with dysphagia. The findings demonstrated potential benefits of gamified swallowing exercises in improving swallowing performance, reducing dysphagia severity, and enhancing adherence among adults with dysphagia, while its effects on dysphagia screening, nutritional status, and quality of life warrant additional research to establish conclusive evidence. Given the limited number and small sample size of included studies, as well as the generally low quality of evidence, these results should be interpreted with caution. This study provides some reference and guidance for healthcare providers in the field of swallowing rehabilitation, such as integrating gamification elements, designing and developing etiology-specific games, and simplifying rehabilitation equipment. In the field of swallowing rehabilitation, helping adults with swallowing difficulties overcome technical barriers and improving the feasibility and scalability of home-based gamified swallowing rehabilitation might be a promising future research direction.

Acknowledgments

We would like to thank Editage for the English-language editing. The authors declare that no generative artificial intelligence (GAI) was used in the research and writing process.

Funding

This study was supported by the Hunan Provincial Natural Science Foundation of China (grant 2025JJ80845) and the High-Level Talent Support Program of Hunan Cancer Hospital (grant 20250730-1002). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

All data for this systematic review and meta-analysis were obtained from published trials. No new primary data were generated.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Deviations from the protocol.

[\[DOCX File, 15 KB - games_v14i1e82017_app1.docx\]](#)

Multimedia Appendix 2

Search strategy for each electronic database.

[\[DOCX File, 24 KB - games_v14i1e82017_app2.docx\]](#)

Multimedia Appendix 3

The details of excluded studies.

[\[DOCX File, 19 KB - games_v14i1e82017_app3.docx \]](#)

Multimedia Appendix 4

Risk of bias summary of all included studies.

[\[DOCX File, 30 KB - games_v14i1e82017_app4.docx \]](#)

Multimedia Appendix 5

Subgroup analysis.

[\[DOCX File, 827 KB - games_v14i1e82017_app5.docx \]](#)

Multimedia Appendix 6

Sensitivity analysis.

[\[DOCX File, 748 KB - games_v14i1e82017_app6.docx \]](#)

Multimedia Appendix 7

Small-study effect analysis.

[\[DOCX File, 341 KB - games_v14i1e82017_app7.docx \]](#)

Checklist 1

PRISMA checklist.

[\[PDF File, 180 KB - games_v14i1e82017_app8.pdf \]](#)

Checklist 2

PRISMA-S checklist.

[\[PDF File, 94 KB - games_v14i1e82017_app9.pdf \]](#)

References

1. Cook J, Sapia M, Walker C, Pittaoulis M. Dysphagia prevalence in Brazil, UK, China, and Indonesia and dysphagic patient preferences. *Healthcare (Basel)* 2024 Sep 12;12(18):18. [doi: [10.3390/healthcare12181827](https://doi.org/10.3390/healthcare12181827)] [Medline: [39337168](https://pubmed.ncbi.nlm.nih.gov/39337168/)]
2. Rajati F, Ahmadi N, Naghibzadeh ZAS, Kazemini M. The global prevalence of oropharyngeal dysphagia in different populations: a systematic review and meta-analysis. *J Transl Med* 2022 Apr 11;20(1):175. [doi: [10.1186/s12967-022-03380-0](https://doi.org/10.1186/s12967-022-03380-0)] [Medline: [35410274](https://pubmed.ncbi.nlm.nih.gov/35410274/)]
3. Zhang M, Li C, Zhang F, et al. Prevalence of dysphagia in China: an epidemiological survey of 5943 participants. *Dysphagia* 2021 Jun;36(3):339-350. [doi: [10.1007/s00455-020-10138-7](https://doi.org/10.1007/s00455-020-10138-7)] [Medline: [32458145](https://pubmed.ncbi.nlm.nih.gov/32458145/)]
4. Wu Y, Chen M, Liu X, et al. Prevalence and risk factors of malnutrition in elderly hospitalized patients with post-stroke dysphagia. *Sci Rep* 2025;15(1). [doi: [10.1038/s41598-025-23448-3](https://doi.org/10.1038/s41598-025-23448-3)]
5. Song W, Wu M, Wang H, Pang R, Zhu L. Prevalence, risk factors, and outcomes of dysphagia after stroke: a systematic review and meta-analysis. *Front Neurol* 2024;15:1403610. [doi: [10.3389/fneur.2024.1403610](https://doi.org/10.3389/fneur.2024.1403610)] [Medline: [39087010](https://pubmed.ncbi.nlm.nih.gov/39087010/)]
6. Karisik A, Dejakum B, Moelgg K, et al. Association between dysphagia and symptoms of depression and anxiety after ischemic stroke. *Eur J Neurol* 2024 May;31(5):e16224. [doi: [10.1111/ene.16224](https://doi.org/10.1111/ene.16224)] [Medline: [38308469](https://pubmed.ncbi.nlm.nih.gov/38308469/)]
7. Maeda T, Kamide N, Ando M, et al. Association between dysphagia and social isolation in community-dwelling older people. *Eur Geriatr Med* 2025 Apr;16(2):605-613. [doi: [10.1007/s41999-024-01140-5](https://doi.org/10.1007/s41999-024-01140-5)] [Medline: [39755780](https://pubmed.ncbi.nlm.nih.gov/39755780/)]
8. Murariu MO, Boia ER, Sitaru AM, et al. Long-term quality of life and functional outcomes in patients with total laryngectomy. *Cancers (Basel)* 2025 Mar 17;17(6):1011. [doi: [10.3390/cancers17061011](https://doi.org/10.3390/cancers17061011)] [Medline: [40149346](https://pubmed.ncbi.nlm.nih.gov/40149346/)]
9. He M, Li C, Xiao L, et al. Dysphagia, cognitive function, and dysphagia-related quality of life among oral cancer patients: a cross-sectional survey. *Geriatr Nurs* 2026 Apr;69:103895. [doi: [10.1016/j.gerinurse.2026.103895](https://doi.org/10.1016/j.gerinurse.2026.103895)]
10. Ye J, Wu C, Chen J, et al. Effectiveness of nurse-delivered stepwise swallowing training on dysphagia in patients with Alzheimer's disease: a multi-center randomized controlled trial. *Int J Nurs Stud* 2024 Feb;150:104649. [doi: [10.1016/j.ijnurstu.2023.104649](https://doi.org/10.1016/j.ijnurstu.2023.104649)] [Medline: [38070229](https://pubmed.ncbi.nlm.nih.gov/38070229/)]
11. Rugaitienė M, Lesauskaitė V, Ulozienė I, Smičius L, Damulevičienė G. Impact of modified diet, swallowing exercises, and electrostimulation on quality of life of older patients suffering from oropharyngeal dysphagia. *Medicina (Kaunas)* 2024 Jun 21;60(7):1021. [doi: [10.3390/medicina60071021](https://doi.org/10.3390/medicina60071021)] [Medline: [39064450](https://pubmed.ncbi.nlm.nih.gov/39064450/)]
12. Lee CL, Banda KJ, Chu YH, et al. Efficacy of swallowing rehabilitative therapies for adults with dysphagia: a network meta-analysis of randomized controlled trials. *Geroscience* 2025 Apr;47(2):2047-2065. [doi: [10.1007/s11357-024-01389-5](https://doi.org/10.1007/s11357-024-01389-5)] [Medline: [39438392](https://pubmed.ncbi.nlm.nih.gov/39438392/)]
13. Hajdú SF, Wessel I, Dalton SO, Eskildsen SJ, Johansen C. Swallowing exercise during head and neck cancer treatment: results of a randomized trial. *Dysphagia* 2022 Aug;37(4):749-762. [doi: [10.1007/s00455-021-10320-5](https://doi.org/10.1007/s00455-021-10320-5)] [Medline: [34117531](https://pubmed.ncbi.nlm.nih.gov/34117531/)]

14. Starmer HM, McCarroll L, Goldsmith T, et al. One-year swallowing outcomes in a head and neck cancer cohort: the impact of adherence to swallowing exercises and feeding tube use. *Dysphagia* 2025 Dec 5. [doi: [10.1007/s00455-025-10918-z](https://doi.org/10.1007/s00455-025-10918-z)] [Medline: [41348337](https://pubmed.ncbi.nlm.nih.gov/41348337/)]
15. Vester S, Muhr A, Meier J, Süß C, Kummer P, Künzel J. Prehabilitation of dysphagia in the therapy of head and neck cancer- a systematic review of the literature and evidence evaluation. *Front Oncol* 2023;13:1273430. [doi: [10.3389/fonc.2023.1273430](https://doi.org/10.3389/fonc.2023.1273430)] [Medline: [38188284](https://pubmed.ncbi.nlm.nih.gov/38188284/)]
16. Baudelet M, Duprez F, Van den Steen L, et al. Increasing adherence to prophylactic swallowing exercises during head and neck radiotherapy: the multicenter, randomized controlled PRESTO-trial. *Dysphagia* 2023 Jun;38(3):886-895. [doi: [10.1007/s00455-022-10513-6](https://doi.org/10.1007/s00455-022-10513-6)] [Medline: [36121560](https://pubmed.ncbi.nlm.nih.gov/36121560/)]
17. Hou M, Zhao Y, Zhao L, Yuan X, Liu Z, Li H. Efficacy of game training combined with surface electromyography biofeedback on post-stroke dysphagia. *Geriatr Nurs* 2024;55:255-262. [doi: [10.1016/j.gerinurse.2023.11.019](https://doi.org/10.1016/j.gerinurse.2023.11.019)] [Medline: [38091711](https://pubmed.ncbi.nlm.nih.gov/38091711/)]
18. Park JS, Lee G, Jung YJ. Effects of game-based chin tuck against resistance exercise vs head-lift exercise in patients with dysphagia after stroke: an assessor-blind, randomized controlled trial. *J Rehabil Med* 2019 Oct 29;51(10):749-754. [doi: [10.2340/16501977-2603](https://doi.org/10.2340/16501977-2603)] [Medline: [31515567](https://pubmed.ncbi.nlm.nih.gov/31515567/)]
19. Zhang B, Wong KP, Liu M, et al. Face recognition-driven video game for dysphagia rehabilitation in stroke patients: a pilot randomized controlled trial. *Arch Phys Med Rehabil* 2025 Mar;106(3):342-350. [doi: [10.1016/j.apmr.2024.10.005](https://doi.org/10.1016/j.apmr.2024.10.005)] [Medline: [39491573](https://pubmed.ncbi.nlm.nih.gov/39491573/)]
20. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021 Mar 29;372:n71. [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
21. Liu J, Xu X, Ou M. Effectiveness of gamified swallowing exercise in adults with dysphagia: protocol for a systematic review and meta-analysis. *Syst Rev* 2026;15(1). [doi: [10.1186/s13643-025-03064-y](https://doi.org/10.1186/s13643-025-03064-y)] [Medline: [41501930](https://pubmed.ncbi.nlm.nih.gov/41501930/)]
22. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev* 2021 Jan 26;10(1):39. [doi: [10.1186/s13643-020-01542-z](https://doi.org/10.1186/s13643-020-01542-z)] [Medline: [33499930](https://pubmed.ncbi.nlm.nih.gov/33499930/)]
23. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Stat Methods Med Res* 2018 Jun;27(6):1785-1805. [doi: [10.1177/0962280216669183](https://doi.org/10.1177/0962280216669183)] [Medline: [27683581](https://pubmed.ncbi.nlm.nih.gov/27683581/)]
24. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014 Dec 19;14:135. [doi: [10.1186/1471-2288-14-135](https://doi.org/10.1186/1471-2288-14-135)] [Medline: [25524443](https://pubmed.ncbi.nlm.nih.gov/25524443/)]
25. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019 Aug 28;366:l4898. [doi: [10.1136/bmj.l4898](https://doi.org/10.1136/bmj.l4898)] [Medline: [31462531](https://pubmed.ncbi.nlm.nih.gov/31462531/)]
26. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008 Apr 26;336(7650):924-926. [doi: [10.1136/bmj.39489.470347.AD](https://doi.org/10.1136/bmj.39489.470347.AD)] [Medline: [18436948](https://pubmed.ncbi.nlm.nih.gov/18436948/)]
27. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011 Apr;64(4):401-406. [doi: [10.1016/j.jclinepi.2010.07.015](https://doi.org/10.1016/j.jclinepi.2010.07.015)] [Medline: [21208779](https://pubmed.ncbi.nlm.nih.gov/21208779/)]
28. Higgins JPT, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions* Version 65: Cochrane; 2024. URL: <https://www.cochrane.org/authors/handbooks-and-manuals/handbook> [accessed 2026-03-07]
29. Mancin S, Sguanci M, Reggiani F, Morengi E, Piredda M, De Marinis MG. Dysphagia screening post-stroke: systematic review. *BMJ Support Palliat Care* 2024 Jan 8;13(e3):e641-e650. [doi: [10.1136/spcare-2022-004144](https://doi.org/10.1136/spcare-2022-004144)] [Medline: [37364991](https://pubmed.ncbi.nlm.nih.gov/37364991/)]
30. IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Med Res Methodol* 2014 Feb 18;14:25. [doi: [10.1186/1471-2288-14-25](https://doi.org/10.1186/1471-2288-14-25)] [Medline: [24548571](https://pubmed.ncbi.nlm.nih.gov/24548571/)]
31. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002 Jun 15;21(11):1539-1558. [doi: [10.1002/sim.1186](https://doi.org/10.1002/sim.1186)] [Medline: [12111919](https://pubmed.ncbi.nlm.nih.gov/12111919/)]
32. Schandelmaier S, Briel M, Varadhan R, et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *CMAJ* 2020 Aug 10;192(32):E901-E906. [doi: [10.1503/cmaj.200077](https://doi.org/10.1503/cmaj.200077)] [Medline: [32778601](https://pubmed.ncbi.nlm.nih.gov/32778601/)]
33. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011 Jul 22;343(5):d4002. [doi: [10.1136/bmj.d4002](https://doi.org/10.1136/bmj.d4002)] [Medline: [21784880](https://pubmed.ncbi.nlm.nih.gov/21784880/)]
34. Nagashima K, Noma H, Furukawa TA. Prediction intervals for random-effects meta-analysis: a confidence distribution approach. *Stat Methods Med Res* 2019 Jun;28(6):1689-1702. [doi: [10.1177/0962280218773520](https://doi.org/10.1177/0962280218773520)] [Medline: [29745296](https://pubmed.ncbi.nlm.nih.gov/29745296/)]
35. Alyanak B, İnanır M, Sade SI, Kablanoglu S. Efficacy of game-based EMG-biofeedback therapy in post-stroke dysphagia: a randomized controlled trial. *Dysphagia* 2025 Dec;40(6):1289-1301. [doi: [10.1007/s00455-025-10819-1](https://doi.org/10.1007/s00455-025-10819-1)] [Medline: [40042633](https://pubmed.ncbi.nlm.nih.gov/40042633/)]
36. Zhang B, Wong KP, Liu M, et al. Effect of artificial intelligence-based video-game system on dysphagia in patients with stroke: a randomized controlled trial. *Clin Nutr* 2025 Feb;45:81-90. [doi: [10.1016/j.clnu.2024.12.022](https://doi.org/10.1016/j.clnu.2024.12.022)] [Medline: [39765160](https://pubmed.ncbi.nlm.nih.gov/39765160/)]

37. Kang XF. Clinical efficacy of surface electromyography-triggered virtual reality games combined with swallowing training on post-stroke dysphagia [Dissertation in Chinese]. : Nanchang University; 2024. [doi: [10.27232/d.cnki.gnchu.2024.003484](https://doi.org/10.27232/d.cnki.gnchu.2024.003484)]
38. Borenstein M. Avoiding common mistakes in meta-analysis: understanding the distinct roles of Q, I-squared, tau-squared, and the prediction interval in reporting heterogeneity. *Res Synth Methods* 2024 Mar;15(2):354-368. [doi: [10.1002/jrsm.1678](https://doi.org/10.1002/jrsm.1678)] [Medline: [37940120](https://pubmed.ncbi.nlm.nih.gov/37940120/)]
39. Mátrai P, Kóti T, Sipos Z, Farkas N. Assessing the properties of the prediction interval in random-effects meta-analysis. *Res synth methods* 2026:1-21. [doi: [10.1017/rsm.2025.10055](https://doi.org/10.1017/rsm.2025.10055)]
40. Edwards EA, Lumsden J, Rivas C, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. *BMJ Open* 2016 Oct 4;6(10):e012447. [doi: [10.1136/bmjopen-2016-012447](https://doi.org/10.1136/bmjopen-2016-012447)] [Medline: [27707829](https://pubmed.ncbi.nlm.nih.gov/27707829/)]
41. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games* 2013 Dec 12;1(1):e3. [doi: [10.2196/games.3139](https://doi.org/10.2196/games.3139)] [Medline: [25658754](https://pubmed.ncbi.nlm.nih.gov/25658754/)]
42. Li X, Wu M, Zhang J, et al. Post-stroke dysphagia: Neurological regulation and recovery strategies. *Biosci Trends* 2025;19(1):31-52. [doi: [10.5582/bst.2025.01029](https://doi.org/10.5582/bst.2025.01029)] [Medline: [39993779](https://pubmed.ncbi.nlm.nih.gov/39993779/)]
43. Sales DS, Hammerle MB, deAlmeida VP, et al. Dysphagia and its impact on quality of life in rare neuromuscular disorders. *Arq Neuropsiquiatr* 2025 Feb;83(2):1-6. [doi: [10.1055/s-0045-1804920](https://doi.org/10.1055/s-0045-1804920)] [Medline: [40107292](https://pubmed.ncbi.nlm.nih.gov/40107292/)]
44. Lumbantobing VBM, Priastomo M, Kosasih CE, Rhamelani P, Anggadiredja K, Adnyana IK. Analysis of factors affecting quality of life (QoL) in middle-aged stroke patients (under 65 years old) in the rehabilitation phase: multivariate regression analysis. *J Multidiscip Healthc* 2025;18:3477-3491. [doi: [10.2147/JMDH.S504851](https://doi.org/10.2147/JMDH.S504851)] [Medline: [40538381](https://pubmed.ncbi.nlm.nih.gov/40538381/)]

Abbreviations

CENTRAL: Cochrane Central Register of Controlled Trials

EMG: electromyography

GRADE: Grading of Recommendations, Assessment, Development, and Evaluation

GUSS: Gugging Swallowing Screen

ICEMAN: Instrument for the Credibility of Effect Modification Analyses

MD: mean difference

PI: prediction interval

PICOS: population, intervention, comparison, outcomes, and study design

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Search Extension

PROSPERO: International Prospective Register of Systematic Reviews

RCT: randomized controlled trial

sEMG: surface electromyography

SMD: standardized mean difference

SSA: Standardized Swallowing Assessment

WST: Water Swallowing Test

Edited by S Brini; submitted 08.Aug.2025; peer-reviewed by L Zhao, S Javorszky; accepted 17.Feb.2026; published 26.Mar.2026.

Please cite as:

Liu J, Ou M, Yan M, Xu X, Qin L

Effectiveness of Gamified Swallowing Exercises in Adults With Dysphagia: Systematic Review and Meta-Analysis of Randomized Controlled Trials

JMIR Serious Games 2026;14:e82017

URL: <https://games.jmir.org/2026/1/e82017>

doi: [10.2196/82017](https://doi.org/10.2196/82017)

© Jijia Liu, Meijun Ou, Mengyao Yan, Xianghua Xu, Li Qin. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 26.Mar.2026. 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.

Review

Game Elements in Military Trauma Care Education: Systematic Review

Natalia Stathakarou¹, PhD; Andrzej A Kononowicz², PhD; Maxine G Harjani¹, MSc; Cara Swain^{1,3,4}, PhD, MBChB, PgDip; Klas Karlgren^{1,5,6}, PhD

¹Department of Learning, Informatics, Management and Ethics (LIME), Karolinska Institutet, Stockholm, Sweden

²Department of Bioinformatics and Telemedicine, Jagiellonian University Medical College, Kraków, Poland

³Royal Centre for Defence Medicine, Birmingham, United Kingdom

⁴Institute of Naval Medicine, Alverstoke, Gosport, United Kingdom

⁵Department of Research, Education and Development and Innovation, Södersjukhuset, Stockholm, Sweden

⁶Faculty of Health and Social Sciences, Department of Health and Functioning, Western Norway University of Applied Sciences, Bergen, Norway

Corresponding Author:

Natalia Stathakarou, PhD

Department of Learning, Informatics, Management and Ethics (LIME)

Karolinska Institutet

Tomtebodavägen 18A

Stockholm, 171 77

Sweden

Phone: 46 707799671

Email: natalia.stathakarou@ki.se

Abstract

Background: Game elements may inform the design of both simulations and games. However, evidence on how individual game elements inform the design of military trauma training simulations and their educational purpose remains limited.

Objective: This systematic review aimed to examine which game elements are used in the design of educational simulations for military trauma management, how they are implemented, for what purpose, and what outcomes are reported related to the game elements.

Methods: This is a systematic review conducted in accordance with PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. We included qualitative, quantitative, mixed methods, and design studies describing simulation-based training for military trauma management that incorporated game elements. Studies focusing solely on assessment, noninteractive interventions, or psychological trauma were excluded. Searches were conducted in Medline (Ovid), PubMed, IEEE Xplore, ERIC, Web of Science, ACM Digital Library, and CINAHL from inception to October 14, 2025, identifying 2487 records. Screening and data extraction were performed independently by 2 reviewers. Methodological quality was assessed using the Medical Education Research Study Quality Instrument (MERSQI) and the Côté and Turgeon grid. Results were synthesized using qualitative thematic synthesis.

Results: Forty-two studies published between 1986 and 2025 were included. Most studies were conducted in the United States and included a wide range of simulation modalities and learner populations. Sixteen game elements were identified, with narrative, sensation, imposed choice, time pressure, and scoring being most prevalent. The thematic synthesis identified multiple categories describing how these game elements were implemented. Justifications for the use of game elements were rarely provided; when present, they were primarily linked to realism, emotional engagement, adaptive learning, and feedback. Elements such as badges and competition were seldom used. No study explicitly linked individual game elements to specific educational outcomes. This review is constrained by heterogeneity across studies, an imperfect fit of quality appraisal tools for some study types, and the possibility of missed studies due to search vocabulary limitations.

Conclusions: This systematic review is innovative in providing the first comprehensive synthesis of how game elements are used in military trauma simulations. Unlike previous reviews, it explicitly focuses on the pedagogical purposes of these elements. It offers an overview of the prevalence of game elements in military trauma care education and synthesizes the pedagogical rationales for their use. The lack of studies explicitly linking individual game elements to learning outcomes highlights the need for more intentional research and transparent reporting. Future studies should treat gamification as a set of targeted design choices

rather than as a single overarching strategy, and explore how its motivational dimensions can be effectively leveraged in military trauma training.

International Registered Report Identifier (IRRID): RR2-10.2196/45969

(*JMIR Serious Games* 2026;14:e79163) doi:[10.2196/79163](https://doi.org/10.2196/79163)

KEYWORDS

gamification; game elements; military trauma care; simulation-based learning

Introduction

Gamification is the use of game elements in an endeavor to nudge participants to perform certain actions by adopting a playful attitude [1], and it is a promising approach in health care education [2]. For example, learning activities might be designed to involve solving problems under time pressure, competing or collaborating with others, and earning points or badges [3]. Gamification has been linked to effects on motivation, behavior, and engagement, and learning [4]. In health care education, gamification and games have been shown to be at least as effective as other educational approaches, and in many studies, more effective for improving knowledge, skills, and satisfaction [2]. Gamification has the potential to improve learning outcomes, especially when it uses game elements that improve learning behaviors and attitudes towards learning [5]. Recent research indicates that gamification offers diverse and flexible strategies for enhancing clinical reasoning education across health care disciplines and settings [6]. In disaster education, gamification may enhance learners' retention of knowledge, ability to cooperate, sense of presence, perceived realism, disaster awareness, decision-making ability, practical skills, and coping capacity [7].

Although gamification is known for its engaging and motivational benefits, its pedagogical value remains controversial [1,5]. This controversy arises because different gamification strategies may use various combinations of game elements, leading to diverse outcomes [8]. Interestingly, much of the literature treats gamification as a single approach solution, rather than acknowledging it as a collection of game elements that may vary considerably in purpose and the learning experiences that they offer.

Gamification and simulation are closely related concepts, but are often also contrasted; based on the spectrum of Qin et al [9], Ricciardi and De Paolis [10] conceptualize simulations and games as 2 extremes on a spectrum. At one end of this spectrum lie classical simulators, which are often designed for skills training and prioritize realism by replicating the real world. At the other end are games developed for fun and entertainment, often situated in entirely fictional or imaginary contexts. Between these extremes lie serious games and simulation games. Serious games are developed with nonentertainment purposes in mind, combining a high degree of realism with the entertainment elements of traditional games to facilitate skills development. In contrast, simulation games often blend imaginative or fictitious environments with simulation-based mechanics, offering engagement and the potential to support learning. Examples of serious games for military trauma training include the French Military Health Service's 3D-SC1 game to

train for and assess forward combat casualty care [11,12] and the US Army's tactical combat casualty care simulation training program, TC3Sim [13,14].

Simulations may allow learners to experience complex situations and act as they would in a real environment. They may take several forms and provide learners with feedback. High-fidelity medical simulations are educationally effective, and simulation-based education complements medical education in patient care settings [15]. Simulation environments range from field exercises to virtual patients and virtual reality and can include a variety of different modalities in place of a human casualty, such as a manikin, a simulated patient, or a cadaver.

In military trauma care, different simulation technologies are used to train a range of technical and nontechnical competencies [16,17]. A recent scoping review [18] examined the use of simulations in military medicine and found that most studies focused on physical simulation modalities, such as manikins and task trainers, whereas only a limited number used augmented or virtual reality interventions. Simulation-based training enables the replication of austere and high-stress environments, providing a safe context for learners to practice trauma management, make decisions under pressure, and learn from errors without compromising patient safety [19]. Kubala and Warnick [20] found that knowing exactly what to expect in combat reduces fear and stress.

Military trauma care is characterized by austere environments, tactical demands, and limited medical resources that fundamentally distinguish it from civilian practice. In deployed settings, medical personnel often operate with minimal equipment and may provide care alone or under hostile conditions far from hospital facilities [21,22]. These circumstances require rapid decision-making, prioritization, and coordination under pressure, frequently with incomplete information, which increases the risk of preventable harm [23]. Military trauma also differs from civilian trauma in its organizational structures, triage systems, and treatment approaches, as well as in the nature of injuries: while civilian trauma commonly involves blunt injuries or low-velocity gunshot wounds, military injuries are often caused by blasts and high-velocity weapons [22-27]. Consequently, treatment protocols developed for civilian settings do not always translate effectively to the battlefield. Recent studies [28] emphasize the often-overlooked conditions of truly austere environments in research and education, indicating a need for trauma simulations to integrate realism and austerity. Yet, most military medical personnel receive training in civilian settings, which may not fully prepare them for managing trauma in hostile settings.

Game elements can be used in both simulations and games, serving different purposes [29]. For instance, a recent study [30] integrated game elements such as varying difficulty levels and scoring within the design of virtual patients to make them both more realistic and engaging as learning activities. Previous work analyzing game elements in education has led to frameworks and taxonomies supporting the design and evaluation of gamification in learning environments [3,31]. Extending this work, a list of game elements with the potential to support education and training in military trauma care was synthesized [30].

Incorporating game elements into simulation-based education in the field of military trauma training has the potential to increase motivation, enhance learners' confidence, and support personalized learning [30]. Yet, literature about how individual game elements can inform the design of military trauma training simulations and influence learning outcomes is underexplored.

This systematic review aims to systematically examine the use of game elements in the design of military trauma training simulations in order to retrieve and synthesize international evidence on their design, educational purpose, and reported outcomes in trauma management training. By doing so, the systematic review seeks to inform simulation design practices and identify directions for future research. The research questions guiding this systematic review are:

1. What game elements are used in the design of educational simulations in the context of military trauma management?
2. How are the identified game elements used?
3. What is the purpose of using game elements in the design of educational simulations in military trauma management?
4. What outcomes are reported related to the game elements?

Methods

Protocol and Registration

A systematic review protocol for this study has been published in JMIR Research Protocols (PMID: 37682596) [32]. No changes were made to the planned synthesis methods after protocol publication. The results are reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [33]. The PRISMA checklist, the PRISMA-S checklist, the PRISMA for Abstracts checklist, and SWiM (Synthesis Without Meta-analysis) reporting items are provided in [Multimedia Appendices 1-4](#), respectively.

Inclusion and Exclusion Criteria

We included both qualitative and quantitative empirical and design studies that addressed different types of simulation interventions, which incorporated game elements. Game elements were identified using published gamification frameworks in a deductive-inductive manner [3,30,31]. We included studies that incorporated game elements designed to enhance the education and training of military trauma management. The included simulation interventions had a clear educational purpose within this context. In this systematic review, the included gamified simulations had some degree of interactivity, allowing the scenarios to unfold in response to the

learners' actions. Studies in which participants passively observed a scenario (eg, a video clip) were excluded. Studies using interventions primarily for patient education, rehabilitation, teleconference, treatment, and decision-support were considered outside the scope of this review. We excluded studies where the training only focused on individual body parts. We also excluded studies in which gamification was used solely for assessment, unless the assessment served an educational purpose. We included studies focusing on military trauma management and excluded those addressing psychological trauma, because it belongs to a different educational context and learner group. We included studies that had a link to military medicine, even if the learner population that received the simulation was linked to a civilian context. Game elements only taking place outside the simulation, such as scoring conducted by instructors who observed or assessed participants after a simulation session, were also excluded. Several simulations incorporated teams to develop team-relevant competencies, and a "team" is recognized as a game element in the Maheu-Cadotte framework [3]. We only included studies that introduced teamwork with a gamification purpose and excluded studies where teamwork was present merely because it was inherently a part of the simulation, such as when practicing communication skills.

Search Methods for Identification of Studies

We explored the databases/search engines: Medline (Ovid), PubMed, IEEE Xplore, ERIC, Web of Science, ACM Digital Library, and CINAHL with the help of librarians at the University Library of Karolinska Institutet. We included all articles regardless of publication language. All articles published up to October 14, 2025, were retrieved. With the help of the librarians, we conducted a search of citations and references in CitationChaser. Each database was searched independently, and no multi-database search function was used. No limits or filters were applied. The search strategies did not undergo a formal peer review, although they were developed by 2 experienced librarians and iteratively tested across several rounds. No study registries were searched. No additional online or print sources were purposefully browsed (eg, tables of contents, conference proceedings, or websites) beyond the citation and reference searching described above. Search strategies were developed specifically for this review in collaboration with librarians and were not adapted or reused from previous literature reviews. The search strategy is included in [Multimedia Appendix 5](#) [34].

Selection of Studies

We imported all identified references to the Rayyan open-source web system [35]. Full-text versions of the included abstracts were retrieved via the university library. We did not contact authors or other stakeholders to obtain additional studies or data. Two researchers independently assessed the identified studies based on the inclusion and exclusion criteria. Any disagreements were resolved through discussion between the 2 reviewers. If no agreement could be reached, a third researcher was consulted. The selection process was represented using a PRISMA flow diagram [33], including the actual number of studies included and excluded at each stage of screening.

Data Extraction and Management

The data extraction sheet and initial coding frame for identifying game elements and outcomes were predefined and published in the registered review protocol [32]. To ensure consistency and a shared understanding of the coding approach, all authors jointly piloted the extraction process on 3 studies during the protocol development stage. This pilot served as a coder calibration exercise to refine code definitions and decision rules.

Throughout the review, the first author was paired with each coauthor during data extraction and coding to maintain alignment and consistency in interpretation. Two researchers independently extracted and managed the data for each included study using the finalized structured form. Any discrepancies were discussed until consensus was reached, and a third author was consulted when necessary.

Data Analysis, Synthesis, and Reporting

A qualitative data analysis combined with thematic synthesis [36] was conducted to answer the research questions. Data synthesis was performed using structured Excel spreadsheets; no qualitative analysis software was used. The simulations described in the included studies were classified according to categories proposed in earlier literature [16,17]. Virtual patients were categorized following the classification framework introduced by Kononowicz et al [37]. To answer the first research question, the reviewers compared extracted data to ensure consistency in coding and interpretation of game elements across studies. The data extraction process was guided by a predefined list of game elements based on existing frameworks [3,30,31].

To address the second research question, we synthesized the data further by grouping realizations of each game element. This involved reviewing how each element was operationalized across studies and clustering similar patterns into subcategories. This synthesis allowed us to develop categories that describe the implementation of each game element. The process was iterative and collaborative, followed by frequent discussion between all authors to refine the categories, ensure consistency, resolve discrepancies, and reach a consensus.

To answer the third research question, a thematic synthesis approach was applied. Paired researchers worked independently to review each study and extract passages of text that provided a rationale, justification, or implied educational purpose for the use of specific game elements. These included both explicit explanations and implicit meanings inferred from the study context. Subsequently, we clustered similar meanings into broader themes that captured the underlying educational or experiential intentions behind the use of the identified game elements. This involved a process of thematic grouping that connected the extracted rationales with the corresponding game elements identified earlier. The mapping was conducted by 2 researchers and discussed by all authors. Themes were constructed inductively from the data. Each theme was then linked to one or more game elements.

To address the fourth research question, we collected information on the reported educational or performance

outcomes of the simulations attributed to individual game elements as described in each study. Data extraction focused on outcomes related to knowledge, skills, and attitudes, specifically engagement, confidence, and secondary outcomes, like accessibility and cost-effectiveness. Although outcomes were recorded alongside the included studies, no formal synthesis was conducted in the located studies to systematically link specific game elements to particular outcomes. This was due to limitations in the way outcomes were reported across studies as a combined effect of all game elements, rather than separately for each game element, and hence the lack of explicit connections between individual game elements and measured results.

No standardized outcome metrics or effect-size transformations were applied. Given heterogeneous outcome reporting and the absence of comparable effect estimates, findings were synthesized using qualitative thematic synthesis, focusing on the implementation and pedagogical purpose of game elements. All included studies contributed to the synthesis; no studies were selected or excluded from the synthesis based on study design or quality assessment, which was used descriptively to support interpretation rather than to weight or filter findings.

Findings were presented using structured tables and narrative summaries. Tables reported key study characteristics (eg, study design, simulation type, learner population) and the presence and implementation of individual game elements; studies were organized descriptively rather than ordered by effect size or risk of bias, as no comparable effect estimates were available.

Quality Appraisal

The quality of the included studies was assessed using 2 established instruments. For quantitative and mixed methods studies, we applied the Medical Education Research Study Quality Instrument (MERSQI) [38]. For research that reported results obtained exclusively through qualitative methods, such as interviews or focus groups, we used the quality appraisal grid developed by Côté and Turgeon [39]. Studies focusing solely on the design and development of simulations were not included in these quality assessments as they fall outside the intended scope of both MERSQI and the Côté and Turgeon grid [39]. Methodological quality was appraised descriptively to support interpretation of the evidence base. Certainty of the synthesis findings was not formally assessed using a grading framework.

Ethical Considerations

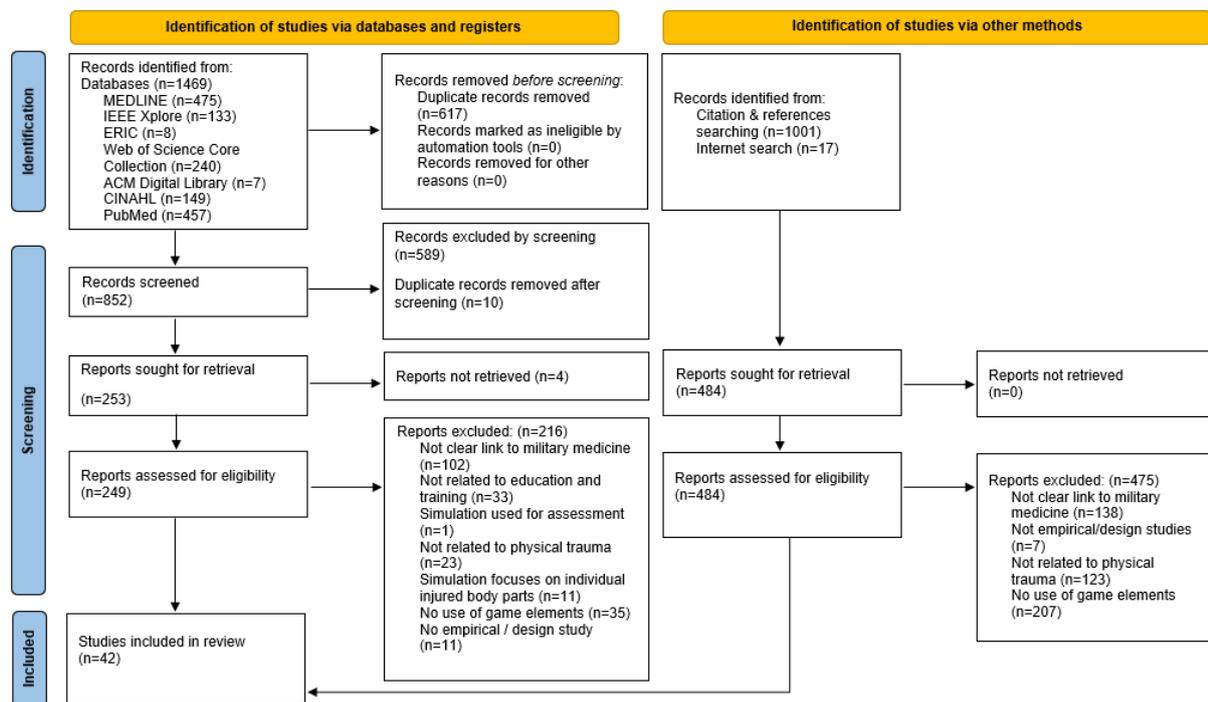
This systematic review does not involve processing of sensitive personal data and therefore ethical approval is not required according to the Swedish Ethical Review Act.

Results

Included Studies

In total, 42 studies were included. [Figure 1](#) presents the PRISMA flowchart showcasing the search and inclusion process. If an article was excluded for multiple reasons, only the first applicable exclusion criterion, as defined in the published review protocol [32], was recorded.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram.



Article and Study Characteristics

The studies covered a variety of simulation types used for training purposes, including nondigital games, virtual patients, virtual patient games, live simulations, standardized patients, mannequins, and virtual and mixed realities. Most of the studies were conducted in the United States, followed by studies

originating from Europe. The learner populations involved were diverse, ranging from hospital decision-makers, surgeons, soldiers, and combat medics to medical students and emergency response teams. Table 1 provides information about the general characteristics of the studies included, specifically the type of simulation, the country in which the study was conducted, and the learner population.

Table 1. Included studies, type of simulation, and learner population.

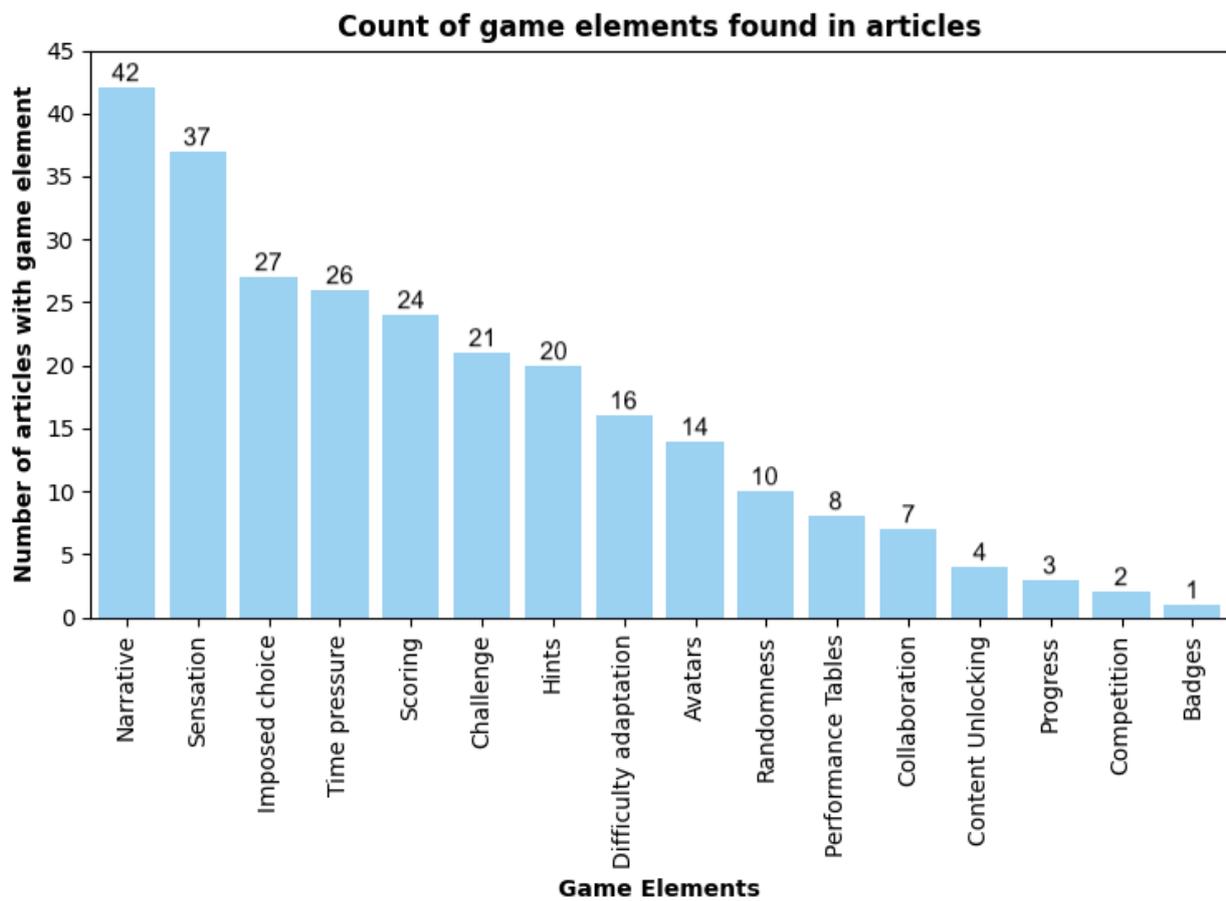
Lead author and year	Simulation type	Learner population	Country
Achatz et al (2020) [40]	Nondigital game/board game	Hospital decision makers	Germany
Arora et al (2014) [41]	Human Standardized Patient	Surgeons, anesthesiologists, physicians, and nurses	United Kingdom
Badler et al (1996) [42]	Virtual patient game	Combat Medics	United States
Baird et al (2020) [43]	Virtual patient - High Fidelity Software Simulation	Thermal injury treatment providers	United States
Beaven et al (2021) [44]	Cadaver and High Fidelity Manikins	National Guard members, doctors, nurses, physicians, operating department practitioners, medical students	United Kingdom
Brown et al (2016) [45]	Virtual and mixed realities	Combat Medics	United States
Chi et al (1996) [46]	Virtual patient game	Combat Medics	United States
Chi et al (1997) [47]	Virtual patient game	Combat Medics	United States
Cohen et al (2013) [48]	Virtual patient game	Clinicians practicing trauma leadership; clinical major incident coordinator/silver commander.	United Kingdom
Couperus et al (2019) [49]	Virtual and mixed realities	Emergency military physicians	United States
Couperus et al (2020) [50]	Virtual and mixed realities	Emergency military physicians	United States
DeFalco et al (2017) [13]	Virtual patient game	Combat medics	United States
de Lesquen et al (2022) [51]	Virtual patient game	Emergency doctor	France
de Lesquen et al (2023) [52]	Virtual patient game	Prehospital physicians	France
Du et al (2022) [53]	Virtual and mixed realities	Military medical students	China
Freeman et al (2001) [54]	Virtual and mixed realities	Navy medical providers (paramedics)	United States
Goolsby et al (2014) [55]	Immersive virtual environment	Military medical students	United States
Hemman (2005) [56]	Virtual patient - High fidelity software simulation	Combat medics	United States
Henderson et al (1986) [57]	Virtual patient - High fidelity software simulation	Medical students	United States
Henderson et al (2020) [58]	Virtual patient game	Combat medics	United States
Kyle et al (2004) [59]	High fidelity Manikins and Human standardized patient	Physicians, nurses, paramedics, professional scientists, military officers, lawyers, career politicians, consultants from non-governmental organizations, administrators, intelligence officers, and logistic personnel.	United States
Lombardo et al (2022) [60]	Virtual and mixed realities	Emergency medicine residents, attendings, medical students, physician assistants, army medics, and nurses	United States
Lu et al (2023) [61]	Virtual and mixed realities	Combat medics and military surgeons	China
Lennquist Montán et al (2014) [62]	A card game in a live exercise	Physicians, nurses, ambulance/paramedics, Military staff, administrators, also collaborating agencies (rescue services, the police)	Sweden
Netzer et al (2015) [63]	High fidelity manikins	Navy Emergency Medical Teams, military physicians	Israel

Lead author and year	Simulation type	Learner population	Country
Pasquier et al (2016) [11]	Virtual patient game	Soldiers, combat medics	France
Planchon et al (2018) [12]	Virtual patient game	Soldiers, combat medics	France
Qin et al (2024) [64]	Virtual and mixed realities	Combat medics, nurses	Israel
Rabotin et al (2023) [65]	Virtual and mixed realities	Paramedics, physicians	Israel
Satava and Jones (1996) [66]	Virtual and mixed realities and wearables	Combat medics	United States
Sonesson et al (2023) [67]	Interactive patient scenarios	Military trauma teams	Sweden
Sotomayor (2008) [68]	Virtual patient game	Combat medics	United States
Sotomayor (2010) [14]	Virtual patient game	Combat medics	United States
Stansfield et al (1998) [69]	Virtual and mixed realities	Combat medics	United States
Stathakarou et al (2024) [30]	Interactive patient scenarios	Combat medics	Sweden
Stone (2005) [70]	Virtual patient game	Military trauma surgeons	United Kingdom
Stone (2011) [71]	Virtual patient game	Military trauma surgeons	United Kingdom
Stone et al (2017) [72]	Virtual and mixed realities	Medical Emergency Response Teams (MERTs)	United Kingdom
Tretyak et al (2025) [73]	Virtual and mixed realities	Medical personnel and trainees involved in tactical emergency or combat casualty care	Austria
Wier et al (2017) [74]	Immersive virtual environment	Medical Emergency Response Teams (MERTs: physicians, nurses, medics)	United States
Willy et al (1998) [75]	Virtual patient - High fidelity software simulation	Military physicians	Germany
Zhu et al (2024) [76]	Virtual patient game	Mobile medical logistics teams: background in medicine, nursing, logistics	China

Game Elements Identified in the Design of Military Trauma Simulations

Figure 2 presents a synthesis of the 16 game elements identified in the design of military trauma simulations across the included studies. Multimedia Appendix 6 [11-14,30,40-76] provides a detailed categorization of the game elements identified in each of the 42 studies. The definitions of the game elements were derived from previous literature [3,30,31]; narrative refers to the structured sequence of events and decisions shaping the learner's experience, while sensation captures the use of visual or auditory stimuli to enhance immersion. Imposed choice describes situations in which learners must select one option to progress, whereas time pressure requires actions or decisions within urgency or a limited timeframe. Scoring provides

quantitative feedback with points, and hints refer to clues that support learners without revealing the correct answer outright. Challenge encompasses mechanisms designed to test abilities, and difficulty adaptation refers to the dynamic complexity of tasks and adjusting. Avatars function as digital representations of learners or patients, and randomness introduces unpredictable aspects in the scenario. Performance tables present detailed summaries of accomplishments across tasks, and collaboration encourages learners to work together toward shared objectives. Content unlocking restricts access to new material until specific criteria are met, progression visualizes the learner's development over time, competition supports comparisons with mechanisms such as leaderboards, and badges serve as symbolic markers of achievement for completing tasks. The definitions of the game elements are summarized in Multimedia Appendix 7.

Figure 2. Synthesis of game elements identified in the design of military trauma simulations.

Application of Game Elements in Military Trauma Simulations

Table 2 provides a thematic synthesis of how game elements were used in the included studies, outlining different categories

of how game elements can be used to inform educational simulations in trauma care.

Table 2. Thematic synthesis of how the game elements were used in the included studies.

Game element	Application of game elements in military trauma simulations
Narrative	<ul style="list-style-type: none"> An interactive scenario that unfolds on the basis of the learners' decisions [11-14,30,40-76]
Sensation	<ul style="list-style-type: none"> High environmental fidelity by replicating battlefield conditions [41,44,59] Virtual fidelity in the digital environment to depict austerity, such as realistic portrayals of injury or digitally recreated battlefield environments [11-14,30,42,43,45-54,57,60-62,64,66,68-71,73,75,76] Mixed reality and interaction with objects in both physical and virtual environments [55,65,72,74]
Imposed choice	<ul style="list-style-type: none"> Presentation of explicit decision options on patient cards [40] Digital menu with predefined decisions [11-13,30,43,45-48,50,52,54,57,58,60,61,64,65,67-71,75,76]
Time pressure	<ul style="list-style-type: none"> Scenarios incorporating explicit time constraints [11,40,70-72,74] Tasks requiring immediate action to simulate urgency [30,48,50-52,55,59,73] Display of visible timers or auditory cues such as ticking clocks to reinforce urgency [45,51,54,57,64,75,76] Delayed decisions negatively affecting patient outcomes, potentially leading to patient death [13,48,50,62,69-71] Timely decision-making contributing positively to performance metrics [42,65]
Scoring	<ul style="list-style-type: none"> Quantifying learner performance based on patient health outcomes, such as trauma scores or health points [40,42,46,47,62,69] Final scoring mechanisms summarizing overall performance or decision quality at the end of the scenario [11-13,30,41,45,51,52,60,61,63,65,70,76] Real-Time Scoring: Continuous assessment during gameplay [30,40,45,49,50,64] Penalty-Based Scoring: Deducting points for incorrect decisions [65] Competitive scoring formats allowing ranking or comparison between participants [57] Scoring linked to resource management, rewarding efficient allocation [57]
Hints	<ul style="list-style-type: none"> Human instructor-guided feedback during simulation sessions [42,44,55,59] Computerized real-time audio feedback [45] Simulated patient cues through physiological responses, appearance, or facial expressions [13,42,48-50,64,66,69,70,75] Computerized feedback via virtual patient interactions or dialogue boxes [30,56,57,68,76] Virtual instructor or colleague intervening when poor decisions occur [11,57] Feedback on inappropriate or excessive resource use [57] Contextual hints through simulated live TV news updates in real time [59]
Challenge	<ul style="list-style-type: none"> Managing complications and unforeseen events [40,43,48,63] Exposure to distracting sounds or visual elements [44,45,72,73] Performing difficult triage decisions, prioritizing treatable patients over those unlikely to survive [61] Decision-making under conflicting or incomplete information [40,59] Identifying concealed or initially nonobvious injuries [30,66] Providing care in austere or unfamiliar environments [30,41,42,44,63,66,69,72,73] Encountering unwinnable cases with inevitable failure outcomes [13,30,58] Experiencing learner death when safety precautions are neglected [30,68]
Difficulty adaptation	<ul style="list-style-type: none"> The "game master" can adjust the scenario's difficulty based on participant performance or situational needs [40,62,63] Learners are able to select the initial difficulty level before the scenario [11,45] Gradually increasing or varying levels of challenge [11,13,30,45,50,51,54,57,58,67,75] Dynamic difficulty in the scenario based on participants' performance [42,63,65]
Avatar	<ul style="list-style-type: none"> Patient cards representing virtual patients [40] Learner avatars enabling player interaction through a virtual self or character [42,43,48,54] Customizable learner avatars [45] Avatars mirroring learners' physical movements in real time [46,47,66,69] Patient avatars representing casualties within the simulation [51-53,73]
Randomness	<ul style="list-style-type: none"> Unexpected resource problems: missing equipment, availability of staff, hospital resources [40,62] Scenario variability in incident type, number of casualties, environmental austerity, weather conditions, and patient injuries or physiology [42,43,45,50-52,62,64] Identical treatments do not always produce the same outcomes [57]
Performance tables	<ul style="list-style-type: none"> User performance summaries comparing actions against predefined "gold standard" treatment procedures [45] Event timing and adherence to trauma resuscitation protocols [65] Interactive display of casualty vital signs and key events with a complete performance log [45] Performance grids integrating scoring and structured debriefing [11,51,52,64,70,76]

Game element	Application of game elements in military trauma simulations
Collaboration	<ul style="list-style-type: none"> Multi-avatar collaboration between human players within the same simulation environment [42,45] Collaborative tasks designed to enhance understanding of different professional roles [45,53] Collaboration with virtual or artificial intelligence-driven team members and avatars [50,53,57,60,72]
Content unlocking	<ul style="list-style-type: none"> Completion of required treatment steps to progress within the scenario [45] Selection of specific choices or actions to advance gameplay [30,40] Demonstration of proficiency or task mastery to unlock subsequent levels [11]
Progress	<ul style="list-style-type: none"> Sequential management and triage flow of patients throughout the scenario [40] Grid-based visualization of patient health and status [51] Rescue progress bar displaying ongoing operations and cumulative training score in real time [76]
Competition	<ul style="list-style-type: none"> Scores used to compare and rank learners, enabling performance competition among peers [11,57]
Badges	<ul style="list-style-type: none"> Bronze, silver, and gold medal graduation according to the scoring system, integrating time and actions delivered [11]

Purposes of Using Game Elements in the Design of Simulation in the Included Studies

While most of the studies did not attempt to attribute the use of game elements to a specific educational purpose, the following studies were identified that justified the use of specific game

elements. Table 3 summarizes the 9 identified themes explaining why specific game elements were integrated into the simulation design. Supporting data excerpts from the included studies are provided in Multimedia Appendix 8 [11,13,30,40,41,43,45-47,52-54,57,60,62,63,65,69,73-75].

Table 3. Justification themes of using game elements in the included studies.

Theme	Game element
Realism and emotional engagement	Narrative [30,41,43,45,51,57,63]; Sensation [11,30,41,43,45,46,51,57,63]; Time pressure [30,40,73]; Avatar [45,51]; Randomness [45,57]; Challenge [45,73]; Imposed choice [30,47,62,75]; Collaboration [53]
Adaptive learning and feedback	Difficulty adaptation [45,51,60,65]; Scoring [11,47,51,60]; Performance tables [11,45,51]
Affective learning	Challenges [12]
Learner agency	Imposed choice [12]
Challenge the learners	Difficulty adaptation [54]
Risk-free experiential learning	Narrative and Sensation [54], Time pressure [69]
Motivation and engagement	Competition and Scoring [11]
Situational awareness	Challenge [69]
Emotional regulation and overcoming anxiety	Challenge and Sensation [74]

Reported Outcomes Associated With Game Elements

None of the included studies correlated individual game elements with the reported outcomes. Instead, the effects on knowledge, skills, and attitudes were attributed to the impact of the simulations as a whole.

Quality Appraisal of Studies

We evaluated 22 studies with MERSQI with scores ranging from 5 to 15.5 and a median score of 9.5 (Table 4). The most common methodological limitations identified included reliance on a single cohort without comparison groups, the use of nonvalidated evaluation instruments, and a focus on outcomes such as satisfaction or basic knowledge and skills acquisition.

Since our inclusion criteria did not restrict studies to pre-post designs, the MERSQI tool items were not applicable in all cases, which may have contributed to lower scores in some studies. We evaluated only 2 studies with the grid by Côté and Turgeon [39] that used qualitative data collection methods. We excluded 18 design and development studies from the quality appraisal due to their descriptive nature, which falls outside the scope of both MERSQI and the Luc Côté grid. The evaluation score of each study is briefly available in Tables 4 and 5. Design studies were not assessed with formal quality appraisal tools [11,42,43,45-47,49-51,54,57,58,66,69-72,75]. Table 5 includes the 2 studies that were appraised using the Côté and Turgeon grid [39].

Table 4. Quality appraisal for studies using the Medical Education Research Study Quality Instrument.

Study	Study Design (max=3)	Sampling (max=3)	Data type (max=3)	Validity (max=3)	Data analysis (max=3)	Outcomes (max=3)	Score (max=18)
Achatz et al (2020) [40]	1	1.5	1	0	2	1	6.5
Arora et al (2014) [41]	1	2	1	1	3	1.5	9.5
Beaven et al (2021) [44]	1.5	2	3	1	3	1.5	12
Cohen et al (2013) [48]	1	0.5	3	2	3	1.5	11
DeFalco et al (2017) [13]	3	1	1	1	3	1.5	10.5
de Lesquen et al (2023) [52]	2	2	3	1	3	1.5	12.5
Du et al (2022) [53]	3	2	1	0	2	1.5	9.5
Goolsby et al (2014) [55]	1.5	2	1	0	2	1	7.5
Hemman (2005) [56]	1.5	2	1	0	3	1.5	8.5
Kyle et al (2004) [59]	1	0	1	0	2	1	5
Lombardo et al (2022) [60]	1	0.5	1	0	2	1	5.5
Lu et al (2023) [61]	1	1.5	1	0	2	1	6.5
Lenquist Montán et al (2014) [62]	1	3	1	0	2	1	8
Netzer et al (2015) [63]	1.5	2	3	2	3	1.5	13
Planchon et al (2018) [12]	3	2	3	2	3	1.5	14.5
Qin et al (2024) [64]	1.5	0.5	3	0	3	1.5	9.5
Rabotin et al (2023) [65]	1	0.5	1	0	2	1	5.5
Sonesson et al (2023) [67]	2	0.5	1	0	2	1.5	7
Sotomayor (2008) [68]	3	2	3	2	3	1.5	14.5
Sotomayor (2010) [14]	3	0.5	3	1	3	1.5	12
Wier et al (2017) [74]	2	2	3	3	3	1.5	14.5
Zhu et al (2024) [76]	3	2	3	3	3	1.5	15.5

Table 5. Quality appraisal for studies using the Côté and Turgeon grid [39].

Study	Introduction (max=2)	Methods (max=5)	Results (max=2)	Discussion (max=2)	Conclusion (max=1)	Total (max=12)
Stathakarou et al (2024) [30]	2	5	2	2	1	12
Tretyak et al (2025) [73]	2	4	2	2	1	11

Discussion

Principal Findings

In this systematic review, we investigated the use of game elements in simulations for military trauma management. The results provide insights into the common application of the game elements and summarize the justification for their use. However, no study explicitly linked individual game elements to specific learning or performance outcomes.

In health care education, gamification and games have been shown to be at least as effective as other educational approaches, and in many studies, more effective for improving knowledge, skills, and satisfaction [62]. Gamification has the potential to improve learning outcomes, especially when it uses game elements that improve learning behaviors and attitudes towards learning [63]. However, many of the studies included in previous

reviews are of low quality, lack a sufficient focus on specific game elements [62], and have methodological limitations [5].

Although games and simulations have been used in military training for centuries [77], research on how individual game elements can inform the design of military trauma training simulations and influence learning outcomes is underexplored. The question of linking specific game elements to outcomes has been raised in gamification studies [8,78] but received little attention in military trauma contexts. Even if the link between game elements and learning outcomes were established, designers must account for the fact that individual elements can be implemented in multiple ways, leading to variation in results.

Clarification studies, which explore how underlying mechanisms account for the observed effects [79], are uncommon in gamified learning research [5] and are largely absent in military trauma contexts. Such studies could provide insights into the

mechanisms involved in gamified learning, inform the design of gamified learning approaches [4,5], and contribute to the understanding of the relationship between design principles and learning outcomes, such as clinical reasoning [80].

As early as 1957, Caillois [81] described characteristics that distinguish games from other forms of activity, including fun, uncertainty, detachment from real life, nonproductivity, the presence of specific rules, and fictitious settings. Garriss et al [82] proposed 6 dimensions that distinguish games from traditional simulations, including fantasy, mystery, sensory stimuli, rules and goals, controls, and challenge.

In this systematic review, narrative and sensation, as the most frequently used game elements in the reviewed simulations, contribute to the fictitious settings by fostering immersion and enabling environmental recreation. Narrative was frequently combined with sensation, the use of sensory stimuli, such as visual and auditory cues, to enhance immersion, often reflecting the austerity and constraints of military trauma contexts. In the application of digital technology, multimedia has been shown to enhance learning outcomes such as satisfaction, achievement, motivation, and attention [83]. In live simulation exercises, sensation was often conveyed by replicating battlefield conditions. This was achieved by conducting simulations outside the classroom or in rare environments, such as simulated deployed hospitals or ships, and by incorporating battlefield sounds and stress-inducing elements to recreate austerity [41,44,59].

Difficulty adaptation was identified in 16/42 included studies. The level of difficulty is a critical factor in learning; it should remain below the maximum capacity of learners' knowledge to maintain motivation and facilitate effective learning. When difficulty exceeds learners' capabilities, performance and engagement may decrease [84]. In the included studies, difficulty adaptation was implemented in various ways, for instance, in a classroom, with an instructor adjusting the level to the learners' performance [40,62,63] or by allowing the learners to select different levels prior to the start of the game [11,45].

"Challenge" as a game element was identified in several studies. This differs from the level of difficulty, which corresponds to the level of knowledge and the skills the learners need to apply to perform the task, as challenge often involves dealing with complications and unforeseen circumstances [40,43,48,63]. Examples include the requirement to make difficult triage decisions (eg, prioritizing treatable patients over those unlikely to survive) [61], decision-making with conflicting information [40,59], distractors such as background noise and visual disturbances [44,45,72,73], or emotionally charged scenarios and unwinnable cases [13,30,58].

Simulations also frequently included imposed choice, which presented users with predefined options or decision menus. Additionally, scoring and time pressure were commonly used to provide feedback or assess learner performance and simulate the urgency of trauma care, respectively. Scoring was included in the total of 24/42 studies. In several cases, it was based on the performance of the decision-making of the learners, providing learners with performance-based feedback

[11-13,30,41,45,51,52,60,61,63,65,70,76]. A recently published study [85] proposes that learners should be given enough opportunities to fail and try again to support serious gaming for disaster preparedness; however, without direct feedback after failure. While repetitive training enables learners to refine their problem-solving strategies over time, a principle in line with Ericsson et al's [86] theory of deliberate practice, feedback combined with scoring may enable reflection in military medicine [30].

In total, 26/42 studies included time pressure, which was introduced both directly and indirectly within several simulations. For example, in some cases, time pressure was made explicit through the display of a ticking clock on the screen [45,51,54,57,64,75,76], whereas in others it was conveyed more implicitly by prompting learners to make rapid decisions under conditions where delays could negatively affect patient outcomes and performance scores.

While some of the game elements, such as narrative, sensation, and time-pressure that simulate contextual austerity of the military environment were often present, other types of game elements that are commonly associated with the playful and entertaining side of gamification were underrepresented. For example, badges appeared only in the form of medals in Pasquier et al [11]. Games have a long history in military medical training, yet the emphasis has perhaps traditionally been placed on their potential to mimic an austere environment, rather than on entertainment. It is possible that the playful dimensions of gamification were intentionally de-emphasized to maintain the seriousness of military trauma training. Alternatively, such elements may have been included in the simulations but not explicitly described in the study reports.

Competition was identified only in 2 studies [11,57], and in both cases, it was accompanied by scoring. Competition has been widely discussed in educational theory for its potential to enhance motivation and engagement. Malone and Lepper [87] emphasized competition as a mechanism that fosters intrinsic motivation by providing learners with clear goals and immediate feedback, often driving individuals to achieve better outcomes when compared with others. Van Eck and Dempsey [84] extended this idea, suggesting that competition may stimulate both intrinsic and extrinsic motivation: "For learners who are extrinsically motivated by social standing and recognition, competition against other individuals may serve to increase their efforts and perseverance in the instructional game to gain standing among their peers. Learners who are intrinsically motivated may likewise compete against their own score to see how much better they can do" [84]. In the same study, competition is linked to the concept of challenge.

Johnson and Johnson [88] broadened the concept by emphasizing collaborative competition, where individuals or teams compete in a way that supports group cohesion and shared learning objectives. This form of competition minimizes the potential negative effects, such as stress or discouragement. Collaboration was discussed in studies as training team skills for trauma training was the purpose of the intervention. In this review, however, we only included collaboration as a game element when it extended beyond the general goal of team

training and was implemented in a distinct or designed manner. Examples included multi-avatar collaboration between human players within the same simulation environment [42,45], collaborative tasks designed to enhance understanding of different professional roles [45,53], or collaboration with virtual or artificial intelligence-driven team members and avatars [50,53,57,60,72].

Understanding the motivations behind the use of game elements is challenging, particularly in studies where design intentions are not made explicit and game elements are not formally acknowledged. In the studies where an explicit rationale was provided, linking game elements to justifications of their purpose resulted in the identification of 9 themes: realism and emotional engagement [11,30,40,41,43,45-47,51,53,57,62,63,73,75], adaptive learning and feedback [11,45,47,51,60,65], affective learning [12], learner agency [12], challenge the learners [54], risk-free experiential learning [54,69], motivation and engagement [11], situational awareness [69], emotional regulation and overcoming anxiety [74].

When examining the stated purpose behind the use of game elements, we found that elements such as narrative and sensation were most frequently justified in terms of their contribution to realism and immersion. In contrast, most other game elements were either not discussed at all or not explicitly linked to any specific educational function. Interestingly, even in studies that mentioned motivational aims in their background sections, game elements were rarely linked to motivation or engagement. For instance, the study by Achatz et al [40], which used a broad range of game elements mentioned in the background: “The didactic approach and the course structure ensure the interest, motivation, and progress of the target group, namely experienced clinical decision makers.” Despite this statement, looking into the use of game elements and their justification, time pressure was not presented as a motivational design choice but rather as a mechanism to simulate real-time conditions. We only identified one study [11] linking competition and scoring to motivation and engagement: “Furthermore, through the processes of scoring and gamification applied in 3D-SC1, the trainee is motivated to improve his personal experience. He also shares his scores with his peers in a competitive and engaging challenge”.

When examining the stated purpose of the simulation itself in a broader way rather than looking at the scope of specific individual game elements, we found that in the few studies explicitly referring to gamified simulations or serious games, the primary goals were often to attract and sustain learners’ interest, and to enhance motivation and engagement. For instance, Planchon et al [12] emphasized the motivational and entertaining aspects of serious games, noting: “SGs are similar to video games in that they are engaging, rewarding, and fun. However, at the same time, they can also be used to educate or train.” In contrast, studies that did not explicitly use the terms game or gamification tended to incorporate game elements to enhance realism and immersion, rather than playfulness or entertainment. These elements were often integrated to replicate the conditions and constraints of real-world environments. For example, Beaven et al [44] noted: “Our aim was to deliver a highly realistic, immersive simulation training experience,

teaching both technical and nontechnical skills necessary for the management of war injuries in the austere environment of a far forward surgical facility. Recreating the physical and psychological work environment in a realistic way was desirable in order to encourage people to behave as they would in real life. By fostering this real-life behavior, the participants are better able to imagine the authentic scenario, and training becomes more immersive as a result.” In such cases, game elements appeared to function to enhance realism and authenticity rather than as mechanisms to directly foster learner motivation or engagement. Other studies, such as Sotomayor [14], noted that games: “appeal to the younger generation that has been exposed to their use since early age. Motivation is a big factor observed within the training audience”. However, even in such cases, the justification remained general and did not extend to a discussion of how specific game elements serve defined educational purposes and support motivation.

Realism in simulation encompasses physical, conceptual, and emotional fidelity, each contributing to the authenticity of the learning experience [89]. Advances in artificial intelligence have further expanded the potential for realism by enabling more dynamic and immersive environments [30,90]. On the other hand, some research points out a cognitive bias towards highly realistic and technically advanced learning tools; this review included 2 studies [70,71] which noted that while multimedia special effects may appeal to avid gamers, they can also distract serious users and impair performance. Effective multimedia materials require careful attention not only to managing cognitive load, but also to ensuring that the chosen media formats support, rather than hinder, learning [91].

None of the included studies explicitly linked individual game elements to reported learning outcomes. When the outcomes of the simulations were considered more broadly, most studies described positive educational effects. This observation aligns with findings from a recent systematic review on gamification in disaster education [92], which concluded that gamification could enhance competencies such as emergency response, decision-making, and teamwork in disaster nursing education, and can support learning engagement through game elements such as cooperation, competition, scoring, and scenario-based activities. However, that review also did not analyze the direct relationship between specific game elements and particular learning outcomes.

Design Implications

The findings of this systematic review suggest that game elements should be treated as targeted design choices that are selected to serve a clearly stated pedagogical intention, rather than added as generic “gamification.” A practical starting point is to define how a simulation is intended to be supported through game elements, and then to implement the corresponding elements as a coherent configuration. For simulations aiming to strengthen realism and emotional engagement in austere trauma contexts, narrative and sensation could be used to create a realistic and immersive experience, and then intentionally reinforced through time pressure, imposed choice, and selected forms of challenge, randomness, avatar use, or collaboration to recreate uncertainty, constraints, and team demands that

characterize deployed care. These game elements were previously identified to contribute to realistic tactical experiences for civilian and military trauma care [30] and align with the game elements reported in Table 3.

When the intention is adaptive learning and feedback, limited evidence from this review suggests the use of game elements such as difficulty adaptation, scoring, and performance tables. In the studies included in this review, motivation and engagement were rarely justified at the game element level and, when they were, this was linked to competition and scoring [11]. However, when deciding on what game elements to use, educators and designers might want to consider that time pressure in simulation might not correspond to the time required in the field, and some of the elements, such as scoring and competition, might be perceived as misaligned with the learning goals [93]. Therefore, a good practice might be to present such elements as optional and configurable features [93].

Finally, because none of the included studies linked individual game elements to outcomes, educators, designers, and researchers are invited to consider the hypothesized mechanism associated with each chosen element and to align evaluation with those mechanisms. In practice, this means moving beyond whole-simulation outcome claims and using designs and measures that can test element-level contributions, thereby improving understanding of how specific elements shape decision-making and the learning experience.

Methodological Considerations and Limitations

This review was designed and conducted as a systematic literature review, following a published protocol [32] and transparent, reproducible procedures for searching, screening, extraction, synthesis, and quality appraisal [94]. The purpose was to retrieve international evidence of the impact of game elements in trauma management training and inform design practices and future research. Although the first research questions involved descriptive accounting of design features, which might be considered a hallmark of a scoping review [94,95], the subsequent thematic syntheses produced a structured and critically appraised summary of the use and educational justification for game elements in trauma management simulation reported in the literature. However, we acknowledge that the borderlines between different types of systematic reviews are often blurred in practice [94-96].

A striking finding was that none of the included studies correlated individual game elements with the reported outcomes. Instead, outcomes were generally attributed to the simulation as a whole. Consequently, evidence cannot be synthesized quantitatively on the level of specific game elements' effectiveness. This should be interpreted as a finding about the state of the literature, rather than a limitation of the systematic review methodology.

When initiating this review, the original intention was to identify studies that explicitly addressed gamification in the context of military trauma training. Acknowledging that the literature on this topic may be sparse, the research questions were adapted to instead explore the use of game elements in the design of simulations in this context. This approach was based on the

understanding that even in the absence of an explicitly acknowledged gamification strategy, game elements may still be purposefully or unintentionally embedded in the instructional design of simulations, potentially influencing the learning outcomes. As anticipated, direct comparisons between gamified and nongamified conditions or otherwise isolated reported effects of specific game elements were not observed.

We conducted a quality assessment of the included quantitative, qualitative, and mixed methods studies using 2 established tools: the MERSQI [38] and the Côté and Turgeon grid [39]. Design and development studies were not assessed using these instruments because they fall outside their intended scope. The appraisal was used descriptively to characterize methodological features and limitations across the included studies and to support interpretation of the evidence base, rather than to weight the thematic synthesis. This approach is consistent with review typologies describing how quality assessment may be used to mediate interpretation in heterogeneous evidence syntheses rather than determine theme inclusion or exclusion [96]. While some studies scored low on these metrics, this should not necessarily be interpreted as a reflection of poor study design. Rather, it may highlight an imperfect match between the quality appraisal tools and the purpose of some of the included studies.

The substantial heterogeneity across the included studies represents another limitation. The wide variation in study designs, simulation types, and learner populations poses significant challenges for synthesizing the data. Heterogeneity was examined descriptively by comparing study designs, simulation modalities, learner populations, and the implementation and stated purpose of game elements, rather than by analyzing heterogeneity of effect estimates. Additionally, although we collaborated with professional librarians to develop a comprehensive search strategy and conducted a reference and citation search, it is possible that we might have missed relevant studies since several of the game elements identified in this review were not part of the original search vocabulary.

Finally, because this review is based on secondary data, we were limited to what was reported by the study authors. In many cases, the motivations or intentions behind the use of game elements may have existed but were not documented. A more comprehensive understanding could have been achieved through interviews or direct engagement with simulation designers.

Conclusion

This is the first comprehensive synthesis of what and how game elements are applied within military trauma simulations, providing a structured evidence base for more intentional and theory-informed design of educational technologies used in high-stakes medical training. Unlike previous reviews, it explicitly focuses on the pedagogical purposes of these elements. It offers an overview of the prevalence of game elements in military trauma care education and synthesizes the pedagogical rationales for their use. While some elements, such as narrative, sensation, and time pressure, were often used in a way that mimics the austerity of the military trauma setting, game elements like badges and competition were underrepresented. Across the reviewed studies, game elements were typically not

justified in terms of their pedagogical function. When justifications were provided, they were most often linked to environmental fidelity and immersion, followed by intentions to provide adaptive learning and feedback. None of the included studies in this review correlated individual game elements with the reported learning outcomes. These findings highlight the need for more intentional research on gamification design and

transparent reporting, in which the educational purpose of each game element is clearly articulated. Future studies should treat gamification as a set of targeted design choices rather than as a single overarching strategy and further explore how its playful and motivational dimensions can be effectively leveraged in military trauma training to support motivation and learning.

Acknowledgments

The authors would like to acknowledge Narcisa Hannerz, Jonas Pettersson, and Anja Vikingson at Karolinska Institutet Biblioteket (university library) for their support in designing the search strategy, conducting the reference and citation searches, and retrieving the articles. The authors declare the use of generative artificial intelligence in the research and writing process. According to the GAIDeT (Generative AI Delegation Taxonomy) [97], the following tasks were assisted by generative artificial intelligence tools under full human supervision: proofreading and editing. The generative artificial intelligence tool used was GPT-4o. Responsibility for the final manuscript lies entirely with the authors. Generative artificial intelligence tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This systematic review was financially supported by the Swedish Armed Forces. The funder was not involved in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The data sets generated during or analyzed during this systematic review are available from the corresponding author on reasonable request.

Authors' Contributions

NS designed the systematic review with guidance from AAK and KK. NS conducted the drafting of the manuscript, which was reviewed by all coauthors, who provided suggestions and edits. Screening, data extraction, and synthesis were carried out by all authors, with each author working in pairs together with NS.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA 2020 checklist.

[\[DOCX File, 273 KB - games_v14i1e79163_app1.docx \]](#)

Multimedia Appendix 2

PRISMA-S checklist.

[\[DOCX File, 17 KB - games_v14i1e79163_app2.docx \]](#)

Multimedia Appendix 3

PRISMA 2020 abstract checklist.

[\[PDF File \(Adobe PDF File\), 122 KB - games_v14i1e79163_app3.pdf \]](#)

Multimedia Appendix 4

SWiM Checklist.

[\[DOCX File, 21 KB - games_v14i1e79163_app4.docx \]](#)

Multimedia Appendix 5

Search strategy.

[\[PDF File \(Adobe PDF File\), 279 KB - games_v14i1e79163_app5.pdf \]](#)

Multimedia Appendix 6

Game elements per study.

[[PDF File \(Adobe PDF File\), 103 KB - games_v14i1e79163_app6.pdf](#)]

Multimedia Appendix 7

Definition of game elements.

[[PDF File \(Adobe PDF File\), 39 KB - games_v14i1e79163_app7.pdf](#)]

Multimedia Appendix 8

Justification of game elements.

[[PDF File \(Adobe PDF File\), 94 KB - games_v14i1e79163_app8.pdf](#)]

References

1. Pfeiffer A, Bezzina S, König N, Kriglstein S. Beyond classical gamification: in- and around-game gamification for education. 2020 Presented at: ECEL 19th European Conference on e-Learning; October 29-30, 2020; Virtual (Originally Berlin, Germany) p. 28-30.
2. Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, et al. Serious gaming and gamification education in health professions: systematic review. *J Med Internet Res* 2019;21(3):e12994 [FREE Full text] [doi: [10.2196/12994](#)] [Medline: [30920375](#)]
3. Maheu-Cadotte M, Cossette S, Dubé V, Fontaine G, Mailhot T, Lavoie P, et al. Effectiveness of serious games and impact of design elements on engagement and educational outcomes in healthcare professionals and students: a systematic review and meta-analysis protocol. *BMJ Open* 2018;8(3):e019871 [FREE Full text] [doi: [10.1136/bmjopen-2017-019871](#)] [Medline: [29549206](#)]
4. Krath J, Schürmann L, von Korfflesch HF. Revealing the theoretical basis of gamification: a systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Comput Hum Behav* 2021;125:106963. [doi: [10.1016/j.chb.2021.106963](#)]
5. van Gaalen AEJ, Brouwer J, Schönrock-Adema J, Bouwkamp-Timmer T, Jaarsma ADC, Georgiadis JR. Gamification of health professions education: a systematic review. *Adv Health Sci Educ Theory Pract* 2021;26(2):683-711 [FREE Full text] [doi: [10.1007/s10459-020-10000-3](#)] [Medline: [33128662](#)]
6. Lee CY, Lee CH, Lai HY, Chen PJ, Chen MM, Yau SY. Emerging trends in gamification for clinical reasoning education: a scoping review. *BMC Med Educ* 2025;25(1):435 [FREE Full text] [doi: [10.1186/s12909-025-07044-7](#)] [Medline: [40133879](#)]
7. Bai S, Zeng H, Zhong Q, Shen Y, Cao L, He M. Application of gamification teaching in disaster education: scoping review. *JMIR Serious Games* 2024;12:e64939 [FREE Full text] [doi: [10.2196/64939](#)] [Medline: [39661070](#)]
8. Mazarakis A, Bräuer P. Gamification is working, but which one exactly? results from an experiment with four game design elements. *Int J Hum Comput Interact* 2022;39(3):612-627. [doi: [10.1080/10447318.2022.2041909](#)]
9. Qin J, Chui YP, Pang WM, Choi KS, Heng PA. Learning blood management in orthopedic surgery through gameplay. *IEEE Comput Graph Appl* 2010;30(2):45-57. [doi: [10.1109/MCG.2009.83](#)] [Medline: [20650710](#)]
10. Ricciardi F, De Paolis LT. A comprehensive review of serious games in health professions. *Int J Comput Games Technol* 2014;2014:1-11. [doi: [10.1155/2014/787968](#)]
11. Pasquier P, Mérat S, Malgras B, Petit L, Queran X, Bay C, et al. A serious game for massive training and assessment of French soldiers involved in forward combat casualty care (3D-SC1): development and deployment. *JMIR Serious Games* 2016 May 18;4(1):e5 [FREE Full text] [doi: [10.2196/games.5340](#)] [Medline: [27194369](#)]
12. Planchon J, Vacher A, Comblet J, Rabatel E, Darses F, Mignon A, et al. Serious game training improves performance in combat life-saving interventions. *Injury* 2018;49(1):86-92. [doi: [10.1016/j.injury.2017.10.025](#)] [Medline: [29042033](#)]
13. DeFalco JA, Rowe JP, Paquette L, Georgoulas-Sherry V, Brawner K, Mott BW, et al. Detecting and addressing frustration in a serious game for military training. *Int J Artif Intell Educ* 2017;28(2):152-193. [doi: [10.1007/s40593-017-0152-1](#)]
14. Sotomayor TM. Teaching tactical combat casualty care using the TC3 sim game-based simulation: a study to measure training effectiveness. *Stud Health Technol Inform* 2010;154:176-179. [Medline: [20543293](#)]
15. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 2005;27(1):10-28. [doi: [10.1080/01421590500046924](#)] [Medline: [16147767](#)]
16. Quick JA. Simulation training in trauma. *Mo Med* 2018;115(5):447-450 [FREE Full text] [Medline: [30385994](#)]
17. Stathakarou N, Sonesson L, Lundberg L, Boffard KD, Kononowicz AA, Karlgren K. Teams managing civilian and military complex trauma: what are the competencies required in austere environments and the potential of simulation technology to address them? *Health Informatics J* 2021;27(4):14604582211052253 [FREE Full text] [doi: [10.1177/14604582211052253](#)] [Medline: [34821149](#)]
18. Caffery SJ, Ferrari BD, Hackett MG. Military medical simulations-scoping review. *Mil Med* 2025;190(3-4):e554-e560. [doi: [10.1093/milmed/usae468](#)] [Medline: [39487968](#)]

19. Laporta AJ, Hoang T, Moloff A. From trauma in austere environments to combat or medical school: how blended hyper-realism in the real and virtual worlds can better prepare surgeons. *Stud Health Technol Inform* 2014;196:233-237. [doi: [10.3233/978-1-61499-375-9-233](https://doi.org/10.3233/978-1-61499-375-9-233)]
20. Kubala AL, Warnick WL. *A Review of Selected Literature on Stresses Affecting Soldiers in Combat*. Alexandria (VA): US Army Research Institute for the Behavioral and Social Sciences; 1979.
21. Gerhardt RT, Mabry RL, De Lorenzo RA, Butler FK. *Fundamentals of Combat Casualty Care*. Texas: Borden Institute, US Army Medical Department; 2012.
22. Leitch RA, Moses GR, Magee H. Simulation and the future of military medicine. *Mil Med* 2002;167(4):350-354 [FREE Full text] [doi: [10.1093/milmed/167.4.350](https://doi.org/10.1093/milmed/167.4.350)]
23. Gruen RL, Jurkovich GJ, McIntyre LK, Foy HM, Maier RV. Patterns of errors contributing to trauma mortality: lessons learned from 2,594 deaths. *Ann Surg* 2006;244(3):371-380. [doi: [10.1097/01.sla.0000234655.83517.56](https://doi.org/10.1097/01.sla.0000234655.83517.56)] [Medline: [16926563](https://pubmed.ncbi.nlm.nih.gov/16926563/)]
24. Khorram-Manesh A, Lönroth H, Rotter P, Wilhelmsson M, Aremyr J, Berner A, et al. Non-medical aspects of civilian-military collaboration in management of major incidents. *Eur J Trauma Emerg Surg* 2017;43(5):595-603 [FREE Full text] [doi: [10.1007/s00068-017-0778-6](https://doi.org/10.1007/s00068-017-0778-6)] [Medline: [28255610](https://pubmed.ncbi.nlm.nih.gov/28255610/)]
25. Khorram-Manesh A. Facilitators and constrainers of civilian-military collaboration: the Swedish perspectives. *Eur J Trauma Emerg Surg* 2020;46(3):649-656 [FREE Full text] [doi: [10.1007/s00068-018-1058-9](https://doi.org/10.1007/s00068-018-1058-9)] [Medline: [30542745](https://pubmed.ncbi.nlm.nih.gov/30542745/)]
26. Haverkamp FJC, van Leest TAJ, Muhrbeck M, Hoencamp R, Wladis A, Tan ECTH. Self-perceived preparedness and training needs of healthcare personnel on humanitarian mission: a pre- and post-deployment survey. *World J Emerg Surg* 2022;17(1):14 [FREE Full text] [doi: [10.1186/s13017-022-00417-z](https://doi.org/10.1186/s13017-022-00417-z)] [Medline: [35248111](https://pubmed.ncbi.nlm.nih.gov/35248111/)]
27. Givens M, Holcomb JB. Red line the red line: optimizing emergency medicine physicians and surgeons collaborative roles on trauma teams. *J Trauma Acute Care Surg* 2024;97(2S Suppl 1):S27-S30. [doi: [10.1097/TA.0000000000004409](https://doi.org/10.1097/TA.0000000000004409)] [Medline: [38815244](https://pubmed.ncbi.nlm.nih.gov/38815244/)]
28. Anagnostou E, Michas A, Giannou C. Practicing military medicine in truly austere environments: what to expect, how to prepare, when to improvise. *Mil Med* 2020;185(5-6):e656-e661. [doi: [10.1093/milmed/usz467](https://doi.org/10.1093/milmed/usz467)] [Medline: [31990026](https://pubmed.ncbi.nlm.nih.gov/31990026/)]
29. Ellaway RH. A conceptual framework of game-informed principles for health professions education. *Adv Simul (Lond)* 2016;1:28 [FREE Full text] [doi: [10.1186/s41077-016-0030-1](https://doi.org/10.1186/s41077-016-0030-1)] [Medline: [29449997](https://pubmed.ncbi.nlm.nih.gov/29449997/)]
30. Stathakarou N, Kononowicz AA, Mattsson E, Karlgren K. Gamification in the design of virtual patients for Swedish military medics to support trauma training: interaction analysis and semistructured interview study. *JMIR Serious Games* 2024;12:e63390 [FREE Full text] [doi: [10.2196/63390](https://doi.org/10.2196/63390)] [Medline: [39436692](https://pubmed.ncbi.nlm.nih.gov/39436692/)]
31. Toda AM, Klock ACT, Oliveira W, Palomino PT, Rodrigues L, Shi L, et al. Analysing gamification elements in educational environments using an existing gamification taxonomy. *Smart Learn Environ* 2019;6(1):1-17. [doi: [10.1186/s40561-019-0106-1](https://doi.org/10.1186/s40561-019-0106-1)]
32. Stathakarou N, Kononowicz AA, Swain C, Karlgren K. Game elements in the design of simulations in military trauma management training: protocol for a systematic review. *JMIR Res Protoc* 2023;12:e45969 [FREE Full text] [doi: [10.2196/45969](https://doi.org/10.2196/45969)] [Medline: [37682596](https://pubmed.ncbi.nlm.nih.gov/37682596/)]
33. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71 [FREE Full text] [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
34. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc* 2016 Jul;104(3):240-243 [FREE Full text] [doi: [10.3163/1536-5050.104.3.014](https://doi.org/10.3163/1536-5050.104.3.014)] [Medline: [27366130](https://pubmed.ncbi.nlm.nih.gov/27366130/)]
35. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016;5(1):210 [FREE Full text] [doi: [10.1186/s13643-016-0384-4](https://doi.org/10.1186/s13643-016-0384-4)] [Medline: [27919275](https://pubmed.ncbi.nlm.nih.gov/27919275/)]
36. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;8:45 [FREE Full text] [doi: [10.1186/1471-2288-8-45](https://doi.org/10.1186/1471-2288-8-45)] [Medline: [18616818](https://pubmed.ncbi.nlm.nih.gov/18616818/)]
37. Kononowicz A, Zary N, Edelbring S, Corral J, Hege I. Virtual patients--what are we talking about? a framework to classify the meanings of the term in healthcare education. *BMC Med Educ* 2015;15:11 [FREE Full text] [doi: [10.1186/s12909-015-0296-3](https://doi.org/10.1186/s12909-015-0296-3)] [Medline: [25638167](https://pubmed.ncbi.nlm.nih.gov/25638167/)]
38. Cook DA, Reed DA. Appraising the quality of medical education research methods: the medical education research study quality instrument and the newcastle-Ottawa scale-education. *Acad Med* 2015;90(8):1067-1076. [doi: [10.1097/ACM.0000000000000786](https://doi.org/10.1097/ACM.0000000000000786)] [Medline: [26107881](https://pubmed.ncbi.nlm.nih.gov/26107881/)]
39. Côté L, Turgeon J. Appraising qualitative research articles in medicine and medical education. *Med Teach* 2005;27(1):71-75. [doi: [10.1080/01421590400016308](https://doi.org/10.1080/01421590400016308)] [Medline: [16147774](https://pubmed.ncbi.nlm.nih.gov/16147774/)]
40. Achatz G, Friemert B, Trentzsch H, Hofmann M, Blätzing M, Hinz-Bauer S, Deployment, Disaster, Tactical Surgery Working Group of the German Trauma Society. Terror and disaster surgical care: training experienced trauma surgeons in decision making for a MASCAL situation with a tabletop simulation game. *Eur J Trauma Emerg Surg* 2020;46(4):717-724. [doi: [10.1007/s00068-020-01441-x](https://doi.org/10.1007/s00068-020-01441-x)] [Medline: [32725275](https://pubmed.ncbi.nlm.nih.gov/32725275/)]
41. Arora S, Cox C, Davies S, Kassab E, Mahoney P, Sharma E, et al. Towards the next frontier for simulation-based training: full-hospital simulation across the entire patient pathway. *Ann Surg* 2014;260(2):252-258. [doi: [10.1097/SLA.0000000000000305](https://doi.org/10.1097/SLA.0000000000000305)] [Medline: [24263325](https://pubmed.ncbi.nlm.nih.gov/24263325/)]

42. Badler NI, Clarke JR, Hollick M, Kokkevis E, Metaxas D, Bindiwanavale R. MediSim: simulated medical corpsmen and casualties for medical forces planning and training. 1995 Presented at: Proceedings of the National Forum: Military Telemedicine On-Line Today Research, Practice, and Opportunities; 1995 March 27-29; McLean, VA, USA p. 21-28. [doi: [10.1109/mtol.1995.504523](https://doi.org/10.1109/mtol.1995.504523)]
43. Baird A, Serio-Melvin M, Hackett M, Clover M, McDaniel M, Rowland M, et al. BurnCare tablet trainer to enhance burn injury care and treatment. *BMC Emerg Med* 2020;20(1):84 [FREE Full text] [doi: [10.1186/s12873-020-00378-z](https://doi.org/10.1186/s12873-020-00378-z)] [Medline: [33126858](https://pubmed.ncbi.nlm.nih.gov/33126858/)]
44. Beaven A, Griffin D, James H. Highly realistic cadaveric trauma simulation of the multiply injured battlefield casualty: an international, multidisciplinary exercise in far-forward surgical management. *Injury* 2021;52(5):1183-1189. [doi: [10.1016/j.injury.2020.09.023](https://doi.org/10.1016/j.injury.2020.09.023)] [Medline: [33039179](https://pubmed.ncbi.nlm.nih.gov/33039179/)]
45. Brown R, McIlwain S, Willson B, Hackett M. Enhancing combat medic training with 3D virtual environments. : IEEE; 2016 Presented at: IEEE International Conference on Serious Games and Applications for Health (SeGAH); 2016 May 11-13; Orlando, FL, USA p. 1-7. [doi: [10.1109/segah.2016.7586266](https://doi.org/10.1109/segah.2016.7586266)]
46. Chi DM, Clarke JR, Webber BL, Badler NI. Casualty modeling for real-time medical training. *Presence (Camb)* 1996;5(4):359-366. [doi: [10.1162/pres.1996.5.4.359](https://doi.org/10.1162/pres.1996.5.4.359)] [Medline: [11539375](https://pubmed.ncbi.nlm.nih.gov/11539375/)]
47. Chi DM, Kokkevis E, Ogunyemi O, Bindiganavale R, Hollick MJ, Clarke JR. Simulated casualties and medics for emergency training. *ScholarlyCommons*. 1997. URL: <https://repository.upenn.edu/hms/25> [accessed 2026-02-19]
48. Cohen D, Sevdalis N, Taylor D, Kerr K, Heys M, Willett K, et al. Emergency preparedness in the 21st century: training and preparation modules in virtual environments. *Resuscitation* 2013;84(1):78-84. [doi: [10.1016/j.resuscitation.2012.05.014](https://doi.org/10.1016/j.resuscitation.2012.05.014)] [Medline: [22659598](https://pubmed.ncbi.nlm.nih.gov/22659598/)]
49. Couperus K. Virtual reality trauma simulation: an immersive method to enhance medical personnel training and readiness. *Ann Emerg Med* 2019;74(4):S30. [doi: [10.1016/j.annemergmed.2019.08.081](https://doi.org/10.1016/j.annemergmed.2019.08.081)]
50. Couperus K, Young S, Walsh R, Kang C, Skinner C, Essendrop R, et al. Immersive virtual reality medical simulation: autonomous trauma training simulator. *Cureus* 2020;12(5):e8062 [FREE Full text] [doi: [10.7759/cureus.8062](https://doi.org/10.7759/cureus.8062)] [Medline: [32542120](https://pubmed.ncbi.nlm.nih.gov/32542120/)]
51. de Lesquen H, de LVB, Cotte J, Avaro JP. Immersive virtual reality to learn triage: a serious game to be prepared to prehospital massive casualty management. *JMIR Preprints* . [doi: [10.2196/preprints.36724](https://doi.org/10.2196/preprints.36724)]
52. de Lesquen H, Paris R, Fournier M, Cotte J, Vacher A, Schlienger D, et al. Toward a serious game to help future military doctors face mass casualty incidents. *J Spec Oper Med* 2023;23(2):88-93. [doi: [10.55460/IJCP-BLY6](https://doi.org/10.55460/IJCP-BLY6)] [Medline: [37126778](https://pubmed.ncbi.nlm.nih.gov/37126778/)]
53. Du W, Zhong X, Jia Y, Jiang R, Yang H, Ye Z, et al. A novel scenario-based, mixed-reality platform for training nontechnical skills of battlefield first aid: prospective interventional study. *JMIR Serious Games* 2022;10(4):e40727 [FREE Full text] [doi: [10.2196/40727](https://doi.org/10.2196/40727)] [Medline: [36472903](https://pubmed.ncbi.nlm.nih.gov/36472903/)]
54. Freeman KM, Thompson SF, Allely EB, Sobel AL, Stansfield SA, Pugh WM. A virtual reality patient simulation system for teaching emergency response skills to U.S. Navy medical providers. *Prehosp Disaster Med* 2001;16(1):3-8. [doi: [10.1017/s1049023x00025462](https://doi.org/10.1017/s1049023x00025462)] [Medline: [11367936](https://pubmed.ncbi.nlm.nih.gov/11367936/)]
55. Goolsby C, Vest R, Goodwin T. New wide area virtual environment (WAVE) medical education. *Mil Med* 2014;179(1):38-41. [doi: [10.7205/MILMED-D-13-00310](https://doi.org/10.7205/MILMED-D-13-00310)] [Medline: [24402983](https://pubmed.ncbi.nlm.nih.gov/24402983/)]
56. Hemman EA. Improving combat medic learning using a personal computer-based virtual training simulator. *Mil Med* 2005;170(9):723-727. [doi: [10.7205/milmed.170.9.723](https://doi.org/10.7205/milmed.170.9.723)] [Medline: [16261973](https://pubmed.ncbi.nlm.nih.gov/16261973/)]
57. Henderson JV, Pruet RK, Galper AR, Copes WS. Interactive videodisc to teach combat trauma life support. *J Med Syst* 1986;10(3):271-276. [doi: [10.1007/BF00992822](https://doi.org/10.1007/BF00992822)] [Medline: [3537185](https://pubmed.ncbi.nlm.nih.gov/3537185/)]
58. Henderson N, Rowe J, Paquette L, Baker R, Lester J. Improving affect detection in game-based learning with multimodal data fusion. *Artif Intell Educ* 2020;12163:228-239. [doi: [10.1007/978-3-030-52237-7_19](https://doi.org/10.1007/978-3-030-52237-7_19)]
59. Kyle RR, Via DK, Lowy RJ, Madsen JM, Marty AM, Mongan PD. A multidisciplinary approach to teach responses to weapons of mass destruction and terrorism using combined simulation modalities. *J Clin Anesth* 2004;16(2):152-158. [doi: [10.1016/j.jclinane.2003.09.003](https://doi.org/10.1016/j.jclinane.2003.09.003)] [Medline: [15110382](https://pubmed.ncbi.nlm.nih.gov/15110382/)]
60. Lombardo R, Walther N, Young S, Gorbalkin C, Sletten Z, Kang C, et al. Ready medic one: a feasibility study of a semi-autonomous virtual reality trauma simulator. *Front Virtual Real* 2022;2. [doi: [10.3389/frvir.2021.719656](https://doi.org/10.3389/frvir.2021.719656)]
61. Lu J, Leng A, Zhou Y, Zhou W, Luo J, Chen X, et al. An innovative virtual reality training tool for the pre-hospital treatment of cranialmaxillofacial trauma. *Comput Assist Surg (Abingdon)* 2023;28(1):2189047 [FREE Full text] [doi: [10.1080/24699322.2023.2189047](https://doi.org/10.1080/24699322.2023.2189047)] [Medline: [36974947](https://pubmed.ncbi.nlm.nih.gov/36974947/)]
62. Lennquist Montán K, Hreckovski B, Dobson B, Örténwall P, Montán C, Khorram-Manesh A, et al. Development and evaluation of a new simulation model for interactive training of the medical response to major incidents and disasters. *Eur J Trauma Emerg Surg* 2014;40(4):429-443. [doi: [10.1007/s00068-013-0350-y](https://doi.org/10.1007/s00068-013-0350-y)] [Medline: [26816238](https://pubmed.ncbi.nlm.nih.gov/26816238/)]
63. Netzer I, Weiss A, Hoppenstein D. Naval casualty management training using human patient simulators. *Disaster Mil Med* 2015;1:9 [FREE Full text] [doi: [10.1186/2054-314X-1-9](https://doi.org/10.1186/2054-314X-1-9)] [Medline: [28265424](https://pubmed.ncbi.nlm.nih.gov/28265424/)]
64. Qin L, You Z, Liu B, Luo C. Development of virtual reality training system for combat musculoskeletal trauma care. *Simulation* 2024;101(1):3-11. [doi: [10.1177/00375497241238025](https://doi.org/10.1177/00375497241238025)]

65. Rabotin A, Glick Y, Gelman R, Ketko I, Taran B, Fink N, et al. Practicing emergency medicine in the metaverse: a novel mixed reality casualty care training platform. *Surg Innov* 2023;30(5):586-594. [doi: [10.1177/15533506231191576](https://doi.org/10.1177/15533506231191576)] [Medline: [37496258](https://pubmed.ncbi.nlm.nih.gov/37496258/)]
66. Satava RM, Jones SB. An integrated medical virtual reality program. the military application. *IEEE Eng Med Biol Mag* 1996;15(2):94-97, 104. [doi: [10.1109/51.486724](https://doi.org/10.1109/51.486724)]
67. Sonesson L, Boffard KD, Örténwall P, Vekzslar P. Determining the educational impact of virtual patients on trauma team training during a multinational, large-scale civil military simulation exercise. *J Trauma Acute Care Surg* 2023;95(2S Suppl 1):S99-S105 [FREE Full text] [doi: [10.1097/TA.0000000000004017](https://doi.org/10.1097/TA.0000000000004017)] [Medline: [37163456](https://pubmed.ncbi.nlm.nih.gov/37163456/)]
68. Sotomayor T. Evaluating Tactical Combat Casualty Care Training Treatments Effects on Combat Medic Trainees in Light of Select Human Descripti [Dissertation]. Orlando (FL): University of Central Florida; 2008.
69. Stansfield S, Shawver D, Sobel A. MediSim: a prototype VR system for training medical first responders. : IEEE; 1998 Presented at: Proceedings of IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180); 1998 March 14-18; Atlanta, GA, USA p. 198-205. [doi: [10.1109/vrais.1998.658490](https://doi.org/10.1109/vrais.1998.658490)]
70. Stone RJ. Serious gaming - virtual reality's saviour? 2005 Presented at: Proceedings of Virtual Systems and Multi-Media (VSMM) Conference; 2005 October 3-7; Ghent, Belgium p. 773-786.
71. Stone RJ. The (human) science of medical virtual learning environments. *Philos Trans R Soc Lond B Biol Sci* 2011;366(1562):276-285 [FREE Full text] [doi: [10.1098/rstb.2010.0209](https://doi.org/10.1098/rstb.2010.0209)] [Medline: [21149363](https://pubmed.ncbi.nlm.nih.gov/21149363/)]
72. Stone RJ, Guest R, Mahoney P, Lamb D, Gibson C. A 'mixed reality' simulator concept for future medical emergency response team training. *J R Army Med Corps* 2017;163(4):280-287 [FREE Full text] [doi: [10.1136/jramc-2016-000726](https://doi.org/10.1136/jramc-2016-000726)] [Medline: [28062529](https://pubmed.ncbi.nlm.nih.gov/28062529/)]
73. Tretyak VE. TacMedVR: immersive VR training for tactical medicine—evaluating interaction and stress response. : IEEE; 2025 Presented at: 11th International Conference on Virtual Reality (ICVR); 2025 July 09-11; Wageningen, Netherlands p. 345-350. [doi: [10.1109/icvr66534.2025.11172647](https://doi.org/10.1109/icvr66534.2025.11172647)]
74. Wier GS, Tree R, Nusr R. Training effectiveness of a wide area virtual environment in medical simulation. *Simul Healthc* 2017;12(1):28-40. [doi: [10.1097/SIH.0000000000000207](https://doi.org/10.1097/SIH.0000000000000207)] [Medline: [28146451](https://pubmed.ncbi.nlm.nih.gov/28146451/)]
75. Willy C, Sterk J, Schwarz W, Gerngross H. Computer-assisted training program for simulation of triage, resuscitation, and evacuation of casualties. *Mil Med* 1998;163(4):234-238. [doi: [10.1093/milmed/163.4.234](https://doi.org/10.1093/milmed/163.4.234)]
76. Zhu S, Li Z, Sun Y, Kong L, Yin M, Yong Q, et al. A serious game for enhancing rescue reasoning skills in tactical combat casualty care: development and deployment study. *JMIR Form Res* 2024;8:e50817 [FREE Full text] [doi: [10.2196/50817](https://doi.org/10.2196/50817)] [Medline: [39133911](https://pubmed.ncbi.nlm.nih.gov/39133911/)]
77. Smith R. The long history of gaming in military training. *Simul Gaming* 2009;41(1):6-19. [doi: [10.1177/1046878109334330](https://doi.org/10.1177/1046878109334330)]
78. Wilson KA, Bedwell WL, Lazzara EH, Salas E, Burke CS, Estock JL, et al. Relationships between game attributes and learning outcomes. *Simul Gaming* 2008;40(2):217-266. [doi: [10.1177/1046878108321866](https://doi.org/10.1177/1046878108321866)]
79. Cook DA, Bordage G, Schmidt HG. Description, justification and clarification: a framework for classifying the purposes of research in medical education. *Med Educ* 2008;42(2):128-133. [doi: [10.1111/j.1365-2923.2007.02974.x](https://doi.org/10.1111/j.1365-2923.2007.02974.x)] [Medline: [18194162](https://pubmed.ncbi.nlm.nih.gov/18194162/)]
80. Maheu-Cadotte M, Cossette S, Dubé V, Fontaine G, Deschênes M, Lapierre A, et al. Differentiating the design principles of virtual simulations and serious games to enhance nurses' clinical reasoning. *Clin Simul Nurs* 2020;49:19-23. [doi: [10.1016/j.ecns.2020.04.004](https://doi.org/10.1016/j.ecns.2020.04.004)]
81. Caillois R. *Man, play and games*. Urbana (IL): University of Illinois Press; 2001.
82. Garris R, Ahlers R, Driskell JE. Games, motivation, and learning: a research and practice model. *Simul Gaming* 2002;33(4):441-467. [doi: [10.1177/1046878102238607](https://doi.org/10.1177/1046878102238607)]
83. Mayer RE, Dow GT, Mayer S. Multimedia learning in an interactive self-explaining environment: what works in the design of agent-based microworlds? *J Educ Psychol* 2003;95(4):806-812. [doi: [10.1037/0022-0663.95.4.806](https://doi.org/10.1037/0022-0663.95.4.806)]
84. Van Eck R, Dempsey J. The effect of competition and contextualized advisement on the transfer of mathematics skills a computer-based instructional simulation game. *Educ Technol Res Dev* 2002;50(3):23-41. [doi: [10.1007/bf02505023](https://doi.org/10.1007/bf02505023)]
85. Zhou T, Loh CS. The effects of fully and partially in-game guidance on players' declarative and procedural knowledge with a disaster preparedness serious game. In: *Research Anthology on Game Design, Development, Usage, and Social Impact*. New York: IGI Global; 2023:1818-1834.
86. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 1993;100(3):363-406. [doi: [10.1037/0033-295x.100.3.363](https://doi.org/10.1037/0033-295x.100.3.363)]
87. Malone TW, Lepper MR. Making learning fun: a taxonomy of intrinsic motivations for learning. In: Snow RE, Farr MJ, editors. *Aptitude, Learning, and Instruction*. Vol 3: Conative and Affective Process Analyses. Hillsdale (NJ): Lawrence Erlbaum Associates; 1987:223-253.
88. Johnson DW, Johnson RT. *Learning Together and Alone: Cooperative, Competitive, and Individualistic Learning*. Boston (MA): Allyn & Bacon; 1994.
89. Choi W, Dyens O, Chan T, Schijven M, Lajoie S, Mancini ME, et al. Engagement and learning in simulation: recommendations of the simnovate engaged learning domain group. *BMJ Simul Technol Enhanc Learn* 2017;3(Suppl 1):S23-S32. [doi: [10.1136/bmjstel-2016-000177](https://doi.org/10.1136/bmjstel-2016-000177)]

90. Gray M, Baird A, Sawyer T, James J, DeBroux T, Bartlett M, et al. Increasing realism and variety of virtual patient dialogues for prenatal counseling education through a novel application of CHATGPT: exploratory observational study. *JMIR Med Educ* 2024;10:e50705 [FREE Full text] [doi: [10.2196/50705](https://doi.org/10.2196/50705)] [Medline: [38300696](https://pubmed.ncbi.nlm.nih.gov/38300696/)]
91. Chang CC, Liang C, Chou PN, Lin GY. Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? perspective from multimedia and media richness. *Comput Hum Behav* 2017;71:218-227. [doi: [10.1016/j.chb.2017.01.031](https://doi.org/10.1016/j.chb.2017.01.031)]
92. Bai S, Zeng H, Zhong Q, Cao L, He M. Effectiveness of gamified teaching in disaster nursing education for health care workers: systematic review. *J Med Internet Res* 2025;27:e74955 [FREE Full text] [doi: [10.2196/74955](https://doi.org/10.2196/74955)] [Medline: [40635388](https://pubmed.ncbi.nlm.nih.gov/40635388/)]
93. Stathakarou N, Kononowicz AA, Harjani M, Reshetukha D, Mattsson E, Karlgren K. Exploring the potential of gamified virtual patients for military trauma care training: a systematic text condensation analysis. *Injury* 2026;113020 [FREE Full text] [doi: [10.1016/j.injury.2026.113020](https://doi.org/10.1016/j.injury.2026.113020)] [Medline: [41545241](https://pubmed.ncbi.nlm.nih.gov/41545241/)]
94. Smith SA, Duncan AA. Systematic and scoping reviews: a comparison and overview. *Semin Vasc Surg* 2022;35(4):464-469. [doi: [10.1053/j.semvascsurg.2022.09.001](https://doi.org/10.1053/j.semvascsurg.2022.09.001)] [Medline: [36414363](https://pubmed.ncbi.nlm.nih.gov/36414363/)]
95. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol* 2018;18(1):143 [FREE Full text] [doi: [10.1186/s12874-018-0611-x](https://doi.org/10.1186/s12874-018-0611-x)] [Medline: [30453902](https://pubmed.ncbi.nlm.nih.gov/30453902/)]
96. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J* 2009;26(2):91-108 [FREE Full text] [doi: [10.1111/j.1471-1842.2009.00848.x](https://doi.org/10.1111/j.1471-1842.2009.00848.x)] [Medline: [19490148](https://pubmed.ncbi.nlm.nih.gov/19490148/)]
97. Suchikova Y, Tsybuliak N, Teixeira da Silva JA, Nazarovets S. GAIDeT (generative AI delegation taxonomy): a taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. *Account Res* 2025;1-27. [doi: [10.1080/08989621.2025.2544331](https://doi.org/10.1080/08989621.2025.2544331)] [Medline: [40781729](https://pubmed.ncbi.nlm.nih.gov/40781729/)]

Abbreviations

MERSQI: Medical Education Research Study Quality Instrument

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

SWiM: Synthesis Without Meta-analysis

Edited by S Brini; submitted 16.Jun.2025; peer-reviewed by TTO Kwok, L Zhu; comments to author 06.Oct.2025; accepted 14.Feb.2026; published 17.Mar.2026.

Please cite as:

Stathakarou N, Kononowicz AA, Harjani MG, Swain C, Karlgren K
Game Elements in Military Trauma Care Education: Systematic Review
JMIR Serious Games 2026;14:e79163

URL: <https://games.jmir.org/2026/1/e79163>

doi: [10.2196/79163](https://doi.org/10.2196/79163)

PMID:

©Natalia Stathakarou, Andrzej A Kononowicz, Maxine G Harjani, Cara Swain, Klas Karlgren. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 17.Mar.2026. 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.

RedMan-GreenMan: Co-Designed Pedestrian Safety Game Prototype for Children With Autism

Jason Carter Stanton¹, MPhil; Nicole L Peel^{2,3}, MHS, MBA, PhD; Caroline J Mills^{2,3}, MAL, PhD; Paul P Breen^{1,3}, PhD

¹The MARCS Institute for Brain, Behaviour & Development, Western Sydney University, 160 Hawkesbury Road, Westmead, Australia

²School of Health Sciences, Western Sydney University, Campbelltown, Australia

³Translational Health Research Institute, Western Sydney University, Westmead, Australia

Corresponding Author:

Paul P Breen, PhD

The MARCS Institute for Brain, Behaviour & Development, Western Sydney University, 160 Hawkesbury Road, Westmead, Australia

Abstract

This letter presents the conceptualization, design, and technical evaluation of the RedMan-GreenMan game co-designed with carers, aimed to help children with autism spectrum disorder (hereafter autism) acquire pedestrian safety skills. While the system has been implemented and is in active use, no empirical evaluation of learning outcomes or behavioral impact has been conducted to date, and the focus of this work is on system development, functionality, and technical evaluation.

(*JMIR Serious Games* 2026;14:e69260) doi:[10.2196/69260](https://doi.org/10.2196/69260)

KEYWORDS

autism; game; pedestrian; safety; training

Introduction

Children with autism face a higher risk of unintentional injury and death compared to their peers, as challenges with attention, motor control, and cognitive delays in some children can hinder the development of essential safety skills [1,2]. In collaboration with the Skillz4me Family Centre for Disabilities, a not-for-profit organization established to support children with moderate to profound disabilities, it was identified that while some research has been completed in the area [3-5], there are currently no pedestrian road safety programs available that adequately address the specific needs of children with autism.

A specific barrier for children with autism is the difficulty in forming clear associations between related objects [6,7]. For example, a small toy train may not be readily associated with an actual full-sized train. To address this challenge, we co-designed a simple game, RedMan-GreenMan, which importantly integrates a real pedestrian traffic signal. This approach may increase the likelihood of recognition and association by the child, and in turn may support the potential transfer of safety skills from the indoor space and gameplay to the public, real-world environment.

An additional design consideration was the integration of a performance recording system to enable staff to monitor each child's interaction with the game over time.

This paper describes the design and in situ testing of the RedMan-GreenMan system. Testing aimed to validate the technical aspects of the system; it did not include any assessment of children, pedagogical outcomes, or effectiveness of the intervention in teaching safety skills.

Methods

User Requirements

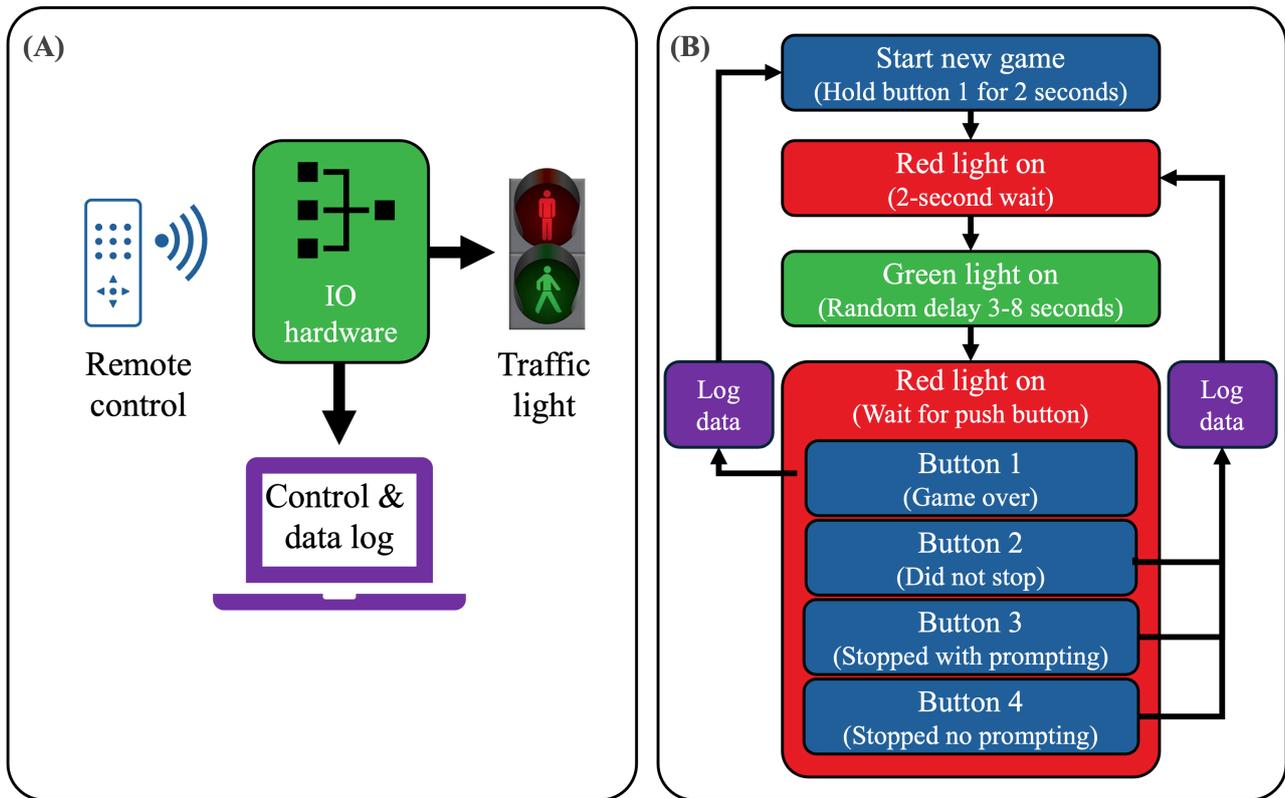
Scope for the design, development, and commissioning of the system was established. Briefly, the user requirements were:

- Use of a real pedestrian traffic light;
- Remote controllable operation of the game;
- Random timing of gameplay in operation, to prevent inadvertent signaling from staff; and
- Date-time stamping of in-game child response.

System Requirements

User requirements led to the system topology and process control as shown in [Figure 1](#).

Figure 1. (A) System topology and (B) process control developed for RedMan-GreenMan. Key requirements included a need for the system to be controllable by a remote control to allow the staff member to interact with children within the center. The categorization of child responses was intentionally limited to three options to simplify decision-making for the busy staff member.



System Evaluation

A series of experiments was conducted to rigorously evaluate system functionality, accuracy, precision, and operational limits.

- System functionality during repeated operation cycles: With the remote positioned 5 meters from the receiver, the program was initiated by pressing and holding the “1” key for 2 seconds. A fixed button selection pattern (2-3-4-3-2) was executed and terminated by pressing the “1” key. This cycle was repeated 40 times to assess system reliability and accuracy under consistent operational conditions.
- Traffic light timing verification: Traffic light change periods were logged to a file and cross-referenced with video recordings for validation. The accuracy and precision of randomly timed light changes (range 3 - 8 s) were analyzed.
- System reliability at increasing range: System reliability was tested with 30 traffic light changes at a 2-meter distance between the remote and receiver. This process was repeated at 5-meter intervals, gradually increasing the distance until complete communication failure was observed.

A single device was used in all testing within the Skillz4me Family Centre for Disabilities with direct line of sight between the remote control and receiver up to a range of 75 m. Distances greater than 75 m lost the line of sight, as measurements were taken in the foyer (80 - 85 m), entry area (90 m), and outside the building (≥ 95 m).

Ethical Considerations

Ethical review was not sought as no human subjects were involved in the research.

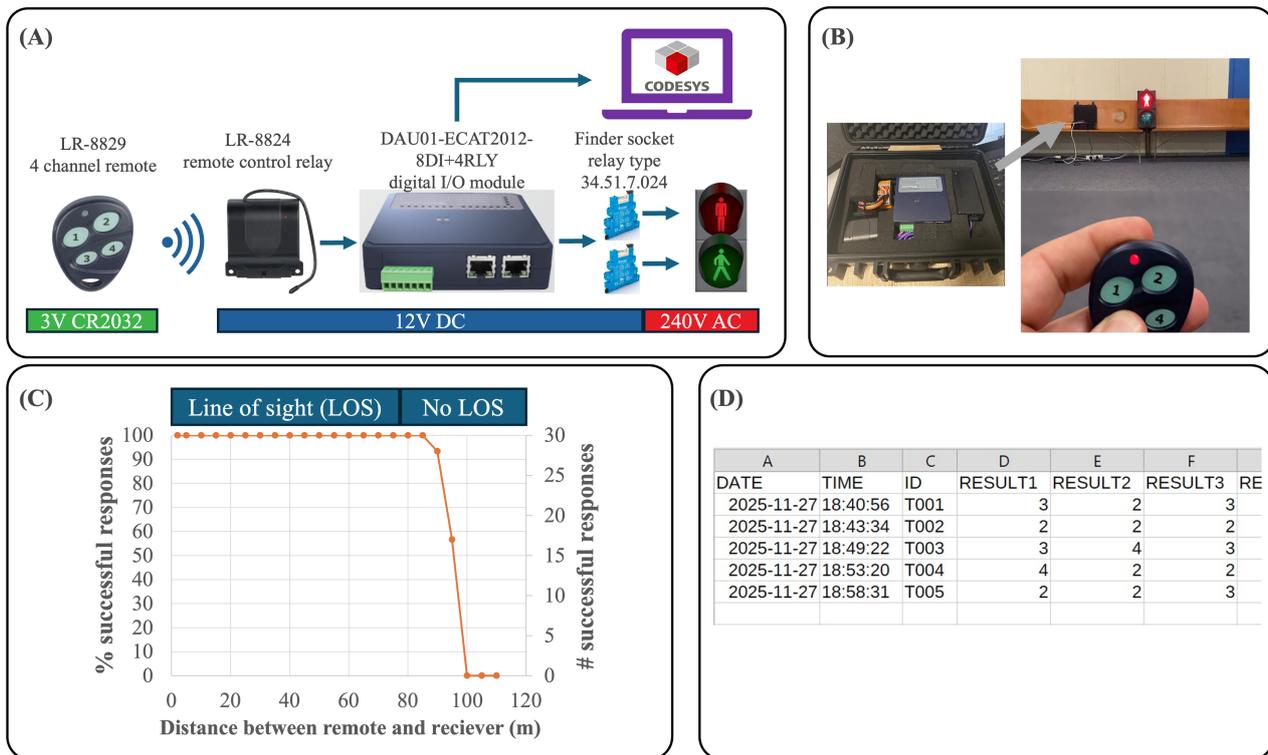
Results

An image of the developed system and its architecture is shown in Figure 2. The system’s performance was evaluated through three distinct experiments. In experiment 1, a predefined sequence pattern was repeated 40 times, achieving 100% accuracy and demonstrating exceptional reliability during repeated operational cycles.

Experiment 2 focused on the accuracy of the traffic light timing, defined as the mean difference between measurements and logged reference values, which was found to be 0.096 seconds, indicating a very slight positive bias. The precision, assessed as the SD of residuals (differences between the secondary measurements and their corresponding references), was 0.072 seconds, reflecting consistent measurement performance.

Experiment 3 assessed the system’s reliability over increasing distances. The system operated with 100% reliability up to a range of 85 m; this includes 10 m (from 75 to 85 m) without a direct line of sight. Beyond this range, performance declined sharply, culminating in total failure at 100 meters. It is important to note that the maximum effective range was constrained by indoor physical limitations, as direct line of sight was obstructed by the layout of the testing environment (Figure 2C). These results highlight the system’s robustness and accuracy under controlled conditions while identifying potential limitations in extended-range applications.

Figure 2. (A) Developed system architecture showing specific components and power requirements. (B) Actual system created: gray arrow points to system electronics next to traffic light. (C) Reliability of system response at increasing range; in order to reach failure, the researcher had to exit the center. (D) Demonstration of a system datalog for 5 consecutive simulated games where the monitored individual (ID) is different for each game (no human subjects involved).



Discussion

This study presents the conceptualization, design, and technical evaluation of RedMan-GreenMan, a gamified intervention co-developed with the Skillz4me Family Centre for Disabilities. The system was designed to support children with autism in acquiring pedestrian safety skills by addressing challenges in associative learning. By incorporating an authentic pedestrian traffic signal, the prototype aims to bridge the perceptual and cognitive divide between the virtual learning environment and real-world traffic scenarios.

Importantly, while the system has been implemented in accordance with user-defined requirements and is currently deployed operationally at the Skillz4me Family Centre for Disabilities, no human data were collected or analyzed for this study, and no educational or pedagogical evaluation has been conducted. The current work should not be interpreted as

evidence of learning impact or behavioral change in the target population.

The system's technical functionality has been validated in situ, but no empirical data is available regarding its effectiveness in teaching safety skills or supporting skill transfer to real-world contexts. Other limitations include the lack of scalability and portability of the prototype, and its construction remains cost-prohibitive (~US \$350 for parts excluding laptop).

Future research should include structured educational trials to rigorously evaluate learning outcomes, skill retention, and transfer to real-world naturalistic settings. Such studies may benefit from mixed methods approaches, including behavioral observations, performance tracking, and input from caregivers or therapists. Parallel system development efforts should focus on improving affordability and portability, with the goal of enabling broader use in diverse educational and therapeutic settings.

Funding

This work received no specific grant from any funding agency.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author. Please see [Multimedia Appendix 1](#) for system hardware design, bill of materials, and software.

Authors' Contributions

JCS led the conceptualization, data curation, original draft, and investigation. PPB led methodology, formal analysis, and visualization. PPB, CJM, and NLP equally assisted with supervision and reviewing and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

System hardware design, bill of materials, and software.

[[PDF File, 668 KB - games_v14i1e69260_app1.pdf](#)]

References

1. Lee LC, Harrington RA, Chang JJ, Connors SL. Increased risk of injury in children with developmental disabilities. *Res Dev Disabil* 2008;29(3):247-255. [doi: [10.1016/j.ridd.2007.05.002](https://doi.org/10.1016/j.ridd.2007.05.002)] [Medline: [17582739](https://pubmed.ncbi.nlm.nih.gov/17582739/)]
2. Wiseman KV, McArdell LE, Bottini SB, Gillis JM. A meta-analysis of safety skill interventions for children, adolescents, and young adults with autism spectrum disorder. *Rev J Autism Dev Disord* 2017 Mar;4(1):39-49. [doi: [10.1007/s40489-016-0096-7](https://doi.org/10.1007/s40489-016-0096-7)]
3. Harriage B, Blair KSC, Miltenberger R. An evaluation of a parent implemented in situ pedestrian safety skills intervention for individuals with autism. *J Autism Dev Disord* 2016 Jun;46(6):2017-2027. [doi: [10.1007/s10803-016-2730-8](https://doi.org/10.1007/s10803-016-2730-8)] [Medline: [26864158](https://pubmed.ncbi.nlm.nih.gov/26864158/)]
4. Dixon DR, Miyake CJ, Nohelty K, Novack MN, Granpeesheh D. Evaluation of an immersive virtual reality safety training used to teach pedestrian skills to children with autism spectrum disorder. *Behav Anal Pract* 2020 Sep;13(3):631-640. [doi: [10.1007/s40617-019-00401-1](https://doi.org/10.1007/s40617-019-00401-1)] [Medline: [32953391](https://pubmed.ncbi.nlm.nih.gov/32953391/)]
5. Steinborn M, Knapp TJ. Teaching an autistic child pedestrian skills. *J Behav Ther Exp Psychiatry* 1982 Dec;13(4):347-351. [doi: [10.1016/0005-7916\(82\)90083-0](https://doi.org/10.1016/0005-7916(82)90083-0)] [Medline: [7166609](https://pubmed.ncbi.nlm.nih.gov/7166609/)]
6. Preissler MA. Associative learning of pictures and words by low-functioning children with autism. *Autism* 2008 May;12(3):231-248. [doi: [10.1177/1362361307088753](https://doi.org/10.1177/1362361307088753)] [Medline: [18445733](https://pubmed.ncbi.nlm.nih.gov/18445733/)]
7. Carr D, Felce J. Teaching picture-to-object relations in picture-based requesting by children with autism: a comparison between error prevention and error correction teaching procedures. *J Intellect Disabil Res* 2008 Apr;52(Pt 4):309-317. [doi: [10.1111/j.1365-2788.2007.01021.x](https://doi.org/10.1111/j.1365-2788.2007.01021.x)] [Medline: [18339093](https://pubmed.ncbi.nlm.nih.gov/18339093/)]

Edited by A Coristine; submitted 12.Jan.2025; peer-reviewed by E Vallefuoco, F Jorge, G Wang; revised version received 27.Nov.2025; accepted 28.Nov.2025; published 15.Jan.2026.

Please cite as:

Stanton JC, Peel NL, Mills CJ, Breen PP

RedMan-GreenMan: Co-Designed Pedestrian Safety Game Prototype for Children With Autism

JMIR Serious Games 2026;14:e69260

URL: <https://games.jmir.org/2026/1/e69260>

doi:[10.2196/69260](https://doi.org/10.2196/69260)

© Jason Carter Stanton, Nicole L Peel, Caroline J Mills, Paul P Breen. Originally published in *JMIR Serious Games* (<https://games.jmir.org>), 15.Jan.2026. 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.

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

<https://www.jmirpublications.com/>