JMIR XR and Spatial Computing (JMXR)

Volume 3 (2026) ISSN 2818-3045 Editor in Chief: Lars Riedemann, MD

Contents

Reviews

Different Delivery Modalities of Virtual Reality Training in Emergency Medicine: Systematic Review (e84310)
Marianna Costa, Martina Codato, Giorgia Brugnera, Silvia Galiazzo, Elisabetta Cantele, Todd Chang, Silvia Bressan. 4

Efficacy of Virtual Reality–Based Mindfulness Interventions: Systematic Review and Meta-Analysis (e90003)
Sarah Barker, Andreas Miles-Novelo, Albert Rizzo, Regina Tuma. 25

Viewpoint

A Conceptual Framework for 4D Holographic Visualization in Gastrointestinal Endoscopy: Toward Externalized Spatiotemporal Cognition (e99446)
Hayato Yamaguchi, Masakatsu Fukuzawa, Takashi Kawai, Fumito Yamanishi, Takahiro Muramatsu, Tadashi Ichimiya, Yasuyuki Kagawa, Kumiko Uchida, Akira Madarame, Takashi Morise, Yoshiya Yamauchi, Shin Kono, Sakiko Naito, Takao Itoi. 41

Research Letters

Exploratory Patterns of Task Difficulty and Visual Feedback in Virtual Reality–Based Upper-Limb Trajectory Training (e94745)
Yu Shi, Sophie Dewil, Zachary Marvin, Noam Harel, Raviraj Nataraj. 49

Immersive Simulation Training in Health Informatics: Qualitative Descriptive Study (e106961)
Jennifer Skelton, Savannah Siebers, Alanna Pelmore. 143

Original Papers

Novel Virtual Reality Methods for the Treatment of Obsessive-Compulsive Disorder Using a User-Centered Design Approach: Focus Group Study (e92410)
Michael Colman, Josie Millar, Bhagyashree Patil, Daniel Finnegan, Mhairi Kristoffersen, Danae Stanton Fraser. 54

Metaverse-Based Psychiatric Consultation for Youths With Mental Health Conditions: Qualitative Descriptive Feasibility Study (e83688)
Mio Ishii, Junichi Fujita, Nao Toyohara, Keiko Ide, Tomoko Moroga, Mizuho Takayama, Takeshi Asami, Tomoyuki Miyazaki. 73

XSL•FO
RenderX

A Virtual Reality Intervention to Promote Uptake of Medications for Opioid Use Disorder in the Emergency Department Following Opioid-Involved Overdose Based on Input From People With Lived Experience: Development Study (e84397)
Zina Trost, Andrea Won, Hari Boyapati, Keri Costa, Josh Davenport, Gail D'Onofrio, Leslie Drager, Phillip Shum, Michael Pierce, Liana Fraenkel, William Soares III. 87

Augmented Reality–Assisted Training Tool for Mental Health Task-Sharers: Pilot Mixed Methods Usability Study (e80711)
Ling Ngiam, Lucas Wozniak, Catherine Dinh-Le, Ayanna Seals, Victoria Ngo, Jose Florez-Arango. 105

User Experience and Socioemotional Evaluation of Virtual Therapist Character Types in a Simulated Virtual Reality Exposure Therapy: Quantitative Study (e81537)
Michelle Hallmann, Nicolina Peperkorn, Francesco Vona, Youssef Shiban, Jan-Niklas Voigt-Antons. 127

Perceived Usability, User Experience, and Technology Acceptance of Role-Specific Augmented Reality Decision Support Tools for Cardiac Arrest Resuscitation: Prospective Observational Pilot Study (e72013)
Ryan Kang, Adam Cheng, Yiqun Lin, Hyeongil Nam, Jennifer Davidson, Donovan Duncan, Johan Siebert, Sergio Manzano, Alexandre De Masi, Ana Rajic, Sharleen Olanka, Frederic Ehrler, Kangsoo Kim. 148

Persuasive Gamified Virtual Reality Experience to Enhance Engagement and Focus in Young Adults With Mild Anxiety Symptoms: Randomized Pilot Experimental Study (e66713)
Mahlet Argaw, Nuru Jingili, Solomon Oyelere, Markus Nyström. 164

Improving Health-Related Symptoms and Behaviors in Children and Adolescents Diagnosed With Diabetes by Using a Virtual Reality–Gamified Self-Care Method: Randomized Controlled Trial (e81402)
Ruhollah Solat, Reza Pourhosein, Hadi Bahrami Ehsan, Ata Pourabbasi, Amin Hashemi, Javad Hatami. 178

Mixed Reality–Based Slit Lamp for Ophthalmic Examination and Telemedicine: Technological Development and Validation Study (e93513)
Rui Zhou, Wei-Chiang Lin, Byron Lam, Rong Wen, Noble Amadi, Shuliang Jiao. 198

Brief Virtual Reality and Mixed Reality Mindfulness Breathing Exercise for Emotional Well-Being and Cognitive Functions in University Students: Within-Subjects Experimental Design Study (e84239)
Zoey Eun, Charmaine Koh, Hwajin Yang, Adalia Goh, Meilan Hu, K Kasturiratna, Andree Hartanto. 212

Immersive Virtual Reality–Supported Cognitive-Behavioral Therapy for Patients With Mild to Borderline Intellectual Disabilities and Substance Use Disorders: Two Exploratory Studies (e82601)
Samantha Murray, Simon Langener, Hanneke Kip, Randy Klaassen, Saskia Kelders, Dirk Heylen, Joanne VanDerNagel. 232

Virtual Reality in Health Professions Education: Qualitative Descriptive Study of Educators' Perspectives (e52925)
Tak Yu, Michael Rowe, Jose Frantz. 247

Predictive Factors of Augmented Reality–Based Clinical Task Performance Among Novice Users: Cross-Sectional Quantitative Study (e81236)
Amogh Vellore, Shovan Bhatia, Michael Kann, Nicolás Kass, Regan Shanahan, Jacquelyn Jardini, Jayne Miner, Sohail Daulat, Griffin Hurt, Rishi Basdeo, Nicole Don, Jacob Biehl, Edward Andrews. 257

User Experience Comparison of a Traditional Mannequin-Based and an Augmented Reality Lumbar Puncture Simulator in Geriatric Medicine Training: Questionnaire Study (e98090)
Emma Combret, Alexis Bourgeais, Frédéric Noublanche, Cédric Annweiler. 271

XSL•FO
RenderX

User Acceptance of a Virtual Reality Beer Pong Intervention in Postoperative Rehabilitation: Exploratory Mixed Methods Study (e82704)
Tim Schneider, Merle Schlender, Verena Uslar, Sirko Pelzl, Daniela Salzmann, Nomi Kreß, Navid Tabriz, Dirk Weyhe. 282

Evaluating the Effects of Clinician Prescribing and Implementation Materials on Adoption of Virtual Reality Therapeutics: Randomized Feasibility Pilot Study (e90626)
Ashlyn Zebrowski, R Fernandez, Cameron Pinkelton, Rebecca Kitzmiller, Adam Kiefer, Laura Stanley, Gita Mody, Carlton Moore, Lukasz Mazur. 298

Patient Perception of the Use of the Apple Vision Pro Virtual Reality Headset in an Anesthesiology Informed Consent Process: Prospective Randomized Controlled Crossover Trial (e95806)
Sabrina Chriqui, John Ryan, Bao Luu, Minh Tran, Byron Fergerson, Alan You, Christian Dameff, Preetham Suresh, Jeffrey Tully. 325

XSL•FO
RenderX

Different Delivery Modalities of Virtual Reality Training in Emergency Medicine: Systematic Review

Marianna Costa[1], MD; Martina Codato[1], MD; Giorgia Brugnera[1], MD; Silvia Galiazzo[1], MD; Elisabetta Cantele[1], MD; Todd P Chang[2], MD, MAcM; Silvia Bressan[1], MD, PhD

[1]Department of Women's and Children's Health, University of Padua, Via Giustiniani 3, Padova, Italy

[2]Children's Hospital Los Angeles, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Silvia Bressan, MD, PhD
Department of Women's and Children's Health, University of Padua, Via Giustiniani 3, Padova, Italy

Abstract

Background: Simulation is a cornerstone of emergency medicine education, and virtual reality (VR) has emerged as a promising immersive and scalable training modality. However, it remains unclear which VR instructional design features most effectively support learning outcomes.

Objective: We aimed to provide a systematic overview of the literature on immersive VR delivery modalities in prehospital and in-hospital emergency medicine training, focusing on how different instructional design features influence educational effectiveness.

Methods: We conducted a systematic review according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines, including English-language articles examining the use of VR for the training of health care professionals, specifically randomized controlled trials (RCTs), non-RCTs, and prospective or retrospective observational comparative studies. PubMed/MEDLINE, Scopus, Embase, and CINAHL were searched up to January 28, 2026. We included studies evaluating immersive headset-based VR for emergency medicine training among qualified health care professionals. Studies involving students, laypersons, nonemergency settings, or augmented, mixed, or web-based reality were excluded. Risk of bias was assessed using the Joanna Briggs Institute critical appraisal tools. Given the wide heterogeneity in methodology across studies, a meta-analysis was not performed, and findings were synthesized using a structured narrative approach based on VR delivery modalities.

Results: Six comparative studies in emergency medicine training were included from 1770 identified records. The included studies compared different immersive VR delivery modalities, including varying exposure schedules, software with or without gamification, machine-guided versus educator-guided training, active versus observational participation with different persuasive feedback strategies, and VR with or without manikin integration. Educational outcomes were heterogeneous, including knowledge retention, procedural performance, and learner-related measures. Overall, the included studies were of moderate to high methodological quality. While observational studies were more susceptible to selection bias, RCTs demonstrated adequate randomization and statistical analysis. Across the included studies, repeated and temporally distributed VR exposure consistently improved knowledge retention and performance, whereas gamification, increased persuasive feedback, educator-guided delivery, and manikin integration showed inconsistent or no additional educational benefit.

Conclusions: Unlike previous reviews focusing on the overall educational value of VR, this review compared different immersive VR delivery strategies in emergency medicine training. The few studies identified suggest that structured and repeated engagement with VR simulations may enhance learning effectiveness, whereas additional technological features do not consistently improve outcomes. The available evidence is limited by the small number of studies, heterogeneous interventions and outcome measures, and moderate risk of bias, restricting the strength of the conclusions. These findings suggest that optimizing instructional design may be more important than increasing technological complexity when developing immersive VR educational programs. Further high-quality studies with standardized outcome measures are needed to identify the most effective VR training strategies for emergency medicine education.

Trial Registration: PROSPERO CRD42024507591; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024507591

(JMIR XR Spatial Comput 2026;3:e84310) doi:10.2196/84310

KEYWORDS

simulation; virtual reality; VR delivery modalities; emergency training; educational strategies

XSL•FO
RenderX

Introduction

Emergency care represents a particularly high-risk setting, characterized by operational complexity, intense time pressure, rapidly evolving situations, and dynamic team structures [1]. For this reason, it is important that both health care professionals working in this field and personnel in training receive proper and repeated emergency training. In some institutions, the number of residents has increased over the years, leading to varying levels of preparedness for critical events and reduced opportunities for direct emergency clinical experience [2-4]. As a result, many residents receive less exposure to emergency scenarios than required, raising concerns that they may not acquire the necessary competencies during their training.

Limited exposure to time-critical emergencies has highlighted the need to expand simulation-based medical education. Simulation has become an essential component of emergency medicine education, providing a safe, controlled environment that ensures psychological safety and enhances skill retention. This educational approach enables health care professionals to practice managing high-stakes scenarios at various levels of complexity, improving their readiness to respond effectively in real-life emergencies [5].

Different simulation techniques and delivery modalities have been employed for various training purposes; as a matter of fact, one educational approach cannot adequately cover trainees' needs at each learning level. Substantial empirical research concerning conventional manikin-based simulation underscores the effectiveness of simulation training compared to traditional educational methods [5]. Manikin-based simulation is recognized as a useful and safe technique; the modalities to conduct health care simulation scenarios (location, equipment, and human resources needed) according to the planned learning objectives are well described [6]. Although manikin simulation has advantages (skill practice, multidisciplinary team training), it requires significant resources (instructors, medical equipment, and a dedicated setting) to be carried out [6].

Virtual reality (VR) represents a more recent technology for simulation training in medical education [7-9]. It is defined as an immersive technology that reproduces an environment that completely replaces the physical context with a 3-dimensional digital space, allowing the user to engage in a simulated setting for a temporary experience, often within a safe or controlled area. Through its hardware components, such as head-mounted displays and joysticks, and software, VR enables trainees to immerse themselves in scenarios that closely match real-life settings and to tackle progressively challenging tasks through tailored learning paths [10].

Since virtual simulation can be self-administered, trainees can access virtual clinical or surgical settings anywhere at any time and develop self-reliance in specialized areas, such as pediatric emergency care. There has been a significant expansion of interest in research on VR and its many applications in medical education. The effects of VR-based training on health care professionals have mostly been studied in relation to the improvement of technical skills, especially in the surgical field (laparoscopic surgery or robot-assisted surgery, and procedural practice) [11,12].

In recent years, simulation through VR in health care has expanded as an educational tool also for nontechnical skills training (eg, decision making, leadership, and situation awareness) [12]. A recent review comparing VR training with traditional manikin-based simulation for health care professionals found that VR experiences are most effective when proctored [13]. Specifically, Foronda et al [13], drawing on both the literature and experience, recommend providing learners with guidance prior to VR simulations—such as instruction on equipment use, preparatory briefings, opportunities for questions, and facilitated debriefing—to enhance the effectiveness of new technology–based training.

Although the use of VR in medical emergency simulation is expanding, little is known about the effectiveness of different delivery modalities of VR simulation. The question is not *if* VR is effective, but *how* VR can be most effective. This is an important area of investigation to identify the optimal balance between educational benefit and resource investment. Our systematic review aims to systematically identify and summarize available scientific evidence describing immersive VR delivery modalities used in prehospital and in-hospital emergency medicine training, with the objective of comparing the different approaches currently adopted in VR-based educational interventions in the field of emergency medicine training.

Methods

Study Design

This systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines (Checklist 1). The study protocol was registered in the PROSPERO (International Prospective Register of Systematic Reviews; CRD42024507591).

Search Strategy

Electronic searches were performed in the following databases: PubMed/MEDLINE, Scopus, Embase, and CINAHL (with MEDLINE records excluded within the CINAHL search to minimize duplication across databases) from January 2000 to January 28, 2026. The initial search was conducted on January 4, 2025, and rerun on January 28, 2026, following peer-review feedback, incorporating the suggested modifications to the search strategy.

The search strategy was developed according to the PICO (Population, Intervention, Comparison, and Outcome) framework. The following concepts were combined: population (P)=health care professionals involved in emergency medicine training; intervention (I)=immersive VR-based training; comparison (C)=alternative VR delivery modalities for the same educational purpose; and outcome (O)=effectiveness of VR delivery strategies.

The search strategy was constructed using MeSH, CINAHL Headings, and free-text keywords according to the key research question. No previously published search filters were used, and

XSL•FO
RenderX

no search strategies from prior reviews were adapted or reused. It was optimized by an experienced librarian.

The search was limited to English-language publications and to studies published between 2000 and 2025, as contemporary literature was considered most representative of current applications of VR in health care education. Due to indexing lag across databases, records published in 2026 may not yet have been fully indexed at the time of the search update. No study registries were searched. Gray literature sources were not searched, and no additional online or print sources were purposefully browsed. We did not contact corresponding authors, first authors of included studies, or subject-area experts to identify ongoing or unpublished studies. The reference lists of eligible articles were screened to ensure inclusion of all relevant articles. The full search strategies for all databases are reported in Checklist 2 according to the PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension) extension for reporting literature searches in systematic reviews. All records identified through database searching were uploaded into the Covidence software for duplicate removal and systematic screening of eligible studies.

Inclusion Criteria

We included English-language articles focusing on VR used for the training of health care professionals (ie, medical students, residents, clinicians, trainees, assistants, interns, nurses, emergency technicians, paramedics, and prehospital physicians) in prehospital and in-hospital medical emergencies. We examined its application both in the context of formal emergency medicine training programs and, more broadly, in educational interventions designed to enhance the competencies of health care professionals in the management of acute and urgent clinical scenarios. Our review included randomized controlled trials (RCTs), non-RCTs, and observational prospective and retrospective comparative studies.

Exclusion Criteria

We excluded studies exclusively enrolling laypeople, nonqualified health care workers, or students. Studies focusing on VR training for scenarios different from medical emergencies and studies assessing augmented reality, mixed reality, or web-based platforms instead of head-mounted display VR technology were excluded.

Augmented reality is defined as a technology that extends or enhances the real physical environment with superimposed virtual content, enabling hybrid experiences that combine real and digital elements, accessible in everyday contexts through mobile devices, wearables, or projectors, whereas mixed reality has been defined as any user experience combining real and virtual objects, with mixed reality's 2 subforms being augmented reality and augmented virtuality [10].

Study Selection

Identified records were screened based on the inclusion and exclusion criteria. Two independent reviewers (GB and SG) initially screened the articles based on their titles and abstracts. Eligible papers underwent full-text assessment to identify papers that eventually met the inclusion criteria. Any disagreement during the study selection process was discussed with a third independent reviewer (SB) to resolve discrepancies.

Data Extraction

Relevant information from the included studies was extracted by 2 independent reviewers (M Codato and EC) using an ad hoc data extraction tool. At this stage, any disagreement on the data to be extracted was resolved through discussion with a third reviewer (SB). We extracted information on the authors, journal, year of publication, country of study conduct, study design, characteristics of the study population, study objectives, VR training characteristics and type of scenarios, VR delivery modalities (intervention/control), and results on the effectiveness of the VR delivery modalities assessed. The outcomes assessed by each included study reflected improvements in both technical skills and nontechnical competencies, such as clinical reasoning, prioritization of interventions, and situational awareness. Effect sizes were interpreted according to Cohen *d* conventional thresholds (small=0.2, moderate=0.5, and large=0.8). Cohen *d* was derived from the reported means and SDs using the standard formula [14].

Data Analysis

Given the substantial heterogeneity in study designs, interventions, comparators, and outcome measures, a quantitative meta-analysis was not performed. Instead, the findings were synthesized using a structured narrative synthesis informed by the Guidance on the Conduct of Narrative Synthesis in Systematic Reviews [15].

The synthesis proceeded in 3 stages. First, studies were grouped according to the educational domain targeted by the intervention, including technical skills acquisition and recognition of emergency clinical scenarios, cardiopulmonary resuscitation (CPR) training and the approach to clinical emergencies, and collaborative training. Within these domains, studies were further organized according to key VR delivery characteristics, including gamification, frequency and spacing of exposure (distributed vs single-session training), level of immersion and participation (active vs observer-based learning), feedback modality (persuasive vs minimal feedback), and integration with manikin-based simulation or traditional training modalities.

Interpretation of the findings incorporated consideration of methodological quality, risk of bias, and contextual factors, such as participant characteristics, training setting, and variability in outcome measurement.

Risk of Bias

Risk of bias assessment was performed by 2 independent reviewers. Studies were appraised according to the individual domains of the Joanna Briggs Institute Critical Appraisal Checklists (see Checklist 3). Any disagreements were resolved through the assessment of a third independent reviewer.

The methodological quality of the included studies informed the interpretation of the findings. In the absence of a quantitative synthesis, no formal weighting was applied; however, greater emphasis was placed on findings derived from RCTs with a lower risk of bias. Findings from studies with an unclear or

XSL•FO
RenderX

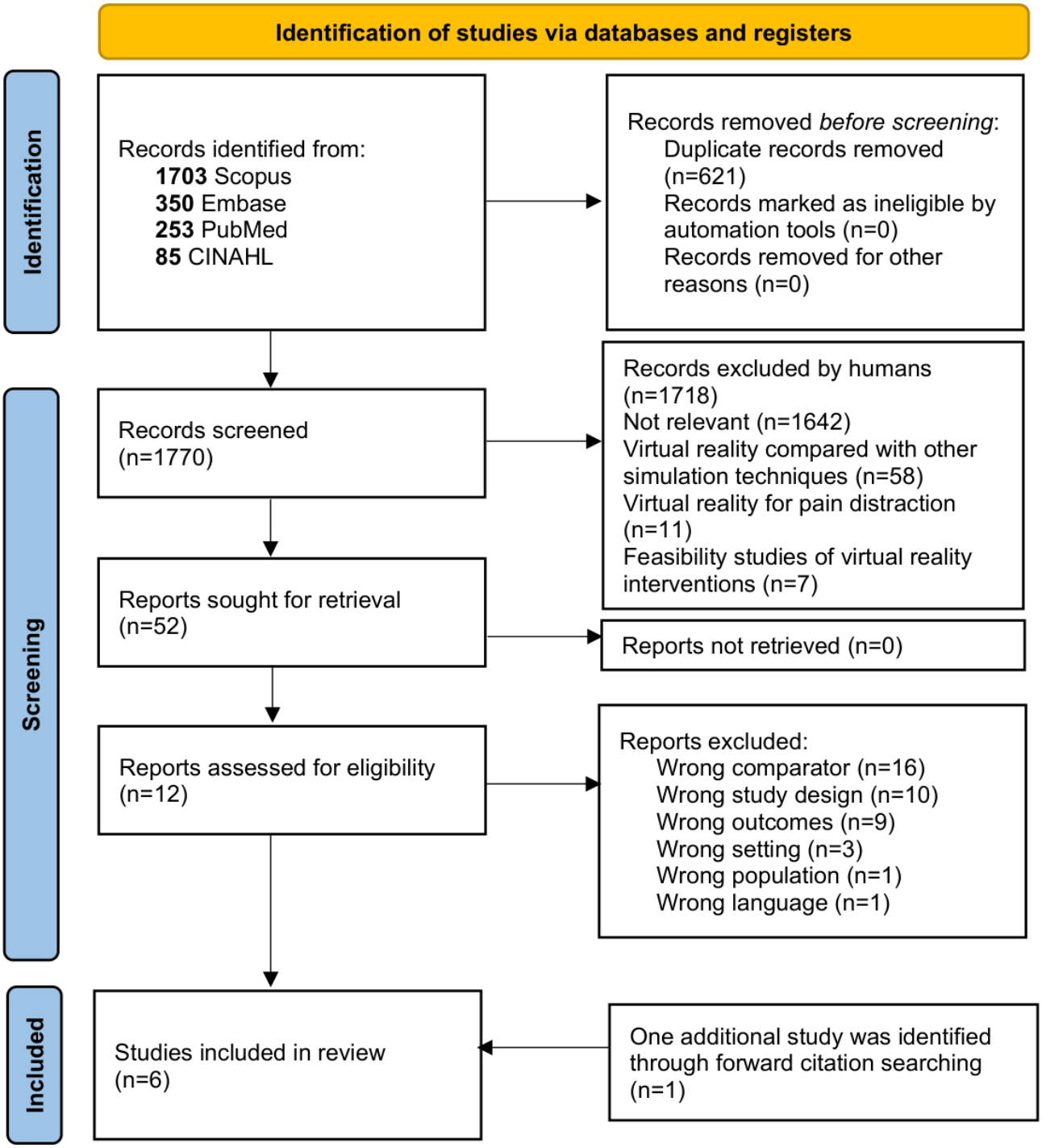
higher risk of bias were interpreted cautiously and considered hypothesis-generating rather than confirmatory.

Results

Study Characteristics

Following the initial search, a total of 1770 articles were retrieved, and 52 unique studies were left after duplicate removal and were screened using the Covidence online software. The PRISMA 2020 flowchart for the study selection process is shown in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 study flow diagram of the study selection process.



After an independent full-text review based on the inclusion and exclusion criteria, only 5 studies met the inclusion criteria for our systematic review. One additional article was identified through hand-searching of the reference lists of the included articles, increasing the total number of included studies to 6.

XSL•FO
RenderX

Their characteristics are described in Table 1. None of the included studies pertained to the prehospital setting.

XSL•FO
RenderX

Table . Characteristics of included studies investigating virtual reality interventions for technical skills acquisition, cardiopulmonary resuscitation, and the recognition and management of clinical scenarios in emergency medicine.

VR[a] for technical skills acquisition and systematic recognition/management of clinical emergencies							
Study	Country/study period	Design/setting	Participants	Learning objectives	VR scenario/VR module features	VR delivery modality	
						Intervention	Control
Larsen et al [16], 2023, *Ultrasound in Med & Biol*	Denmark/Oct 2021-Dec 2021	• RCT[b] • 2 study arms • Double-blinded • Allocation ratio 1:1 • Single-center	• 48 students/residents (24 per study arm) • No previous experience with FLUS[c]	• Train in the use of lung ultrasound (pathology findings detection)	• FLUS module (pleural effusion, pneumonia, etc)	• Gamified VR module	• Nongamified VR module
Lietz et al [17], 2023, *Acta Paed*	Austria/Aug 2021-Mar 2022	• Pilot RCT • 3 study arms • Allocation ratio 1:1:1 • Single-center	• 36 students (12 per study arm)	• Improve proficiency in emergency algorithms and tasks; crisis management in high-intensity emergency; CPR[d] skills	• Infant with shock rapidly moving to nonshockable cardiac arrest • Single-player	• VR training interval after baseline training • Group A: monthly for 5 times • Group B: once after 3 months • Group C: baseline training	• No further VR sessions after baseline VR training
Kitapcioglu et al [18], 2024, *JMIR Serious Games*	Turkey/2023 - 2024	• RCT • 2 study arms • Single-center	• 62 students (31 per study arm)	• Provide ACLS[e] training	Both groups undertook VR-based basic training for ACLS. Afterward: • MG[f] group were trained with a VR-based advanced training module • EG[g] group attended VR-based, educator-guided training in the metaverse	• VR-EG training after VR basic ACLS training	• VR-MG full • ACLS training (based + advanced)

XSL•FO
RenderX

VR[a] for technical skills acquisition and systematic recognition/management of clinical emergencies

Study	Country/study period	Design/setting	Participants	Learning objectives	VR scenario/VR module features	VR delivery modality	
						Intervention	Control
Jaskiewicz et al [19], 2020, *Medicine*	Poland/April 2019-May 2019	• Cross-over study • Single-center	• 91 medical students completed BLS[h] course	• Assess CPR quality during sudden cardiac arrest scenario (SCA)	• 3-h BLS course (45 min theoretical background + 90 min of CPR on the manikin) • Single-player	• 2 min hands-on-only CPR • VR SCA scenario versus traditional SCA scenario	• 2 min hands-on-only CPR • Traditional SCA scenario versus VR SCA scenario
Mühling et al [20], 2023, *Multimedia System*	Germany/Oct 2020-Jul 2021	• Observational study • 2 groups (observers and active participants) • Single-center	• 227 medical students • 97 active participants, 130 observers (8 - 10 students per VR session)	• Train cognitive abilities (ie, executive functions, situational awareness, time management, and prioritize tasks)	• 120-min VR sessions including 3 - 4 scenarios and guided by an assistant + debriefing (60 min) of the cases with a tutor (ie, sepsis, acute myocardial infarction, etc) • Single-player	• Active participants • Only 1 student per scenario (3 - 4 students per VR session) was exposed to the VR environment	• Observers watched from the first-person perspective on a monitor and were able to assist verbally the colleague student undergoing the VR scenario
Khanal et al [21], 2014, *J Biomed Informatics*	United States/—	• RCT • 3 study arms • Single-center	• 156 ACLS-certified clinicians (26 teams of 6 participants)	• Provide team ACLS training	• VF[i] or PEA[j] case scenarios (5 min ACLS team training + 30 min) • VR groups received training in a VR environment (20 min guided single-user tutorial to familiarize with the user interface) • Multiplayer	• Persuasive VR ACLS training (feedback components available for all team members, besides alerts specific to individual member roles) • Minimally persuasive VR ACLS training (limited feedback components)	• Traditional manikin-based in-person training

[a]VR: virtual reality.

[b]RCT: randomized controlled trial.

[c]FLUS: focus lung ultrasound.

[d]CPR: cardiopulmonary resuscitation.

[e]ACLS: advanced cardiac life support.

[f]MG: machine guidance.

[g]EG: educator guidance.

XSL•FO
RenderX

[h]BLS: basic life support.

[i]VF: ventricular fibrillation.

[j]PEA: pulseless electrical activity.

As indicated in Table 1, medical students were the primary study population in 4 of the 6 included studies. The VR scenarios, the outcomes, and the delivery modalities assessed were different.

Three of the included RCTs investigated the use of VR for technical skills acquisition, disease recognition, and management of clinical emergencies, operationalizing the intervention through distinct modalities. These included varying time intervals between VR training sessions [15], the use of VR software with or without gamification elements [17], and educator-guided versus machine-guided VR sessions [16].

Two studies implemented the use of VR for CPR training [18] and for the approach to clinical emergencies, comparing the use of VR with or without a manikin during CPR and passive versus active exposure to VR-based training, respectively [19]. Only 1 study described collaborative VR sessions delivered through VR software incorporating varying levels of persuasive feedback intensity.

As shown in Table 2, the learning outcomes and the measures of effectiveness assessed by the included studies were also diverse.

XSL•FO
RenderX

Table . Outcomes, measurement instruments, and effect sizes of included studies evaluating virtual reality interventions in technical skills training, cardiopulmonary resuscitation performance, and clinical scenario management in emergency medicine.

Study	Country	Outcomes	Measures of outcome	Key findings	Cohen *d* for learning outcomes
Virtual reality for technical skills acquisition and systematic recognition and management of clinical emergencies					
Larsen et al [16], 2022, *Ultrasound in Med & Biol*	Denmark	• Primary: Educational impact of novice FLUS[a] users participating in a gamified/nongamified VR[b] training module in FLUS • Secondary: Assessment of learning effect by comparing the scores of each group with known test scores of novices, intermediates, and experienced users in FLUS	• Assessment by simulation test on US Simulator • Test scores were calculated as the sum of points, given for correct assessment of final diagnosis in each case. A correct final diagnosis=2 points, partially correct=1 point, and incorrect=0 points	• No difference was found between gamified (mean=15.5 points, 95% CI 14.5 - 16.5) and nongamified (mean=15.2 points, 95% CI 14.3 - 16.1) • Gamified elements did not affect learning outcome (mean, SD): gamified: 15.5 (SD 2.4) vs nongamified: 15.2 (SD 2.1; *P*=.66) • Known test score on US mentor simulator lung ultrasound (mean, SD) • Novices: 12 (SD 3.46); intermediate: 15.7 (SD 1.74); experienced: 17 (SD 0.85) • Gaming/nongaming learning effect: most comparable test score for known competence groups in lung ultrasound • Novice: Both VR groups had significantly higher test scores than known novice (gamified: *P*<.001 vs nongamified: *P*<.001) • Intermediate: Both VR groups had test scores that were not significantly different from those of known intermediates (gamified: *P*=.63 vs nongamified: *P*=.24) • Expert: Both VR groups had significantly lower test scores than known experts (gamified: *P*<.001 vs nongamified: *P*<.001)	• <0.2 (small)

XSL•FO
RenderX

Study	Country	Outcomes	Measures of outcome	Key findings	Cohen *d* for learning outcomes
Lietz et al [17], 2023, *Acta Paed*	Austria	• Primary: Individual performances after each VR session • Secondary: Evaluation of subjective stress and workload after each training session • User experience	• Individual performance score (15 variables coded into the program and including all measures sets and the duration and initial timing of each intervention) • NASA Task Load score • Questionnaire regarding various aspects of participants' individual experiences with the VR headsets after each training session	• Group A had the highest score improvement compared with the other 2 study groups • Assessment training (mean, SD): • Group A[c]: 11.08 (SD 1.84) vs Group B[d]: 8.54 (SD 1.85; *P*=.004, Cohen *d*=1.38) • Group A: 11.08 (SD 1.84) vs Group C[e]: 9.25 (SD 1.80; *P*=.03) • The average perceived mental demand of participants in Group A after the baseline training was 11.92 (SD 4.92) out of 20 points (*P*=.42). It reduced significantly in the final assessment training to 8.17 (SD 4.59; *P*=.06) • Compared satisfaction with individual performance in Group A: improvement of 0.67 from a mean score of 2.17 (SD 0.83) to 1.50 (SD 0.90); Group A felt the most comfortable during the VR simulation, with an improvement of 0.92 from a mean score of 2.75 (SD 1.48) to 1.83 (SD 0.83) • Participants stated that VR interface was easy to use (mean, SD): Group A (1.25, SD 0.83), Group B (1.50, SD 0.67), Group C (1.75, SD 0.62)	• 1.38 (vs one-time VR) • 1.01 (vs control group) (large)
	Turkey				• 0.44 (moderate)

XSL•FO
RenderX

Study	Country	Outcomes	Measures of outcome	Key findings	Cohen *d* for learning outcomes
Kitapcioglu et al [18], 2024, *JMIR Serious Games*		• Primary: Comparison of VR ACLS[f] test scores between the MG[g] and EG[h] groups	• Virtual reality examination score • Duration of MG vs EG training for each participant • Correlation between examination scores and time spent on training	• Virtual reality exam score (total 100 points): • MG group (mean, SD) (53, SD 33) vs EG group (mean, SD) (68, SD 35; *P*=.08) • Time spent for MG vs EG training (in min): • MG group (median, range) (66, range 13-100) vs EG group (median, range) (15, range 5-30; *P*=.002) • MG group rho=0.569, *P*=.005 vs EG group rho=0.298 and *P*=.10. The post hoc power analysis (80%) supported these findings	
Virtual reality for CPR training and approach to clinical emergencies					

XSL•FO
RenderX

Study	Country	Outcomes	Measures of outcome	Key findings	Cohen *d* for learning outcomes
Jaskiewicz et al [19], 2020, *Medicine*	Poland	• Primary: Assess CCs[i] quality during a SCA[j] scenario performed with the use of VR prototype • Secondary: Assess if additional equipment may cause CCs subjectively harder to provide and analyze impressions of students	• CCs were conducted on CPR mannequin; quality was measured with SimPad equipped with Skillreporter software (Laerdal, Stavanger, Norway) • Five questions regarding previous VR experience • Questionnaire about the applicability of VR in the resuscitation training	• No statistically significant differences between the VRS[k] and TS[l] groups neither in CCs rate (mean, SD): TS 114.9 (SD 11.8) vs VRS 114.2 (SD 11.9) nor in CCs depth (mean, SD): TS 49.3 (SD 7.4) vs VRS 47.8 (SD 8.8). • A statistically significant difference was found in the percentage of FCR[m] component (for VRS group mean 60.2, SD 37.5 vs TS group mean 83.7, SD 27.9; *P*<.001). • Majority of participants (97/99, 97.8%, 95% CI 0.92 - 0.99; *P*<.01) rated the educational value of VR in CPR training at 3 or 4 points in 1- to 4-point scale (1=very low, 4=very high)	• 0.71 (moderate)
Mühling et al [20], 2023, *Multimedia System*	Germany	• Primary: Evaluation of STEP-VR[n] into the medical curriculum • Secondary: Perception of VR-based training, psychological distress users experience and perceived learning benefits	• Questionnaires for evaluation in both groups of: • Level of presence • Estimated learning success • Motivation pre-training and post-training • Questionnaires for evaluation of perceived stress on estimated learning success in AP[o] group		• 0.20 (small)

XSL•FO
RenderX

Study	Country	Outcomes	Measures of outcome	Key findings	Cohen *d* for learning outcomes
				• No differences in demographic characteristics between observers (n=130) and AP (n=97) • AP rated the degree of presence significantly higher than observers (mean, SD) (3.81, SD 0.98, n=97 and 3.26, SD 0.74, n=130, respectively, *P*<.001) • Estimated learning success was rated higher in AP than observers (mean, SD) (3.87, SD 0.75, n=97 and 3.72, SD 0.77, n=130, respectively, *P*=.05) • Motivation pretraining was not significant, posttraining (mean, SD) 3.61, SD 0.84 and 3.07, SD 0.88, *P*>.01 • No statistically significant correlations were identified	
Collaborative VR training					

XSL•FO
RenderX

Study	Country	Outcomes	Measures of outcome	Key findings	Cohen *d* for learning outcomes
Khanal et al [21], 2014, *J Biomed Informatics*	United States	• Primary: Team performance tested during high-fidelity simulation performed in front of ACLS human experts • Pretest phase: assessment of baseline performance (directly observing teams) • Posttest phase: assessment of performance after the intervention (directly observing teams and reviewing video-recorded scenario)	• Pre/postintervention tests • Pretest phase: time to perform critical actions • Posttest phase: electronic checklist assessment tool developed and validated by a team of expert ACLS trainers. It included items that correspond to AHA[p] guidelines for ACLS	• Pretest team performance: no statistically significant difference between control and persuasive groups. AHA guidelines adherence (141/360, 39.4%); of total 360 tasks: • control: 39.1% (47/120), persuasive: 35.8% (43/120), minimally persuasive: 43.3% (52/120) • Posttest team performance: increased AHA guidelines adherence (58.3%); of total 360 tasks: • control: 68.3% (82/120), persuasive 57.5% (69/120), minimally persuasive: 49.1% (59/120) • PEA[q]/Vfib[r]/VTach[s] (control vs minimally persuasive: *P*=.05 and *P*=.02); • Pre/post comparison of performances of groups showed significantly improved performances for control (*P*=.02 for PEA and *P*=.01 for VFib/VTach) and persuasive groups (*P*=.02 for PEA, *P*=.048 for VFib/VTach)	• N/A[t]

[a]FLUS: focus lung ultrasound.

[b]VR: virtual reality.

[c]Group A: monthly training for 5 months.

[d]Group B: once after 3 months.

[e]Group C: baseline training.

[f]ACLS: advanced cardiac life support.

[g]MG: machine-guided.

XSL•FO
RenderX

[h]EG: educator guided.
[i]CCs: chest compressions.
[j]SCA: sudden cardiac arrest.
[k]VRS: virtual reality scenario.
[l]TS: traditional scenario.
[m]FCR: full chest relaxation.
[n]STEP-VR: Simulation-Based Training of Medical Emergencies for Physicians.
[o]AP: active participants.
[p]AHA: American Heart Association.
[q]PEA: pulseless electrical activity.
[r]VFib: ventricular fibrillation.
[s]VTach: ventricular tachycardia.
[t]N/A: not applicable.

Regarding the use of VR for technical skills acquisition, disease recognition, and the management of clinical emergencies, no significant differences were identified in the RCTs comparing gamification strategies [17] or automated machine-guided versus educator-guided VR training modalities [16].

In contrast, a study examining different temporal distributions of VR training exposure demonstrated significant differences in outcomes [15]. Specifically, monthly VR training sessions conducted over 5 months resulted in superior individual performance scores compared with 2 sessions delivered within a 3-month period [15].

With respect to CPR training, the study findings indicated no statistically significant differences between groups in terms of chest compression rate or depth. However, a statistically significant difference was observed in full chest recoil, favoring the control group that initially underwent traditional CPR simulation training [18].

Similarly, considering the approach to emergencies as a learning objective, comparisons between active and passive VR exposure revealed that passive participation—defined as observing the VR scenario from a first-person perspective via a monitor while peers engaged in the VR training—was associated with lower self-reported learning success, as measured by an ad hoc questionnaire [19].

Finally, the study investigating varying intensities of persuasive feedback embedded within VR software during collaborative VR sessions found no significant differences between the study groups [20]. To further synthesize findings across studies, results were organized according to VR delivery modalities, focusing on shared instructional design characteristics rather than individual study outcomes (Table 3).

Table . Synthesis of virtual reality modalities.

Virtual reality delivery modality	Study	Effect on learning	Interpretation
Gamification	Larsen et al [16]	No consistent effect	Not determinant
Distributed exposure	Lietz et al [17]	Positive	Likely beneficial
Active participation	Mühling et al [20]	Positive	Enhances engagement
Feedback intensity	Khanal et al [21]	Mixed	Context-dependent
Educator involvement	Kitapcioglu et al [18]	No clear benefit	Not decisive
Manikin integration	Jaskiewicz et al [19]	Neutral	No added value

Quality Assessment

The summary risk-of-bias assessment of the included studies is displayed in Table 4.

Four of the 6 included studies had an unclear risk of bias in one or more domains. The evaluation of bias related to participant retention (question 10) was not applicable to most included studies. Only one RCT [17] and one observational study [19] had a low risk of bias. Findings from studies with a higher risk of bias must be interpreted with caution.

XSL•FO
RenderX

Table . Risk of bias assessment of included studies evaluating virtual reality–based educational interventions.

Study	1	2	3	4	5	6	7	8	9	10	11	12	13[a]
RCT[b]													
Khanal et al [21], 2014	?[c]	?	+[d]	+	–[e]	+	–	+	+	N/A[f]	+	+	+
Larsen et al [16], 2023	+	+	+	+	+	+	+	+	+	N/A	+	+	+
Lietz et al [17], 2023	+	+	+	?	?	+	?	+	+	+	+	+	+
Kitapcioglu et al [18], 2024	+	–	?	–	+	?	+	+	+	N/A	+	+	+
Observational study													
Jaskiewicz et al [19], 2020	+	+	+	–	–	?	+	N/A	N/A	N/A	+	N/A	N/A
Mühling et al [20], 2023	+	+	+	+	+	+	+	N/A	N/A	N/A	+	N/A	N/A

[a]Numbered column headings (1-13) represent the corresponding questions (Items 1-13) of the Joanna Briggs Institute Critical Appraisal Checklist used to assess the methodological quality of the included studies.

[b]RCT: randomized controlled trial.

[c]?: Unclear.

[d]+: Yes (criterion met).

[e]–: No (criterion not met).

[f]N/A: not applicable.

Discussion

This systematic review aimed to compare different VR delivery modalities for emergency medicine training, focusing on how specific instructional design features influence educational effectiveness. Despite the growing adoption of VR in medical education, only 6 comparative studies met the inclusion criteria, highlighting the limited evidence specifically addressing *how* immersive VR should be delivered rather than whether VR is effective per se. Overall, the available evidence suggests that active learner engagement and repeated, temporally distributed VR exposure are the instructional strategies most consistently associated with improved learning outcomes, whereas additional technological features, including gamification, persuasive feedback, and educator-guided versus machine-guided delivery, did not demonstrate consistent educational benefits. However, the small number of studies, together with their methodological and conceptual heterogeneity, limits the strength of these conclusions and warrants cautious interpretation.

Notably, several excluded studies compared VR with alternative educational approaches (eg, conventional lectures [21], manikin-based simulations [22], video-based training [23], traditional technique guides [24], etc), highlighting how research has focused more on whether VR is effective compared to other teaching techniques than on how VR can best be effective in enhancing learning.

Timely and active VR exposure appeared to be most effective for training. Repetitive VR training exposures over monthly time windows showed a higher dose-response effectiveness in terms of performance compared with longer time windows of 3 months [15]. This observation [15] aligns with broader educational theory and simulation literature. Beyond VR-specific interventions, the importance of training frequency reflects broader educational principles related to spaced learning, which have consistently been associated with improved knowledge consolidation and long-term skill retention in medical education [25,26].

Prior research has consistently demonstrated that repeated exposure to high-fidelity simulation and CPR retraining

XSL•FO
RenderX

improves both knowledge and skill retention [27]. While a previous study reported a significant decline in CPR skills within 6 to 12 months after training [28]—suggesting the need for refresher sessions within that timeframe—a more recent scoping review on pediatric CPR retraining strategies indicates that shorter, more frequent refreshers may be more effective for sustaining performance [29].

Although definitive evidence regarding the optimal retraining interval remains limited, even brief monthly refresher sessions, lasting as little as 2 minutes, appear to enhance skill retention and support higher-quality CPR performance over time [29]. While optimal retraining intervals remain unclear, VR appears to offer an effective platform for facilitating distributed learning, enabling repeated, spaced practice to enhance skill acquisition and retention.

The addition of gamification elements did not result in measurable improvements in technical performance, suggesting that engagement-enhancing features alone may not translate into superior learning outcomes [17]. Although gamification did not improve objective technical performance in the included study, previous literature suggests that game elements may still enhance learner engagement, motivation, and training adherence. Their educational effectiveness may therefore depend less on the presence of gamification itself and more on how these elements are pedagogically integrated within the learning experience [30].

Likewise, machine-guided VR training achieved outcomes comparable to educator-guided training [16], raising important considerations regarding scalability and resource allocation.

In settings with limited faculty availability, automated VR modules may represent a feasible alternative without substantial loss of effectiveness. This finding may have important implications for emergency medicine education, where increasing learner volumes, limited faculty time, and restricted access to simulation centers frequently constrain the delivery of simulation-based training. Recent evidence suggests that AI- or machine-guided VR systems may offer scalable and accessible alternatives to traditional facilitator-dependent simulation by supporting autonomous, repeatable, and remotely accessible training experiences without substantial reductions in educational effectiveness [31].

With respect to CPR training, VR alone did not consistently improve objective performance measures, such as compression depth or rate when compared with traditional approaches [18]. These findings suggest that, for psychomotor skills requiring tactile feedback, VR may need to be integrated with physical manikins or complementary training modalities. These findings differ from those of a study that evaluated manual skill acquisition during CPR training using both guided and semiguided VR modes [32]. This paper demonstrated that repeated practice of chest compressions in VR significantly improved performance—particularly compression rate—regardless of whether a manikin was used. However, this study [32] was excluded from the present analysis as it involved undergraduate computer science students, which does not align with the target group of this review.

Conversely, in cognitive domains such as prioritization, situational awareness, and emergency algorithms, VR appears to offer potential benefits, particularly when learners are actively immersed rather than passively observing scenarios [19].

Recent studies have reported VR sessions involving both learners immersed in the virtual environment and others observing the scenario on external displays while contributing in an observational role [33,34].

These approaches may support collaborative learning, reflective observation, and peer-assisted learning, which are recognized educational components of simulation-based training. However, current evidence remains insufficient to determine whether multi-learner immersive formats provide meaningful advantages over simpler single-user models.

Collaborative VR environments and persuasive feedback systems have not consistently demonstrated superior outcomes compared with less complex approaches [20]. This may reflect the still-evolving educational integration of collaborative VR modalities rather than an actual lack of educational benefit.

Notably, few studies [15,19] incorporated structured debriefing into VR-based training, despite its well-established role in simulation-based education. An observational study [19] cites a 60-minute debriefing phase, which occurred after VR-based training; it is described as a tutor-guided review to summarize and discuss problem-solving issues. The articles mentioned above described the debriefing phase as a useful moment for feedback; however, they did not include a comparison between 2 different debriefing modalities.

Debriefing phase in manikin-based simulation has a well-defined role and structure, focusing mainly on nontechnical aspects of clinical case management; in VR simulation, it still seems to have the main aim of providing technical corrections [35,36]. The educational role of debriefing within immersive VR therefore remains poorly defined and represents an important area for future research.

Overall, VR offers important educational advantages, including standardized scenarios, scalability, and flexible access independent of physical simulation centers [37,38]. These characteristics may be particularly relevant in contexts with limited faculty resources or high learner volumes. However, its educational value depends less on technological complexity than on how it is integrated into simulation curricula. Future educational programs should therefore incorporate VR within evidence-based instructional designs that promote active learner engagement, distributed practice, meaningful feedback, and the acquisition of both technical and nontechnical competencies, particularly decision-making in complex emergency scenarios.

This review has several limitations. First, only a small number of studies met the inclusion criteria, limiting the robustness of comparisons across modalities. Second, substantial conceptual and methodological heterogeneity was present in interventions, comparators, and outcome measures, preventing quantitative synthesis. Additionally, a risk of bias was present in several studies, particularly in blinding and allocation procedures. Third, most studies were conducted among undergraduate medical students, limiting generalizability to practicing clinicians and

XSL•FO
RenderX

multidisciplinary emergency teams. Fourth, many outcomes were subjective or based on short-term assessments, and none evaluated long-term retention or patient-level clinical outcomes. Lastly, only a few studies explored the cost-effectiveness of various VR modalities. These limitations should be considered when interpreting the present findings and highlight the need for more standardized comparative research.

Unlike previous reviews focusing on the overall educational value of VR, this review compared different immersive VR delivery strategies in emergency medicine training.

By systematically differentiating and comparing distinct VR delivery modalities, this review addresses a critical gap in the literature. Although interest in VR in this field has grown substantially, robust comparative evidence across different delivery modalities remains scarce and fragmented. The few studies identified suggest that timely and structured engagement with VR-based training may enhance educational outcomes. Collectively, this work contributes to advancing simulation-based education and supports the development of more effective, scalable, and strategically implemented emergency training programs. Further high-quality research is needed to determine the most effective VR strategies for emergency medicine education. Rather than investing primarily in increasingly sophisticated VR technologies, future educational programs should prioritize evidence-based instructional design principles that promote active participation, distributed practice, and meaningful feedback. Standardized outcome measures and multicenter comparative trials will be essential to establish best practices for immersive VR implementation in emergency medicine education.

Acknowledgments

This work was supported by the Research4Residents project within the Residency Program in Pediatrics at the University of Padua, Padua, Italy. The authors declare the use of generative AI in the research and writing process. During the preparation of this systematic review, AI-assisted tools were used to support language editing, text refinement, and formatting. All study selection, data extraction, analysis, and interpretation were conducted solely by the authors, and the AI tools did not influence the scientific content or conclusions of the review.

According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing, translation, reformatting. The generative AI tool used was ChatGPT (OpenAI; GPT-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 manuscript. Declaration submitted by: collective responsibility.

Funding

The Research4Residents project within the Residency Program in Pediatrics funded the research training and open access publication fees.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary materials.

Authors' Contributions

SB, TPC, and M Costa contributed to the study design. The literature search, study selection, data extraction, and quality assessment were performed by GB, SG and M Codato. Data synthesis and analysis were conducted by M Costa, M Codato and EC. M Costa and M Codato prepared the initial draft of the manuscript. All authors critically revised the manuscript, contributed to its intellectual content, and approved the final version. M Codato and M Costa are co-first authors.

Conflicts of Interest

TPC serves as an unpaid medical advisor to i3Simulations and Health-on-Cloud. The other authors have no conflicts of interest.

Checklist 1
PRISMA checklist.
[PDF File, 109 KB - xr_v3i1e84310_app1.pdf]

Checklist 2
PRISMA-S checklist.
[PDF File, 362 KB - xr_v3i1e84310_app2.pdf]

Checklist 3
Joanna Briggs Institute checklist.
[DOCX File, 702 KB - xr_v3i1e84310_app3.docx]

XSL•FO
RenderX

References

1. Carlin E, Dubash R, Ho J, Bertenshaw C. Simulation in emergency medicine. Emerg Med Australas 2021 Apr;33(2):357-361. [doi: 10.1111/1742-6723.13756] [Medline: 33675104]
2. Mergen M, Meyerheim M, Graf N. Reviewing the current state of virtual reality integration in medical education - a scoping review protocol. Syst Rev 2023 Jun 20;12(1):97. [doi: 10.1186/s13643-023-02266-6] [Medline: 37337293]
3. Horne EF, Thornton SW, Leraas HJ, O'Brian R, Greenwald E, Tracy ET. Use of pediatric trauma simulations to facilitate exposure to pediatric trauma resuscitations during training. Surgery 2023 Dec;174(6):1334-1339. [doi: 10.1016/j.surg.2023.08.004] [Medline: 37748976]
4. Topjian AA, Raymond TT, Atkins D, et al. Part 4: Pediatric basic and advanced life support: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2020 Oct 20;142(16_suppl_2):S469-S523. [doi: 10.1161/CIR.0000000000000901] [Medline: 33081526]
5. McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. Acad Med 2011 Jun;86(6):706-711. [doi: 10.1097/ACM.0b013e318217e119] [Medline: 21512370]
6. Roussin C, Sawyer T, Weinstock P. Assessing competency using simulation: the SimZones approach. BMJ Simul Technol Enhanc Learn 2020;6(5):262-267. [doi: 10.1136/bmjstel-2019-000480] [Medline: 35517396]
7. Zhao G, Fan M, Yuan Y, Zhao F, Huang H. The comparison of teaching efficiency between virtual reality and traditional education in medical education: a systematic review and meta-analysis. Ann Transl Med 2021 Feb;9(3):252. [doi: 10.21037/atm-20-2785] [Medline: 33708879]
8. Kyaw BM, Saxena N, Posadzki P, et al. Virtual reality for health professions education: systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959. [doi: 10.2196/12959] [Medline: 30668519]
9. Barteit S, Lanfermann L, Bärnighausen T, Neuhann F, Beiersmann C. Augmented, mixed, and virtual reality-based head-mounted devices for medical education: systematic review. JMIR Serious Games 2021 Jul 8;9(3):e29080. [doi: 10.2196/29080] [Medline: 34255668]
10. Rauschnabel PA, Felix R, Hinsch C, Shahab H, Alt F. What is XR? Towards a framework for augmented and virtual reality. Comput Human Behav 2022 Aug;133:107289. [doi: 10.1016/j.chb.2022.107289]
11. Bracq MS, Michinov E, Jannin P. Virtual reality simulation in nontechnical skills training for healthcare professionals: a systematic review. Simul Healthc 2019 Jun;14(3):188-194. [doi: 10.1097/SIH.0000000000000347] [Medline: 30601464]
12. Choudhury N, Gélinas-Phaneuf N, Delorme S, Del Maestro R. Fundamentals of neurosurgery: virtual reality tasks for training and evaluation of technical skills. World Neurosurg 2013 Nov;80(5):e9-e19. [doi: 10.1016/j.wneu.2012.08.022] [Medline: 23178917]
13. Foronda CL, Gonzalez L, Meese MM, et al. A comparison of virtual reality to traditional simulation in health professions education: a systematic review. Simul Healthc 2024 Jan 1;19(1S):S90-S97. [doi: 10.1097/SIH.0000000000000745] [Medline: 37651101]
14. Goulet-Pelletier JC, Cousineau D. A review of effect sizes and their confidence intervals, part I: the Cohen's d family. Quant Methods Psychol 2018;14(4):242-265. [doi: 10.20982/tqmp.14.4.p242]
15. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A Product from the ESRC Methods Programme: Lancaster University; 2006. [doi: 10.13140/2.1.1018.4643]
16. 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] [Medline: 36535832]
17. Lietz A, Kraller J, Hoffelner A, Ritschl V, Berger A, Wagner M. Dose-response of virtual reality training of paediatric emergencies in a randomised simulation-based setting. Acta Paediatr 2023 Sep;112(9):1995-2005. [doi: 10.1111/apa.16847] [Medline: 37195147]
18. Kitapcioglu D, Aksoy ME, Ozkan AE, Usseli T. Comparing learning outcomes of machine-guided virtual reality-based training with educator-guided training in a metaverse environment: randomized controlled trial. JMIR Serious Games 2024 Aug 7;12:e58654. [doi: 10.2196/58654] [Medline: 39110497]
19. Jaskiewicz F, Kowalewski D, Starosta K, Cierniak M, Timler D. Chest compressions quality during sudden cardiac arrest scenario performed in virtual reality: a crossover study in a training environment. Medicine (Baltimore) 2020 Nov 25;99(48):e23374. [doi: 10.1097/MD.0000000000023374] [Medline: 33235109]
20. Mühling T, Späth I, Backhaus J, et al. Virtual reality in medical emergencies training: benefits, perceived stress, and learning success. Multimed Syst 2023 Aug;29(4):2239-2252. [doi: 10.1007/s00530-023-01102-0]
21. Khanal P, Vankipuram A, Ashby A, et al. Collaborative virtual reality based advanced cardiac life support training simulator using virtual reality principles. J Biomed Inform 2014 Oct;51:49-59. [doi: 10.1016/j.jbi.2014.04.005] [Medline: 24732098]
22. Ma J, Wang Y, Joshi S, et al. Using immersive virtual reality technology to enhance nursing education: a comparative pilot study to understand efficacy and effectiveness. Appl Ergon 2024 Feb;115:104159. [doi: 10.1016/j.apergo.2023.104159] [Medline: 37924578]

XSL•FO
RenderX

23. Abulfaraj MM, Jeffers JM, Tackett S, Chang T. Virtual reality vs. high-fidelity mannequin-based simulation: a pilot randomized trial evaluating learner performance. Cureus 2021 Aug;13(8):e17091. [doi: 10.7759/cureus.17091] [Medline: 34527478]
24. Shah AS, Sobolewski B, Chon S, et al. Just-in-time, just-in-place virtual training in the pediatric emergency department: a novel approach to impact the perfusion exam. Adv Med Educ Pract 2023;14:901-911. [doi: 10.2147/AMEP.S414022] [Medline: 37614829]
25. Lamb A, McKinney B, Frousiakis P, Diaz G, Sweet S. A comparative study of traditional technique guide versus virtual reality in orthopedic trauma training. Adv Med Educ Pract 2023;14:947-955. [doi: 10.2147/AMEP.S395087] [Medline: 37693298]
26. Cardoso SA, Suyambu J, Iqbal J, et al. Exploring the role of simulation training in improving surgical skills among residents: a narrative review. Cureus 2023 Sep;15(9):e44654. [doi: 10.7759/cureus.44654] [Medline: 37799263]
27. Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. Acad Emerg Med 2008 Nov;15(11):988-994. [doi: 10.1111/j.1553-2712.2008.00227.x] [Medline: 18778378]
28. Au K, Lam D, Garg N, et al. Improving skills retention after advanced structured resuscitation training: a systematic review of randomized controlled trials. Resuscitation 2019 May;138:284-296. [doi: 10.1016/j.resuscitation.2019.03.031] [Medline: 30928503]
29. Anderson R, Sebaldt A, Lin Y, Cheng A. Optimal training frequency for acquisition and retention of high-quality CPR skills: a randomized trial. Resuscitation 2019 Feb;135:153-161. [doi: 10.1016/j.resuscitation.2018.10.033] [Medline: 30391370]
30. Gugelmin-Almeida D, Tobase L, Maconochie I, Polastri T, Rodrigues Gesteira EC, Williams J. What can be learned from the literature about intervals and strategies for paediatric CPR retraining of healthcare professionals? A scoping review of literature. Resusc Plus 2022 Dec;12:100319. [doi: 10.1016/j.resplu.2022.100319] [Medline: 36337082]
31. 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 May;26(2):683-711. [doi: 10.1007/s10459-020-10000-3] [Medline: 33128662]
32. Liaw SY, Tan JZ, Bin Rusli KD, et al. Artificial intelligence versus human-controlled doctor in virtual reality simulation for sepsis team training: randomized controlled study. J Med Internet Res 2023 Jul 26;25:e47748. [doi: 10.2196/47748] [Medline: 37494112]
33. Buttussi F, Chittaro L, Valent F. A virtual reality methodology for cardiopulmonary resuscitation training with and without a physical mannequin. J Biomed Inform 2020 Nov;111:103590. [doi: 10.1016/j.jbi.2020.103590] [Medline: 33039589]
34. Berg H, Steinsbekk A. The effect of self-practicing systematic clinical observations in a multiplayer, immersive, interactive virtual reality application versus physical equipment: a randomized controlled trial. Adv Health Sci Educ Theory Pract 2021 May;26(2):667-682. [doi: 10.1007/s10459-020-10019-6] [Medline: 33511505]
35. Abensur Vuillaume L, Goffoy J, Dubois N, et al. Collaborative virtual reality environment in disaster medicine: moving from single player to multiple learners. BMC Med Educ 2024 Apr 19;24(1):422. [doi: 10.1186/s12909-024-05429-8] [Medline: 38641770]
36. Alam F, LeBlanc VR, Baxter A, et al. Age and its impact on crisis management performance and learning after simulation-based education by acute care physicians: a multicentre prospective cohort study. Br J Anaesth 2024 Feb;132(2):383-391. [doi: 10.1016/j.bja.2023.10.040] [Medline: 38087740]
37. Elendu C, Amaechi DC, Okatta AU, et al. The impact of simulation-based training in medical education: a review. Medicine (Baltimore) 2024 Jul 5;103(27):e38813. [doi: 10.1097/MD.0000000000038813] [Medline: 38968472]
38. Mergen M, Graf N, Meyerheim M. Reviewing the current state of virtual reality integration in medical education - a scoping review. BMC Med Educ 2024 Jul 23;24(1):788. [doi: 10.1186/s12909-024-05777-5] [Medline: 39044186]

Abbreviations

CPR: cardiopulmonary resuscitation
PICO: Population, Intervention, Comparison, and Outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Literature Search Extension
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
VR: virtual reality

XSL•FO
RenderX

Edited by I Steenstra; submitted 17.Sep.2025; peer-reviewed by N Stathakarou, Y Yang; revised version received 06.Jul.2026; accepted 06.Jul.2026; published 31.Jul.2026.

Please cite as:
Costa M, Codato M, Brugnera G, Galiazzo S, Cantele E, Chang TP, Bressan S
Different Delivery Modalities of Virtual Reality Training in Emergency Medicine: Systematic Review
JMIR XR Spatial Comput 2026;3:e84310
URL: https://xr.jmir.org/2026/1/e84310
doi:10.2196/84310

© Marianna Costa, Martina Codato, Giorgia Brugnera, Silvia Galiazzo, Elisabetta Cantele, Todd P Chang, Silvia Bressan. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 31.Jul.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy of Virtual Reality–Based Mindfulness Interventions: Systematic Review and Meta-Analysis

Sarah Alicia Barker[1], MFA, PhD; Andreas J Miles-Novelo[1], PhD; Albert Rizzo[2], PhD; Regina M Tuma[1], PhD

[1]Psychology Department, Fielding Graduate University, 2020 De La Vina Street, Santa Barbara, CA, United States

[2]Department of Psychiatry and Behavioral Sciences, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Sarah Alicia Barker, MFA, PhD
Psychology Department, Fielding Graduate University, 2020 De La Vina Street, Santa Barbara, CA, United States

Abstract

Background: Virtual reality–based mindfulness interventions (VRbMIs) increasingly populate studies as scalable tools for stress and emotion regulation. However, findings across psychological outcomes are heterogeneous, and methodological variation in intervention design, outcome measurement, and reporting practices limits cross-study comparability and cumulative synthesis.

Objective: A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020–compliant systematic review and meta-analysis was conducted to evaluate the psychological effects of VRbMIs published between 2023 and 2025, examining methodological quality and outcome consistency across diverse study designs.

Methods: A PRISMA 2020–compliant systematic search was conducted across PubMed, PsycINFO, Scopus, Semantic Scholar, and CORE from January 2023 to April 2025. Eligible studies included quantitative VRbMIs reporting psychological outcomes. We assessed risk of bias using the Mixed Methods Appraisal Tool and conducted random-effects meta-analyses using Hedges *g* where data were sufficient.

Results: VRbMIs were associated with a large, statistically significant reduction in negative affect and a statistically significant increase in state mindfulness, while effects on depression, stress, and anxiety were small to moderate and nonsignificant. Trait mindfulness was described narratively rather than meta-analyzed and showed limited, inconsistent change across studies. For anxiety, we included 13 studies contributing 27 effect size estimates in the quantitative synthesis. A random-effects model indicated a small-to-moderate, nonsignificant pooled effect (g=0.44, 95% CI –0.08 to 0.96; t_{26}=1.74; P=.09), with substantial between-study heterogeneity (Q_{26}=381.60; P<.001; τ^2=0.96). Physiological outcomes were reported across a subset of studies and generally aligned with self-reported psychological findings; however, inconsistent measurement and incomplete reporting precluded quantitative synthesis. This analysis represents 35 studies with a combined sample of approximately 1550 participants (1 study did not report sample size).

Conclusions: VRbMIs demonstrate statistically significant short-term benefits for negative affect and state mindfulness, with consistently positive but nonsignificant trends toward improvement in stress, anxiety, and depression. Effects on positive affect and trait mindfulness were small and less consistent. Physiological findings were promising but limited by inconsistent reporting. These results support the use of VRbMIs as accessible tools for emotional regulation across diverse populations, while highlighting the need for larger trials, standardized outcome reporting, and more rigorous control conditions to strengthen the evidentiary foundation of this rapidly evolving field.

Trial Registration: PROSPERO CRD420251178827; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251178827

(JMIR XR Spatial Comput 2026;3:e90003) doi:10.2196/90003

KEYWORDS

virtual reality–based mindfulness interventions; systematic review; meta-analysis; stress; psychological; affect; mindfulness

Introduction

Background

Mindfulness-based interventions (MBIs) have consistently demonstrated benefits for stress reduction, emotion regulation, and overall well-being across clinical and nonclinical populations [1-3]. Virtual reality (VR) technologies provide an immersive delivery modality for mindfulness interventions through multisensory environments that support sustained attention and embodied engagement. Virtual reality–based mindfulness interventions (VRbMIs) integrate contemplative instruction with 3D or 360° environments, allowing participants to practice focused attention and relaxation within simulated natural or therapeutic contexts [1,4,5]. Early evidence indicates that VRbMIs are associated with physiological markers of

XSL•FO
RenderX

regulation, including increased heart rate variability (HRV) and alpha electroencephalography (EEG) activity, alongside improvements in stress, anxiety, and mood [4,6-8]. However, findings across psychological outcomes are heterogeneous, and methodological variation in intervention design, outcome measurement, and reporting practices limits cross-study comparability and cumulative synthesis.

Prior Systematic Reviews and Rationale for This Study

Several systematic reviews have examined VRbMIs, though each is constrained in scope in ways that limit cumulative synthesis. Ma et al [9] conducted the first narrative systematic review of immersive VR mindfulness training, including 7 studies focused specifically on whether immersion level moderates mindfulness outcomes. While foundational, that review included no meta-analysis and predates the substantial growth in the literature from 2023 onward. Wieczorek et al [4] conducted the first systematic review and evidence mapping of VRbMIs (search to September 2022), documenting consistent psychological and physiological benefits across 22 studies, including improvements in anxiety, mindfulness, affect, and stress, but explicitly did not perform a meta-analysis due to high methodological heterogeneity across studies. Milasi et al [10] reviewed 16 randomized controlled trials in nonclinical populations only, excluding the clinical, occupational, and mixed methods designs that characterize much of the recent literature. Xie et al [11] conducted a meta-analysis of 25 studies focused specifically on mindfulness as an outcome, without synthesizing the broader psychological profile of VRbMI effects across multiple outcome domains.

No completed review has synthesized empirical studies published between 2023 and 2025, incorporated psychological outcomes across diverse study designs, or evaluated methodological quality using the Mixed Methods Appraisal Tool (MMAT). This review addresses these gaps directly. By extending the search window to April 2025 and including quantitative, quasi-experimental, and mixed methods designs, this study provides an updated and more comprehensive assessment of the VRbMI evidence base than prior reviews have offered.

Literature Context

MBIs have demonstrated benefits for psychological well-being, including reductions in stress, anxiety, and negative affect, across clinical and nonclinical populations [1-3]. As immersive technologies have matured, VR has emerged as a delivery modality for mindfulness practices that integrates guided contemplative instruction, such as breath awareness or body-focused attention, with immersive 3D or 360° environments [4]. VRbMIs typically combine standardized mindfulness guidance with simulated natural or restorative settings, enabling controlled, multisensory delivery of practice within laboratory, clinical, and applied contexts [4]. Empirical studies have shown that brief and multisession VR mindfulness exposures can reduce self-reported stress, anxiety, and negative affect, while also producing physiological indicators associated with relaxation, including changes in heart rate (HR) and HRV [4,6-8]. However, effect magnitudes vary substantially across studies, reflecting differences in intervention duration, guidance style, population characteristics, and the measurement of outcomes. These variations complicate direct comparison across trials and limit conclusions about the consistency and durability of effects.

Wieczorek et al [4] conducted the first systematic review and evidence mapping of VRbMIs, documenting psychological benefits across 22 studies, including improvements in anxiety, mindfulness, affect, and stress—alongside smaller, less consistently reported physiological effects. Notably, the authors explicitly declined to perform meta-analytic pooling due to high methodological heterogeneity across included studies. That review identified several methodological constraints, including small sample sizes, heterogeneity in intervention formats, and limited standardization in physiological reporting [4]. Notably, the review concluded before the publication of a substantial wave of studies between 2023 and 2025. More recent investigations have expanded the evidence base through the inclusion of multisession protocols, active control conditions, and improved VR hardware fidelity, reflecting maturation in both study design and implementation [5-8,10].

Across the emerging literature, researchers have applied VRbMIs in health care, educational, and occupational settings, most commonly targeting anxiety, depression, stress, affective state, and self-reported mindfulness. Physiological end points have included HRV, electrodermal activity, EEG, and cortisol measures, intended to complement subjective outcomes [4]. Despite this breadth, methodological variability remains high. Many studies rely on single-session exposures, convenience samples, or passive control conditions, while others differ in intervention dose and sensory complexity. This diversity contributes to heterogeneity in reported effects and limits the aggregation of cross-study results.

Measurement practices further constrain comparability. Researchers most often assess psychological outcomes using symptom-focused self-report instruments such as the Depression Anxiety Stress Scales, the State-Trait Anxiety Inventory, and the Positive and Negative Affect Schedule. These measures emphasize changes in distress and affective state, reflecting dominant operationalizations used in intervention research. Physiological outcomes, by contrast, are frequently reported incompletely, with missing descriptive statistics or inconsistent baselines that preclude quantitative synthesis [4]. Greater consistency in outcome selection and reporting is therefore necessary to support cumulative evaluation of VRbMIs.

To address these limitations, this meta-analysis integrates findings from the most recent wave of empirical VRbMI research, using PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 reporting standards and random-effects procedures that are suitable for heterogeneous, small-sample literatures [12]. The review evaluates methodological quality using the MMAT [13] and quantifies between-study heterogeneity. By updating the cumulative evidence base, this analysis clarifies the strength and consistency of VRbMI effects and identifies methodological priorities for future digital mindfulness research.

XSL•FO
RenderX

Methods

Study Design

A systematic review and quantitative meta-analysis of empirical studies was conducted to examine VRbMIs. We followed the PRISMA 2020 guidelines [12] and applied methodological standards appropriate for small-sample random-effects modeling [12,14].

Eligibility Criteria

We included peer-reviewed empirical studies that evaluated VRbMIs in human participants. Eligible studies implemented a mindfulness-based practice delivered through immersive VR technology (eg, head-mounted displays presenting 3D or 360° environments) and reported at least 1 psychological or physiological outcome relevant to stress, affect, or well-being. We required studies to provide sufficient quantitative data to permit the calculation of effect sizes. To ensure consistent application of inclusion criteria, a mindfulness centrality rubric was developed and applied during full-text screening to evaluate whether each study's intervention met threshold criteria for a recognized mindfulness-based practice. The mindfulness centrality rubric is provided in Multimedia Appendix 1.

While qualitative methods offer valuable insights into the subjective and phenomenological dimensions of mindfulness practice, including participants' lived experience of presence, immersion, and contemplative engagement, the primary aim of this review was to quantify effect sizes across standardized psychological and physiological outcomes to enable meta-analytic pooling. Qualitative data do not permit the calculation of standardized mean differences or the estimation of between-study heterogeneity that meta-analysis requires. This decision necessarily limits the review's capacity to capture experiential dimensions of VRbMIs that may not be adequately represented by symptom-focused instruments. A complementary qualitative synthesis or mixed methods review remains an important direction for future research.

This study encompasses randomized controlled trials, nonrandomized controlled studies, and pre- and postintervention designs to reflect the methodological diversity of the emerging VRbMI literature. Qualitative-only studies, theoretical papers, reviews, protocols, conference abstracts, and dissertations without peer-reviewed publication were excluded. Nonimmersive digital interventions (eg, mobile or desktop mindfulness programs without VR), studies that did not involve a mindfulness-based practice, and studies lacking adequate reporting of outcomes were also excluded, and we limited inclusion to studies involving adult participants (aged 18 years and older), published in English.

Search Strategy

A comprehensive search was conducted across 5 databases: PubMed, PsycINFO, Semantic Scholar, CORE, and Scopus—covering the period from January 2023 to April 2025. The Boolean framework used for all databases was:

- ("mindfulness-based intervention" OR "MBSR" OR "MBCT")
- AND ("virtual reality" OR "VR" OR "immersive virtual reality")
- AND ("stress" OR "anxiety" OR "depression" OR "HRV" OR "EEG" OR "cortisol")

We manually searched reference lists of included studies and prior reviews (eg, Wieczorek et al [4]) to identify additional relevant publications and removed duplicates before screening.

Screening and Study Selection

We screened titles and abstracts for relevance using predefined inclusion criteria and screened full text independently, applying the same criteria to confirm eligibility. A PRISMA 2020 flow diagram summarizes the number of records identified, screened, excluded, and included in the final synthesis. The author (SAB) represents the single screener, having performed the title, abstract, and full-text screening independently using predefined inclusion criteria. Artificial intelligence (AI)–assisted tools (Rayyan, SciSpace, and large language models) were used to support record organization and retrieval; however, we made all inclusion and exclusion decisions.

Of the 2943 records initially retrieved, 45 duplicates were removed, yielding 2898 unique titles and abstracts for screening. At the title and abstract stage, 2764 records were excluded. The most common reasons for exclusion were that the study did not use immersive VR technology, the study did not deliver a structured MBI, the publication type was ineligible (eg, review paper, editorial, theoretical piece, or protocol), or the study lacked sufficient data for synthesis. In cases where abstracts were unavailable, exclusion at this stage was limited to titles that clearly indicated misalignment with inclusion criteria. Of the 134 full-text papers assessed for eligibility, 99 were excluded using a structured tagging system within Rayyan. Exclusion reasons included mindfulness was peripheral or undefined within the intervention (excluded—score 1 peripheral), the study did not use immersive VR (excluded—not VR), the study lacked a structured MBI component (excluded—not MBI), or the full text was inaccessible due to paywall restrictions. A complete record of exclusion tags is available in the Open Science Framework (OSF) repository.

Data Extraction

For each eligible study, data were extracted into a standardized spreadsheet, including authors, publication year, and country; study design and sample size; participant characteristics (age and population type); intervention characteristics (duration, frequency, headset type, environment description, and guidance mode); comparator type; psychological and physiological outcome measures; and pre- and postintervention means, SDs, and sample sizes. When statistics (eg, SDs) were missing, we estimated them from other reported values (eg, 2-tailed t statistics or η^2) using formulas recommended in the Cochrane Handbook for Systematic Reviews of Interventions [15]. Additional variables extracted included participant characteristics, intervention features, and outcome measures. When data were missing or unclear, we applied conservative assumptions and standard estimation procedures and documented them during extraction. All extracted values were double-checked for accuracy. For each outcome domain, all

XSL•FO
RenderX

reported results that permitted effect size calculation were extracted. When multiple measures or time points were reported within a study, we chose the most comparable pre-post or between-group results aligned with the primary analysis. In cases where studies reported results narratively as statistically significant without sufficient statistics for direct calculation, we applied a conservative estimate of g=0.5; for marginal or nonsignificant outcomes, we used g=0.30, following the precedent established by Borenstein et al [16]. Studies for which effect sizes could not be calculated or estimated were excluded from the quantitative synthesis and reported narratively.

Data Preparation and Effect Size Computation

Following data extraction, all variables were organized with the explicit aim of computing standardized mean differences (Hedges g) for psychological outcomes across studies. Extracted data included pre- and postintervention means, SDs, and sample sizes for both within-subject and between-group designs, where available. When studies did not report sufficient statistical detail for direct calculation, consistent estimation procedures were applied to enable inclusion in the quantitative synthesis. We derived SDs from reported SEs or CIs using established formulas ($SD=SE \times \sqrt{n}$; $SD=[\text{upper CI}-\text{lower CI}]/[2\times t_{crit}]$). For within-subject designs lacking a reported pre- and postcorrelation, a conservative estimate of r=0.50 was applied, consistent with recommendations for meta-analytic practice [15,17].

In cases where only test statistics or P values were reported, effect sizes were calculated using established transformation procedures [15,17]. When studies reported outcomes narratively without sufficient statistical detail, conservative estimates were assigned (g=0.50 for statistically significant effects; g=0.30 for marginal or nonsignificant findings), following precedent in prior meta-analytic work [4]. For outcome measures in which higher scores indicated greater symptom severity, effect sizes were reverse-coded so that positive values consistently reflected improvement. When multiple instruments assessed the same construct within a study, we selected the most used measure across the dataset to maintain comparability. All effect sizes were calculated as Hedges g with corresponding SEs and variances, which were subsequently used in random-effects meta-analytic models [14,18,19]. All pooled effects were evaluated using 2-tailed tests (α=.05). We used random-effects models to account for between-study variability.

Risk-of-Bias Assessment

The methodological quality was evaluated using the MMAT 2018 [13]. Certainty of evidence was assessed qualitatively based on study quality (MMAT ratings), consistency of findings, and the extent of between-study heterogeneity. We rated each study across 5 domains: sampling, measurement, confounding, data completeness, and analysis. The ratings were summarized as low, moderate, or high risk of bias. When discrepancies arose during coding, we resolved them by verifying the source texts. Risk-of-bias assessments were conducted by a single reviewer (SAB) using the MMAT 2018 criteria. All ratings were performed independently, with AI-assisted tools used only to support identification of methodological details and not to determine final judgments.

Effect-Size Calculation and Data Synthesis

We computed effect sizes for continuous outcomes as standardized mean differences (Hedges g) with corresponding SEs and 95% CIs. For within-subject designs, we assumed a conservative pre- and postcorrelation of r=0.50 when studies did not report this value. Random-effects models were used for all pooled analyses to account for expected between-study heterogeneity, estimating between-study variance with the DerSimonian-Laird method [14]. Although DerSimonian-Laird estimation is standard in small-sample meta-analyses, it may underestimate uncertainty under conditions of extreme heterogeneity; results should therefore be interpreted conservatively [14]. We included studies in the quantitative synthesis when sufficient statistical data were available to compute effect sizes. We retained studies that met the inclusion criteria but did not report adequate data for narrative synthesis and prepared data for synthesis by converting reported statistics into standardized mean differences (Hedges g). When summary statistics were incomplete, conservative estimation procedures were applied to derive effect sizes.

Outcome Classification

We selected the 6 primary psychological outcome domains synthesized in this review—anxiety, depression, stress, negative affect, positive affect, and mindfulness—a priori based on 2 converging rationales. First, these domains were retained from Wieczorek et al [4] to ensure methodological continuity and enable direct comparison with the prior evidence base [4]. Second, they represent the outcome categories most consistently reported across the VRbMI literature, as identified during preliminary screening. Stress, anxiety, and depression reflect the dominant clinical targets of MBIs in Western psychological research [2,4]; positive and negative affect capture the broader affective profile of intervention response [4]; and mindfulness—assessed as both state and trait—directly indexes the contemplative construct central to all included interventions [1,2]. Domains such as sleep, cognition, and pain were reported by a subset of studies but were insufficiently represented across the sample to support meta-analytic pooling and were therefore retained for narrative description only. The distribution of outcome domains across the 35 included studies was as follows: anxiety (n=17, 49%), depression (n=17, 49%), stress (n=10, 29%), mindfulness (n=18, 51%), negative affect (n=4, 11%), and positive affect (n=4, 11%).

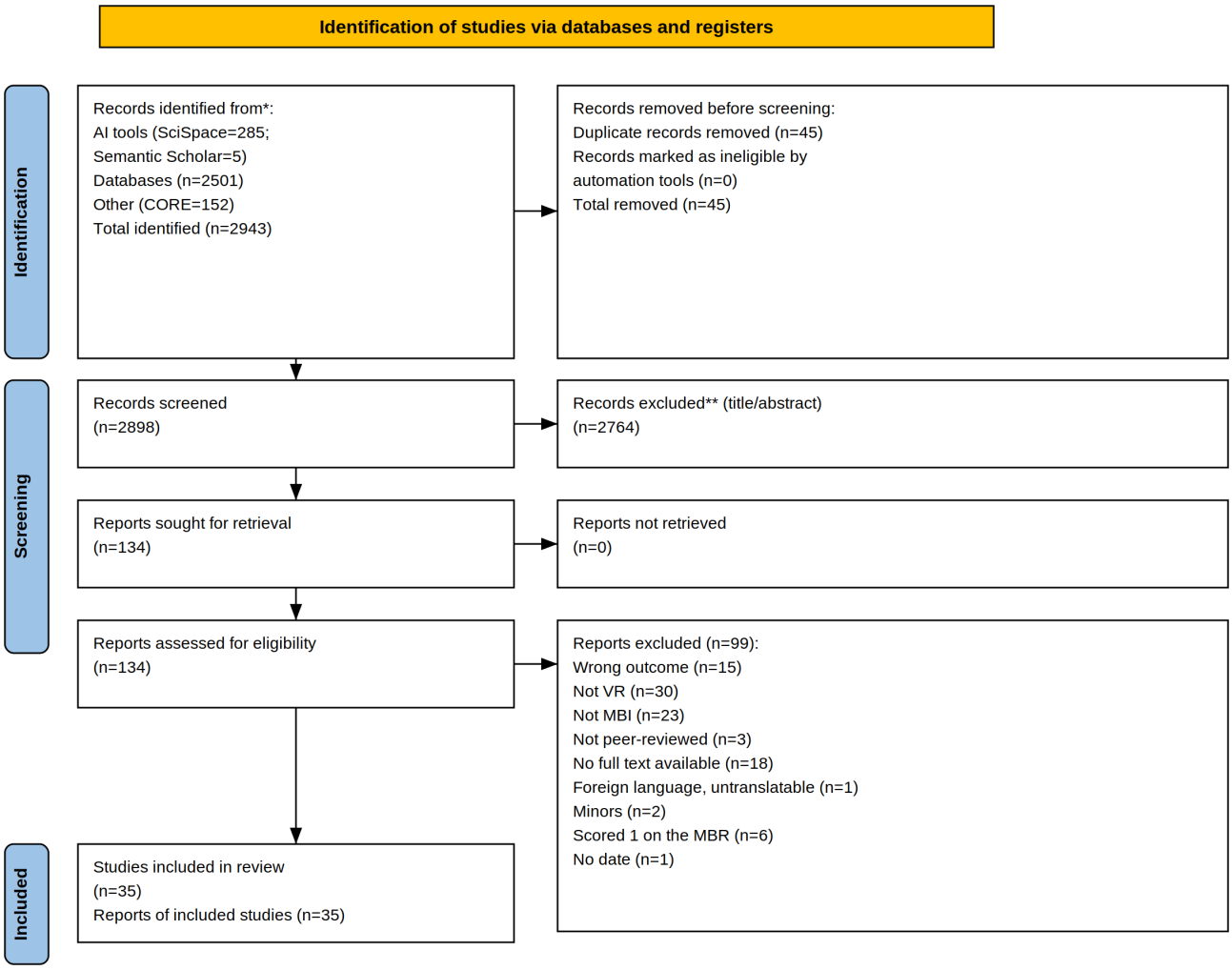
Results

Study Selection

Following the PRISMA 2020 reporting guidelines, study identification, screening, and inclusion were documented using a transparent flow diagram [12]. The search identified 2943 records from bibliographic databases and AI-assisted retrieval tools. After removing duplicates, we screened 2898 unique titles and abstracts and assessed 134 full-text papers for eligibility. Of these, 35 studies met all inclusion criteria and were included in the final synthesis. Figure 1 presents the PRISMA flow diagram summarizing the screening process and study selection. The completed PRISMA 2020 checklist is provided in Checklist 1.

XSL•FO
RenderX

Figure 1. PRISMA 2020 flow diagram of study selection. The diagram depicts the number of records identified, screened, excluded, and retained in the final synthesis. *Records identified using AI-assisted tools (SciSpace, Semantic Scholar) and traditional databases. **Reasons for exclusion at the title/abstract stage were not individually categorized; the total reflects records not meeting VR-MBI criteria, empirical study requirements, or relevance thresholds. AI: artificial intelligence; MBI: mindfulness-based intervention; MBR: mindfulness centrality rubric; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses; VR: virtual reality.



Study Characteristics

The 35 included studies encompassed a range of quantitative and mixed methods designs, including randomized controlled trials, quasi-experimental studies, and nonrandomized formats. The studies examined clinical, educational, and occupational populations, including individuals with depression, anxiety disorders, or psychosis [20,21], health care professionals experiencing burnout [22,23], patients with cancer [24,25], and university students engaged in mindfulness-based training [6,26,27]. McConnell et al [28] examined VR mindfulness for chronic low back pain in a 2-arm randomized trial, measuring disability and pain intensity across 12 sessions. Torres García et al [29] examined emotional discomfort in patients with breast cancer using the Emotional Discomfort Detection scale prior to first chemotherapy; while the VR group showed improvement trends, statistical reporting was insufficient for effect size computation, and the study was retained for narrative description only.

Geographically, the evidence base spanned North America, East Asia, and Europe, reflecting the growing international interest in VRbMIs. Studies were conducted in the United States [22,25,30-33], China [6,24,34], Taiwan [7], Finland [35], Spain [5,26,36], the United Kingdom [37], South Korea [21,38,39], and Turkey [40]. Supplementary visual summaries of study characteristics, including demographic figures and geographic distribution, are available in the OSF repository. Table 1 presents detailed characteristics of all included studies, including study design, population, intervention features, and outcome measures, in accordance with PRISMA 2020 reporting standards.

XSL•FO
RenderX

Table . Characteristics of included studies (N=35).

Study	Design	Values, n	Population	Intervention (dose)	Model	Guide	Comparator	Outcomes	Time	MMAT[a]	*g*
Ch et al (2023) [13]	Experiment	20	Remote workers	VR[b] mindfulness (9 weeks)	360° video	Mixed	Control	Stress, creativity	Pre or post	Low-moderate	−0.57
Barton et al (2024) [35]	Mixed	58	Adults	VR mindfulness (1×6 minutes)	Mixed or multimodal	Guided	VR conditions	Mindfulness, fatigue, mood	Pre or post	Moderate	1.22
Blackmore et al (2024) [20]	Mixed	27	Clinical	VR mindfulness (1×15 minutes)	360° video	Guided	None	Anxiety, mood, mindfulness	Pre or post	Moderate	1.26
Cano et al (2024) [5]	Quasi-experiment	31	Post-COVID-19	VR mindfulness (16×60 minutes)	CAVE[c]	Guided	None	Cognition, attention	Pre or post	High	0.01
Cawley and Tejeiro (2024) [27]	Quasi-experiment	67	Students	VR mindfulness (1×10 minutes)	360° video	Guided	Audio, coloring	Stress, well-being	Pre or post	Low-moderate	0.34
Chen et al (2024) [41]	RCT[d]	76	Adults	VR mindfulness (8×18 minutes)	App-based VR	Guided	Control	Anxiety, depression	Pre or post+FU	Low	0.99
Ferrer Costa et al (2024) [23]	Pilot	83	Health care workers	VR mindfulness (8×13 minutes)	Immersive VR	Guided	None	Burnout, engagement	Pre or post	Low-moderate	−0.56
Fonseca et al (2024) [31]	RCT	26	Caregivers	VR mindfulness (1×6 minutes)	Immersive VR	Guided	Control	Anxiety	Pre or post	Moderate	−0.75
Franklin et al (2023) [25]	RCT	36	Patients with cancer	VR mindfulness (1×10 minutes)	Immersive VR	Guided	Control	Anxiety, depression	Pre or post	Low	NR[e]
Garland et al (2024) [32]	Pilot	34	SUD[f]	VR mindfulness (8×15 minutes)	App-based VR	Guided	None	Craving, mood	Pre or post	Low	0.53
Han et al (2023) [39]	Experiment	40	Students	VR meditation (1×10 minutes)	Immersive VR	Guided	Control	Anxiety	Pre or post	Moderate	N/A[g]
Hanna et al (2025) [37]	Quasi-experiment	32	Students	VR mindfulness (4 - 8 weeks)	Immersive VR	Guided	No VR	Mindfulness	Pre or post	Moderate	N/A

XSL•FO
RenderX

Study	Design	Values, n	Population	Intervention (dose)	Model	Guide	Comparator	Outcomes	Time	MMAT[a]	g
Jimenez-Barragan et al (2025) [36]	RCT	70	Pregnant women	VR mindfulness (6 weeks, daily 14 minutes)	Immersive VR	Guided	Standard care	Pain, anxiety	Pre or post	Moderate	3.92
Kamada et al (2025) [42]	RCT	10	Patients undergoing surgery	VR mindfulness (NR[e])	Immersive VR	Guided	Standard care	Pain, anxiety	Pre or post	Moderate	−1.49
Kim et al (2024) [38]	Experiment	8	Dementia	VR mindfulness (6 sessions)	Immersive VR	Guided	None	Anxiety, depression	Pre or post	Lo	N/A
Kim et al (2024) [43]	Experiment	60	Students	VR meditation (5×30 minutes)	360° video	Guided	Control	Sleep, HRV[h]	Pre or post	Low	1.28
Kumar et al (2024) [44]	RCT	40	Students	VR meditation (3×10 minutes)	Immersive VR	Guided	Control	EEG[i], stress	Pre or post	Low-moderate	1.33
Lee et al (2023) [21]	RCT	64	Psychosis	VR mindfulness (8×30 minutes)	Mixed or multimodal	Guided	Control	Symptoms	Pre or post	Moderate	0.52
Liu et al (2025) [34]	RCT	90	Postpartum	VR mindfulness (8 weeks)	Immersive VR	Guided	Control	Anxiety, depression	Pre or post+FU[j]	Moderate	0.83
Mancini et al (2024) [45]	Experiment	21	Adults	VR mindfulness (NR)	Interactive	Guided	Audio	Mindfulness	Pre or post	Moderate	N/A
Mao et al (2024) [24]	Experiment	48	Patients with cancer	VR mindfulness (4×60 minutes)	360° video	Guided	None	Anxiety, fatigue	Pre or post+FU	Low-Moderate	1.93
McConnell et al (2024) [28]	RCT	52	Chronic pain	VR mindfulness (12	Mixed or multimodal	Guided	Standard care	Pain, disability	Pre or post	High	1.33
Modrego-Alarcón et al (2025) [26]	RCT	93	Students	VR mindfulness (6×7.5 minutes)	Immersive VR	Guided	Control	Mindfulness	Pre or post	Low	1.73
Murray et al (2024) [46]	Mixed	38	TBI[k]	VR mindfulness (4-minute sessions)	Immersive VR	Guided	None	Distress	Pre or post	Low	0.68
Ng et al (2024) [7]	Experiment	51	Students	VR breathing (4×6 minutes)	Immersive VR	Mixed	Control	EEG	Pre or post	Moderate	N/A

XSL•FO
RenderX

Study	Design	Values, n	Population	Intervention (dose)	Model	Guide	Comparator	Outcomes	Time	MMAT[a]	*g*
Olasz et al (2024) [8]	RCT	56	Students	VR mindfulness (2×20 minutes)	Immersive VR	Guided	Tablet	Anxiety, HR[l]	Pre or post	Moderate	1.07
Ong et al (2025) [47]	Mixed	35	Educators	VR mindfulness (1×10 minutes)	360° video	Guided	None	Mindfulness, UX[m]	Pre or post	Low	N/A
Poetar et al (2023) [48]	RCT	47	Adults	VR mindfulness (1×30 minutes)	App-based VR	Guided	Desktop	Mood	Pre or post	Moderate-high	0.58
Sexton-Radek et al (2024) [49]	Mixed	3	Students	VR mindfulness (7×15 minutes)	Immersive VR	Guided	None	Sleep	Pre or post	High	N/A
Sezer et al (2025) [40]	Experiment	54	Adults	VR mindfulness (1×20 minutes)	Immersive VR	Guided	Control	Anxiety, HRV	Pre or post	Moderate	0.50
Spitz et al (2024) [33]	Pilot	27	Dysphoria	VR mindfulness (2 weeks)	Immersive VR	Guided	None	Mood	Pre or post	Moderate	N/A
Torres García et al (2024) [29]	Quantitative (NR)	NR	Adults	VR mindfulness (NR)	Immersive VR	Guided	None	Emotional discomfort	Pre or post	High	N/A
Van Doren et al (2024) [30]	Mixed	32	Veterans	VR mindfulness (1×60 minutes)	360° video	Guided	None	Mindfulness	Pre or post	Moderate	−0.98
Williams et al (2024) [22]	Experiment	61	Health care workers	VR mindfulness (5 - 15 minutes)	Immersive VR	Guided	None	Stress, engagement	Pre or post	Moderate	1.01
Zheng et al (2024) [6]	RCT	60	Students	VR mindfulness (2 weeks)	Immersive VR	Guided	Control	Anxiety, depression	Pre or post+FU	Low	1.81

[a]MMAT: Mixed Methods Appraisal Tool.

[b]VR: virtual reality.

[c]CAVE: room-scale CAVE system.

[d]RCT: randomized controlled trial.

[e]NR: not reported. Effect size could not be calculated due to the absence of a control or comparator group.

[f]SUD: substance use disorder.

[g] N/A: not available; effect size could not be calculated; study reported only *P* values and percentage change without group-level means and SDs necessary to compute Hedges *g*.

[h]HRV: heart rate variability.

[i]EEG: electroencephalography.

[j]FU: follow-up.

[k]TBI: traumatic brain injury.

[l]HR: heart rate.

[m]UX: user experience.

XSL•FO
RenderX

Risk of Bias Assessment

Using the MMAT [13], we evaluated the methodological quality of all 35 included studies across 5 domains: sampling, measurement, confounding, data completeness, and analysis. Overall, most studies were rated as having a moderate risk of bias (n=22), reflecting partial adherence to quality criteria, commonly due to limited reporting of randomization procedures, incomplete follow-up data, or small sample sizes. In total, 11 studies met most or all MMAT criteria and were rated low risk. In contrast, we rated 3 studies as high risk due to inadequate control of confounding variables or incomplete reporting of outcomes. Among randomized controlled trials, methodological quality was generally strong, with clear descriptions of participant allocation and intervention fidelity. However, many trials did not report whether outcome assessors were blinded, and adherence monitoring was often limited. Nonrandomized designs exhibited a higher risk of bias in sampling and confounding domains, primarily due to convenience recruitment and the absence of active control conditions. Mixed methods studies demonstrated a moderate to high risk of bias, primarily due to the limited integration between the qualitative and quantitative components.

AI-assisted tools supported this appraisal by facilitating the identification of relevant methodological details within full-text papers; we reviewed all AI-assisted outputs before assigning final MMAT ratings. Table 2 summarizes the distribution of risk-of-bias ratings by study design.

Table . Risk of bias summary of study design.

Domain and score (low, moderate, or high)	Count, n
Quantitative randomized controlled trial	
High risk	1
Moderate risk	5
Low risk	8
Quantitative nonrandomized	
High risk	1
Moderate risk	13
Low risk	0
Quantitative nonrandomized or mixed methods	
High risk	1
Moderate risk	4
Low risk	2

Quantitative Results

Overview

Following the risk-of-bias appraisal, quantitative outcomes were synthesized across the 35 included studies using standardized mean difference estimates. We conducted analyses using random-effects models to account for between-study variability and expressed all effect sizes as Hedges *g*, along with corresponding 95% CIs. The random-effects weighting was applied to minimize bias from unequal sample sizes and heterogeneous designs. When studies did not report complete statistics (eg, pre-post means, SDs, or sample sizes), we applied conservative estimation rules based on the direction and significance of reported effects, following Hedges [15] and guidelines for small-sample meta-analyses. All statistical transformations and pooling procedures were performed in JASP (version 0.19; University of Amsterdam) using the DerSimonian-Laird estimator for between-study variance [14], consistent with Hedges [15] framework and contemporary guidelines for small-sample meta-analyses [18].

The quantitative synthesis focused on 6 primary psychological outcome domains reported across multiple studies: anxiety, depression, stress, negative affect, positive affect, and mindfulness. Each outcome is summarized below, with pooled effect sizes, CIs, and heterogeneity statistics. Forest plots for all meta-analytic models, including study-level and pooled estimates, are available in the project's OSF repository. The following sections present each outcome domain sequentially, beginning with anxiety, the most frequently reported psychological variable in the included studies. Studies contributing to each synthesis varied in design, sample characteristics, and risk of bias, with most rated low to moderate quality using MMAT criteria.

Anxiety

In total, 27 effect size estimates from 13 studies provided sufficient data for meta-analysis. A random-effects model indicated a small-to-moderate, nonsignificant pooled effect on anxiety (g=0.44, 95% CI –0.08 to 0.96; t_{26}=1.74; P=.09). Between-study heterogeneity was substantial (Q_{26}=381.60; P<.001; τ^2=0.96), reflecting considerable variability in intervention design, population characteristics, and outcome measurement [15-17].

Depression

In total, 10 studies yielding 18 effect size estimates provided sufficient statistical data for meta-analysis. A random-effects

XSL•FO
RenderX

model yielded a moderate pooled effect (g=0.62, 95% CI –0.13 to 1.36]; t_{17}=1.75; P=.10); however, the CI crossed 0, indicating that the effect was not statistically significant. Between-study heterogeneity was substantial (Q_{17}=236.77; P<.001; τ^2=1.09) [15-17].

Stress

In total, 6 studies yielding 10 effect size estimates reported stress-related outcomes using validated self-report instruments, including the Perceived Stress Scale, the Depression, Anxiety and Stress Scale-21 stress subscale, or comparable measures [6,13,22,27,44,46]. Participant populations included university students, health care professionals, and community adults across both single-session and multiweek interventions. Individual study effect sizes varied across intervention formats. Some studies reported large reductions in perceived stress following immersive MBIs (eg, Kumar et al [44], g=1.33; Zheng et al [6], g=1.81), whereas others reported moderate or small effects following brief or lower-intensity interventions (eg, Murray and Shmidheiser [46], g=0.68; Cawley and Tejeiro [27], g=0.34). In occupational samples, effects were small or negligible [22]. A random-effects meta-analysis yielded a small-to-moderate pooled effect that did not reach statistical significance (g=0.45, 95% CI −0.08 to 0.99; t_9=1.90; P=.09). Between-study heterogeneity was substantial (Q_9=54.26; P<.001; τ^2=0.35) [15-17].

Negative Affect

In total, 3 studies yielding 4 effect size estimates reported changes in negative affect using validated self-report instruments, including the Positive and Negative Affect Schedule (PANAS) and the Profile of Mood States [20,30,48]. Participant samples included university students, individuals enrolled in recovery programs, and nonclinical adult populations. Across studies, individual effect sizes indicated reductions in negative affect following MBIs delivered in VR and non-VR formats. Van Doren et al [30] reported a large pre- and postreduction following a brief immersive VR mindfulness session (g=–0.98), whereas Blackmore et al [20] observed a moderate reduction following a VR-supported mindfulness practice emphasizing curiosity and decentering (g=–0.60). Poetar et al [48] reported similar decreases in negative affect in both VR-based (g=–0.58) and desktop-based (g=–0.68) mindfulness conditions. A random-effects meta-analysis yielded a large and statistically significant pooled effect on negative affect (g=–0.67, 95% CI –0.92 to −0.42; t_3=–8.71; P=.003). Between-study heterogeneity was negligible (Q_3=0.93; P=.82; τ^2=0.00) [15-17].

Positive Affect

In total, 3 studies yielding 4 effect size estimates assessed positive affect outcomes using the PANAS or comparable mood indices [20,30,48]. Participant samples included university students, clinical groups, and general adult populations. Individual study findings were variable. Van Doren et al [30] reported a moderate increase in positive affect following a single immersive mindfulness session (g=0.80). Poetar et al [48] reported a slight increase in the VR condition (g=0.27) and a slight decrease in the desktop-based comparison condition (g=–0.31). Blackmore et al [20] observed a slight, nonsignificant increase in positive affect (g=0.28). A random-effects meta-analysis yielded a small, pooled effect that did not reach statistical significance (g=0.24, 95% CI –0.45 to 0.92; t_3=1.10; P=.35). Between-study heterogeneity was modest (Q_3=6.58; P=.09; τ^2=0.09) [15-17].

Mindfulness

In total, 18 studies assessed mindfulness-related outcomes using validated self-report instruments, including the State Mindfulness Scale, Five Facet Mindfulness Questionnaire, and Toronto Mindfulness Scale [20,26,27,30,35,37,45,47]. These measures captured both state and trait aspects of mindfulness following exposure to VRbMIs. Individual study findings varied across intervention formats and outcome measures. Cawley and Tejeiro [27] reported significant gains in state mindfulness following a brief VR session (d=0.69), whereas Blackmore et al [20] observed large pre- and posteffects for curiosity (g=1.26) and decentering (g=1.51). Modrego-Alarcón et al [26] reported smaller but measurable effects across multiple VR environments. Other studies reported modest changes following repeated exposure protocols.

In total, 2 studies yielding 5 effect size estimates using the State Mindfulness Scale provided sufficient data for quantitative synthesis. A random-effects meta-analysis yielded a large, statistically significant pooled effect on state mindfulness (g=1.00, 95% CI 0.32-1.68; t_4=4.11; P=.02). Between-study heterogeneity was moderate (Q_4=12.37; P=.02; τ^2=0.21), reflecting variability in VR environment type and delivery format across conditions [15-17].

Physiological Outcomes

Several studies included physiological measures alongside psychological outcomes; however, reporting was inconsistent and often incomplete. Standard indices included HR, HRV, electrodermal activity (or skin conductance level), EEG, and cortisol. Because most studies did not report sufficient statistics for quantitative pooling, physiological outcomes were summarized descriptively. Across studies, reported physiological changes generally corresponded with reductions in autonomic arousal following VR-based mindfulness exposure. Kim et al [38] reported increased HRV and reduced HR during guided VR meditation compared with a neutral VR control condition. Ng et al [7] observed decreased skin conductance levels and increased alpha EEG activity during mindfulness-based VR sessions. Williams et al [22] documented reductions in HR and respiration rate among health care professionals during immersive breath-awareness practice. Kamada et al [42] evaluated immersive VR for reducing intraoperative pain and anxiety in patients undergoing surgery in a single-center pilot randomized controlled trial conducted in Helsinki. Sexton-Radek et al [49] measured salivary melatonin, subjective calm, and sleep efficiency in a small mixed methods study, with participants reporting greater relaxation and modest improvements in sleep scores over 5 days; however, no statistical testing was conducted.

XSL•FO
RenderX

Only 5 studies provided analyzable physiological statistics, and baseline definitions varied substantially across designs. As a result, we did not conduct a formal meta-analysis of physiological outcomes. Heterogeneity in study design, small sample sizes, and variability in outcome reporting limited overall confidence in the evidence. Findings for stress-related outcomes supported greater confidence, while substantial heterogeneity across anxiety studies reduced confidence in those pooled estimates.

Discussion

Overview

This review synthesized evidence from 35 empirical studies examining VRbMIs to evaluate psychological outcomes and methodological quality. Across studies, VRbMIs were associated with favorable short-term psychological outcomes, most notably a large, significant reduction in negative affect and a large, significant increase in state mindfulness, alongside nonsignificant trends toward improvement in stress, anxiety, and depression. Pooled effect sizes ranged from small to moderate (g=0.24 - 0.45) for affect, stress, anxiety, and depression outcomes, and large for negative affect (g=–0.67) and state mindfulness (g=1.00). Physiological outcomes, including changes in HR, HRV, and electroencephalographic markers, generally aligned with self-reported psychological findings; however, inconsistent and incomplete reporting precluded a quantitative synthesis. Taken together, the evidence suggests that VRbMIs can function as short-term supports for emotion regulation and attentional training. Although several pooled effects did not reach statistical significance, the overall direction and consistency of findings across studies, particularly for stress-related outcomes, support a cautious interpretation of short-term benefit. At the same time, these findings highlight methodological constraints that continue to shape how mindfulness is operationalized and evaluated within this emerging body of research.

Psychological and Physiological Findings

Across diverse populations, including students, health care professionals, and clinical groups, VRbMIs showed directionally consistent improvements in stress, anxiety, and depressive symptoms, though pooled effects for these outcomes did not reach statistical significance consistent with the direction of prior meta-analytic findings [4]. Effects were generally larger in interventions that combined immersive environments with explicit mindfulness instruction or interactive guidance [6,20]. In contrast, outcomes related to positive affect and trait mindfulness were smaller and less consistent, suggesting that VRbMIs primarily support short-term emotional regulation rather than stable dispositional change. Most studies operationalized mindfulness using standardized psychological instruments such as the Depression, Anxiety and Stress Scale-21 (DASS-21), State-Trait Anxiety Inventory, and PANAS. While these measures facilitate cross-study comparability, they predominantly frame mindfulness in terms of symptom reduction and self-regulation. The inclusion rubric applied in this review excluded VR applications that focused solely on relaxation or entertainment, ensuring conceptual alignment with mindfulness-based practice.

Nevertheless, the resulting evidence base largely reflects a biomedical orientation that emphasizes stress management over broader contemplative dimensions. Of the 35 included studies, stress, anxiety, and depression were the 3 most frequently measured outcomes, while positive affect and mindfulness trait measures appeared in fewer than half of the studies, and no included study explicitly measured contemplative outcomes such as compassion, equanimity, insight, or nonattachment. This pattern mirrors broader critiques of secular mindfulness research, which has been characterized as prioritizing symptom reduction and clinical utility over the cultivation of wisdom, ethical development, and the liberative dimensions central to traditional Buddhist practice [3,50,51]. The dominance of distress-focused outcome measures reflects the influence of Western psychological frameworks, particularly cognitive-behavioral models on how mindfulness is operationalized and evaluated in digital health research.

Compared with Wieczorek et al [4], the present synthesis incorporates a broader and more methodologically diverse set of studies, including recent randomized trials and multisession interventions. While overall effect patterns remain broadly consistent, the inclusion of newer studies increases heterogeneity and reduces confidence in pooled estimates for several outcomes, particularly anxiety and mindfulness. These findings suggest that earlier estimates may have reflected smaller, more homogeneous samples.

Methodological and AI Contributions

Despite rapid growth in VRbMI research between 2023 and 2025, methodological rigor across studies remained uneven. Common limitations included small sample sizes, brief intervention durations, limited follow-up, and reliance on single-group pre- and postdesigns. These features constrained statistical power and reduced confidence in causal inference. Even among randomized controlled trials, control conditions often failed to match VR interventions in terms of duration, engagement, or sensory load, complicating the interpretation of intervention-specific effects. Heterogeneity in outcome measures, intervention protocols, and reporting formats further limited comparability across studies. The methodological and epistemic dimensions of the AI-assisted review process used in this study are examined in depth in a companion paper currently under review.

Conceptual Implications for Measurement and Design

Beyond procedural improvements, this review highlights a broader conceptual pattern in how mindfulness is represented within VR research. Nearly all included studies framed mindfulness as a culturally neutral attentional or stress-reduction technique, operationalized through symptom-focused instruments and evaluated against biomedical outcome criteria. While this approach supports cross-study comparability and clinical applicability, it narrows the construct of mindfulness to outcomes measurable within Western psychological paradigms. The epistemological and cultural dimensions of this narrowing, including how AI-assisted retrieval tools may

XSL•FO
RenderX

reproduce dominant publication biases and how outcome measures embed assumptions about selfhood and well-being, are examined in depth in a companion paper currently under review. This analysis focuses on the methodological and empirical implications of these patterns for the design and synthesis of VRbMI research.

Limitations and Future Directions

Several limitations constrain the interpretation of these findings. We did not conduct a sensitivity analysis due to the limited number of studies per outcome and substantial heterogeneity in study design and reporting, which constrained the feasibility of meaningful robustness testing. Many studies relied on small samples, brief follow-up periods, and incomplete statistical reporting, which reduced precision and limited the estimation of effect sizes. Substantial heterogeneity across interventions, outcome measures, and study designs further restricted comparability. Physiological outcomes showed potential value but were rarely reported with sufficient detail to support interpretation or synthesis. We did not conduct analyses to explore sources of heterogeneity (eg, subgroup analysis or meta-regression) due to the limited number of studies per outcome and variability in study designs. Further, we did not assess reporting bias (eg, publication bias) due to the small number of studies per outcome and heterogeneity in study designs.

At the review level, reliance on English-language publications likely introduced selection bias aligned with the dominant biomedical literature [52]. AI-assisted tools supported literature retrieval and screening. A complete methodological and epistemic analysis of this process appears in a separate paper currently under review and is therefore not reproduced here. Although the use of a structured inclusion rubric helped limit conceptual drift, future research should address both methodological and conceptual gaps. Larger trials with active control conditions and clearer reporting standards are necessary to clarify mechanisms and the durability of effects. For physiological measures in particular, a minimal reporting set specifying devices, sampling rates, preprocessing decisions, and measurement windows would enable cross-study comparison and longitudinal analysis. Beyond methodological rigor, future work should expand its conceptual scope by incorporating culturally grounded frameworks and qualitative or phenomenological methods capable of capturing the ethical, relational, and experiential dimensions of practice.

A further limitation concerns the sole-author design of this review. Although sole-authored systematic reviews are not uncommon in emerging fields, they introduce heightened risk of selection bias, confirmatory screening, and inconsistent application of inclusion criteria across a large study set. To mitigate these risks, we applied a structured, prespecified inclusion rubric documented in the PROSPERO registration and available in the OSF repository. We operationalized each criterion as a binary decision rule prior to screening and implemented a structured tagging system within Rayyan to document inclusion and exclusion decisions transparently at each stage. Each record was tagged with a specific reason code (eg, “not immersive VR,” “no mindfulness component,” and “insufficient data”) to ensure consistent and auditable decision-making throughout the screening process. Screenshots of the tagging workflow are available in the OSF repository. In cases where a study’s eligibility remained ambiguous, for example, when a VR application was described as “relaxation-based” but incorporated explicit breath-awareness or body-scan elements, we resolved uncertainty by returning to the primary source, applying the most conservative interpretation of the inclusion criterion, and documenting the decision rationale in the extraction log. Where ambiguity persisted after this process, we excluded the study. This approach prioritized consistency and transparency.

Implications

These findings have implications for practice, research design, and theory. In applied contexts, VRbMIs provide a flexible format for supporting stress and emotion regulation in educational, clinical, and workplace settings. Their value lies less in novelty than in their capacity to scaffold attention and engagement for populations that may struggle with conventional practice formats. For researchers, the synthesis highlights the importance of standardized outcome reporting and transparent documentation of review processes. Without shared reporting conventions, especially for physiological measures, the field limits its ability to build cumulative knowledge. Transparency in both measurement and retrieval practices remains essential for interpretability and reproducibility. At a theoretical level, the review demonstrates that the tools used to study mindfulness also participate in defining it. Aligning intervention design and measurement with contemplative concerns such as attention, care, and interdependence may help balance methodological precision with conceptual integrity. Under these conditions, technological innovation can support the continuity rather than dilute the mindfulness practice.

Conclusions

This synthesis of 35 empirical studies indicates that VRbMIs produce a large, statistically significant reduction in negative affect (g=–0.67; P=.003) and a large, statistically significant increase in state mindfulness (g=1.00; P=.02). Effects on depression (g=0.62; P=.10), stress (g=0.45; P=.09), and anxiety (g=0.44; P=.09) were small to moderate and did not reach statistical significance, though their direction was consistently positive. Effects on positive affect (g=0.24; P=.35) and trait mindfulness were small and less consistent, and physiological outcomes, while directionally aligned with self-reported findings, could not be formally synthesized due to incomplete reporting. Despite substantial heterogeneity across outcomes, the overall pattern of findings supports the use of VR-based mindfulness as a low-risk and accessible intervention across diverse populations. At the same time, the evidence base remains shaped by limited sample sizes, inconsistent physiological reporting, and a dominant biomedical framing that prioritizes symptom reduction over practice context. Future research that integrates rigorous trial design, transparent reporting practices, and culturally situated conceptual frameworks will better support both scientific credibility and theoretical coherence. Such integration may help ensure that VR-based mindfulness research advances not only technical effectiveness but also the broader

XSL•FO
RenderX

purpose of mindfulness as a practice oriented toward awareness, responsibility, and human flourishing.

Acknowledgments

SAB thanks her husband for his steady support and patience throughout the completion of this work. We used artificial intelligence (AI) tools (Rayyan, SciSpace, and ChatGPT) to support literature screening, data organization, and statistical transformation. All AI-assisted outputs were manually reviewed and verified against original source materials. The methodological and epistemic dimensions of this AI-assisted process are examined in full in a companion manuscript currently under review.

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 openly available in the Open Science Framework repository [53]. These materials include the extracted study data, computed effect sizes (Hedges *g*), analytic outputs, and supporting documentation. All data were derived from previously published studies; no individual participant data were collected or used.

Authors' Contributions

Conceptualization: SAB (lead)
Methodology: SAB
Software: SAB
Formal analysis: SAB
Investigation: SAB
Data curation: SAB
Writing—original draft: SAB
Writing—review and editing: SAB, AJM-N, RMT, AR
Visualization: SAB
Project administration: SAB
Supervision: AJM-N, RMT, AR

Conflicts of Interest

None declared.

Multimedia Appendix 1
Mindfulness centrality rubric: inclusion criteria and screening procedures for virtual reality–based mindfulness interventions.
[PDF File, 169 KB - xr_v3i1e90003_app1.pdf]

Checklist 1
PRISMA checklist.
[DOCX File, 281 KB - xr_v3i1e90003_app2.docx]

References

1. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003 Apr;84(4):822-848. [doi: 10.1037/0022-3514.84.4.822] [Medline: 12703651]
2. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. Clin Psychol Sci Pract 2003;10(2):144-156. [doi: 10.1093/clipsy.bpg016]
3. Van Dam NT, van Vugt MK, Vago DR, et al. Mind the hype: a critical evaluation and prescriptive agenda for research on mindfulness and meditation. Perspect Psychol Sci 2018 Jan;13(1):36-61. [doi: 10.1177/1745691617709589] [Medline: 29016274]
4. Wieczorek A, Schrank F, Renner KH, Wagner M. Psychological and physiological health outcomes of virtual reality-based mindfulness interventions: a systematic review and evidence mapping of empirical studies. Digit Health 2024;10:20552076241272604. [doi: 10.1177/20552076241272604] [Medline: 39484656]
5. Cano N, Gómez-Hernández J, Ariza M, et al. A multimodal group-based immersive virtual reality intervention for improving cognition and mental health in patients with post-COVID-19 condition. A quasi-experimental design study. Front Psychol 2024;15:1441018. [doi: 10.3389/fpsyg.2024.1441018] [Medline: 39131859]

XSL•FO
RenderX

6. Zheng L, Li W, Song S, et al. Effectiveness of mindfulness-based virtual reality training on stress, anxiety, and depression among Chinese university students. Mindfulness (N Y) 2024 Apr;15(4):899-913. [doi: 10.1007/s12671-024-02335-4]
7. Ng HYH, Wu CW, Hsu HC, et al. Neurological evidence of diverse self-help breathing training with virtual reality and biofeedback assistance: extensive exploration study of electroencephalography markers. JMIR Form Res 2024 Dec 6;8:e55478. [doi: 10.2196/55478] [Medline: 39642375]
8. Olasz O, Erdős S, Horváth K. The effects of virtual reality-based mindfulness exercises on the perception of time, psychological and physiological states of young people: a randomized crossover trial. Mindfulness (N Y) 2024 Sep;15(9):2347-2354. [doi: 10.1007/s12671-024-02438-y]
9. Ma J, Zhao D, Xu N, Yang J. The effectiveness of immersive virtual reality (VR) based mindfulness training on improvement mental-health in adults: a narrative systematic review. Explore (NY) 2023;19(3):310-318. [doi: 10.1016/j.explore.2022.08.001] [Medline: 36002363]
10. Milasi C, Bonanno M, Bernardelli L, et al. Enhancing mindfulness effects on well-being using immersive virtual reality in non-clinical populations: where are we going? Mindfulness (N Y) 2025 Dec;16(12):3399-3423. [doi: 10.1007/s12671-025-02710-9]
11. Xie F, Zheng H, Hu X. The effectiveness of virtual reality in enhancing mindfulness: a systematic review and meta-analysis. Front Public Health 2025;13:1709782. [doi: 10.3389/fpubh.2025.1709782] [Medline: 41334408]
12. 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] [Medline: 33782057]
13. Ch NAN, Ansah AA, Katrahmani A, et al. Virtual nature experiences and mindfulness practices while working from home during COVID-19: effects on stress, focus, and creativity. Int J Hum Comput Stud 2023 Mar;171:102982. [doi: 10.1016/j.ijhcs.2022.102982] [Medline: 36536937]
14. Hong QN, Pluye P, Fàbregues S, et al. Improving the content validity of the mixed methods appraisal tool: a modified e-Delphi study. J Clin Epidemiol 2019 Jul;111:49-59. [doi: 10.1016/j.jclinepi.2019.03.008] [Medline: 30905698]
15. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. J Educ Stat 1981 Jun;6(2):107-128. [doi: 10.3102/10769986006002107]
16. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis, 2nd edition: Wiley; 2021.
17. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986 Sep;7(3):177-188. [doi: 10.1016/0197-2456(86)90046-2] [Medline: 3802833]
18. Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions Version 63: Cochrane; 2022.
19. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol 2013 Nov 26;4:863. [doi: 10.3389/fpsyg.2013.00863] [Medline: 24324449]
20. Blackmore R, Giles C, Tremain H, et al. Examining the use of virtual reality to support mindfulness skills practice in mood and anxiety disorders: mixed methods study. J Med Internet Res 2024 Dec 6;26:e45640. [doi: 10.2196/45640] [Medline: 39641990]
21. Lee BM, Kim SW, Lee BJ, et al. Effects and safety of virtual reality-based mindfulness in patients with psychosis: a randomized controlled pilot study. Schizophrenia (Heidelb) 2023 Sep 13;9(1):57. [doi: 10.1038/s41537-023-00391-8] [Medline: 37704650]
22. Williams MS, Ordonez K, King D, et al. A brief mindfulness-based intervention pilot study: lessons learned in feasibility and acceptability among healthcare professionals experiencing burnout. OBM Integr Complement Med 2024;09(4):1-21. [doi: 10.21926/obm.icm.2404060]
23. Ferrer Costa J, Moran N, Garcia Marti C, Colmenares Hernandez LJ, Radu Ciorba Ciorba F, Ciudad MJ. Immediate impact of an 8-week virtual reality educational program on burnout and work engagement among health care professionals: pre-post pilot study. JMIR XR Spatial Comput 2024;1:e55678. [doi: 10.2196/55678] [Medline: 38845403]
24. Mao W, Chen W, Wang Y. Effect of virtual reality-based mindfulness training model on anxiety, depression, and cancer-related fatigue in ovarian cancer patients during chemotherapy. Technol Health Care 2024;32(2):1135-1148. [doi: 10.3233/THC-230735] [Medline: 37781832]
25. Franklin DM, Silvestro C, Carrillo RA, et al. The impact of meditation aided by VR technology as an emerging therapeutic to ease cancer related anxiety, stress, and fatigue. Front Virtual Real 2023;4:1195196. [doi: 10.3389/frvir.2023.1195196]
26. Modrego-Alarcón M, Morillo H, Campos D, et al. Effects and acceptability of virtual reality to facilitate mindfulness practice in university students. J Comput High Educ 2025 Mar;37(1):273-298. [doi: 10.1007/s12528-023-09393-y]
27. Cawley A, Tejeiro R. Brief virtual reality mindfulness is more effective than audio mindfulness and colouring in reducing stress in university students. Mindfulness (N Y) 2024 Feb;15(2):272-281. [doi: 10.1007/s12671-024-02306-9]
28. McConnell R, Lane E, Webb G, LaPeze D, Grillo H, Fritz JA. A multicenter feasibility randomized controlled trial using a virtual reality application of pain neuroscience education for adults with chronic low back pain. Ann Med 2024 Dec;56(1):2311846. [doi: 10.1080/07853890.2024.2311846] [Medline: 38354690]
29. Torres García A, Morcillo Serra C, Argilés Huguet M, González Gardó L, Abad Esteve A, Ramos Quiroga JA. Efficacy of a virtual reality intervention for reducing anxiety, depression, and increasing disease coping in patients with breast cancer before their first chemotherapy dose. Cogn Ther Res 2024 Jun;48(3):451-465. [doi: 10.1007/s10608-023-10440-2]

XSL•FO
RenderX

30. Van Doren N, Ng H, Rawat E, McKenna KR, Blonigen DM. Virtual reality mindfulness training for veterans in residential substance use treatment: pilot study of feasibility and acceptability. J Subst Use Addict Treat 2024 Jun;161:209315. [doi: 10.1016/j.josat.2024.209315] [Medline: 38378078]
31. Fonseca A, Qian D, Forbes T, et al. Reducing preoperative caregiver anxiety with virtual reality: a pragmatic, randomized controlled study. J Patient Exp 2024;11:23743735231220190. [doi: 10.1177/23743735231220190] [Medline: 38188533]
32. Garland EL, Recasens M, Taple BJ, Donaldson GW, Weisberg RB. Virtual reality-based Mindfulness-Oriented Recovery Enhancement (MORE-VR) as an adjunct to medications for opioid use disorder: a Phase 1 trial. Ann Med 2024 Dec;56(1):2392870. [doi: 10.1080/07853890.2024.2392870] [Medline: 39172534]
33. Spitz AM, Senda MC, Johnson KA, Taylor IM, Jensen MM, Kozel FA. The relationship of anxious arousal with treatment of dysphoria using virtual reality mindfulness and 2 accelerated transcranial magnetic stimulation protocols. J Clin Psychiatry 2024 May 22;85(2):23m15195. [doi: 10.4088/JCP.23m15195] [Medline: 38780528]
34. Liu N, Deng J, Lu F, Xiao J. Virtual reality enhanced mindfulness and yoga intervention for postpartum depression and anxiety in the post COVID era. Sci Rep 2025;15(1):11766. [doi: 10.1038/s41598-025-96165-6]
35. Barton AC, Do M, Sheen J, Byrne LK. The restorative and state enhancing potential of abstract fractal-like imagery and interactive mindfulness interventions in virtual reality. Virtual Real 2024 Mar;28(1):53. [doi: 10.1007/s10055-023-00916-7]
36. Jimenez-Barragan M, Del Pino Gutierrez A, Sauch Valmaña G, et al. Immersive virtual reality eHealth intervention to reduce anxiety and depression in pregnant women: randomized controlled trial. JMIR Hum Factors 2025 Apr 30;12:e71708. [doi: 10.2196/71708] [Medline: 40306641]
37. Hanna J, Mehta J, West H, Keane P, Wilson N, Bridge P. Evaluation of a novel interactive virtual reality environment for mindfulness skills training. BMC Digit Health 2025;3(1):6. [doi: 10.1186/s44247-024-00141-5]
38. Kim J, Lee J, Kim Y, Ory M, Welsh-Bohmer K. Investigating the impact of immersive virtual reality meditation on coherence achievement score, anxiety, and depression among people living with dementia. J Health Soc Sci 2024;9(1):100-109. [doi: 10.19204/2024/NVST4]
39. Han D, Kim D, Kim K, Cho I. A comparison study on stress relief in VR. Presented at: 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); Mar 25-29, 2023. [doi: 10.1109/VRW58643.2023.00292]
40. Sezer I, Moreau P, El Sayed Hussein Jomaa M, et al. Enhanced brain-heart connectivity as a precursor of reduced state anxiety after therapeutic virtual reality immersion. Adv Sci (Weinh) 2025 Oct;12(38):e03606. [doi: 10.1002/advs.202503606] [Medline: 40788089]
41. Chen Y, Wang T, Tan Y, Li D. Feasibility and effectiveness study of applying a hallucinogen harm reduction and integration model to a mindfulness thinking intervention using virtual reality: a randomized controlled trial. Digit Health 2024;10:20552076241249869. [doi: 10.1177/20552076241249869] [Medline: 38665888]
42. Kamada T, Ohdaira H, Nishie R, et al. Immersive virtual reality for reducing intraoperative pain: a pilot randomized controlled trial. In Vivo 2025;39(3):1638-1646. [doi: 10.21873/invivo.13964] [Medline: 40294990]
43. Kim KY, Hur MH, Kim WJ. Effects of virtual reality (VR)-based meditation on sleep quality, stress, and autonomic nervous system balance in nursing students. Health Care (Don Mills) 2024 Aug 4;12(16):1581. [doi: 10.3390/healthcare12161581] [Medline: 39201140]
44. Kumar A, Sankhyan S, Tripathi K, Thakur S, Bhavsar A, Dutt V. VRZM: exploring the effect of zen meditation on EEG patterns in immersive environments. Presented at: 2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC); Oct 6-10, 2024. [doi: 10.1109/SMC54092.2024.10831479]
45. Mancini M, Chirico A, Niewiadomski R, et al. Multimodal interactive VR mindfulness experience. 2024 Jun 3 Presented at: Proceedings of the 2024 ACM International Conference on Interactive Media Experiences; Jun 3, 2024. [doi: 10.1145/3656650.3656753]
46. Murray G, Shmidheiser M. Virtual reality immerses you in your mind: the experience and stress-reduction benefits of VR mindfulness modules in persons with TBI. Brain Inj 2024 Apr 15;38(5):355-360. [doi: 10.1080/02699052.2024.2311334] [Medline: 38318793]
47. Ong J, Doshi K, Chew KH, Tay JHM, Foo JJC. The effectiveness of immersive VR 360 mindfulness for the reduction of stress and burnout among teachers. MHDT 2025 Mar 17;2(1):20-31. [doi: 10.1108/MHDT-10-2024-0026] [Medline: 40112291]
48. Poetar CR, Bradley N, Voinescu A. Immersive virtual reality or computerised mindfulness meditation for improving mood? Preliminary efficacy from a pilot randomised trial. Front Psychol 2023;14:1157469. [doi: 10.3389/fpsyg.2023.1157469] [Medline: 37954176]
49. Sexton-Radek K, Jensen D, Hardy E, et al. Mindful space: mixed methods study. J Behav Health 2024;13(4). [doi: 10.33425/2832-4579/24072]
50. Purser RE. McMindfulness: How Mindfulness Became the New Capitalist Spirituality: Repeater Books; 2019.
51. Wilson J. Mindful America: The Mutual Transformation of Buddhist Meditation and American Culture: Oxford University Press; 2014.
52. Dobrescu AI, Nussbaumer-Streit B, Klerings I, et al. Restricting evidence syntheses of interventions to English-language publications is a viable methodological shortcut for most medical topics: a systematic review. J Clin Epidemiol 2021 Sep;137:209-217. [doi: 10.1016/j.jclinepi.2021.04.012] [Medline: 33933579]

XSL•FO
RenderX

53. Barker SA. Efficacy of virtual reality–based mindfulness interventions: a systematic review and meta-analysis. Open Science Framework. 2026. URL: https://osf.io/kfjzc/ [accessed 2026-05-29]

Abbreviations

AI: artificial intelligence
EEG: electroencephalography
HR: heart rate
HRV: heart rate variability
MBI: mindfulness-based intervention
MMAT: Mixed Methods Appraisal Tool
OSF: Open Science Framework
PANAS: Positive and Negative Affect Schedule
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
VR: virtual reality
VRbMI: virtual reality–based mindfulness intervention

Edited by I Steenstra; submitted 19.Dec.2025; peer-reviewed by A Arya; revised version received 27.Apr.2026; accepted 28.Apr.2026; published 24.Jun.2026.

Please cite as:
Barker SA, Miles-Novelo AJ, Rizzo A, Tuma RM
Efficacy of Virtual Reality–Based Mindfulness Interventions: Systematic Review and Meta-Analysis
JMIR XR Spatial Comput 2026;3:e90003
URL: https://xr.jmir.org/2026/1/e90003
doi:10.2196/90003

© Sarah Alicia Barker, Andreas J Miles-Novelo, Albert Rizzo, Regina M Tuma. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 24.Jun.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Conceptual Framework for 4D Holographic Visualization in Gastrointestinal Endoscopy: Toward Externalized Spatiotemporal Cognition

Hayato Yamaguchi[1,2], MD, PhD; Masakatsu Fukuzawa[1], MD, PhD; Takashi Kawai[3], MD, PhD; Fumito Yamanishi[1], MD; Takahiro Muramatsu[1], MD, PhD; Tadashi Ichimiya[1], MD; Yasuyuki Kagawa[1], MD, PhD; Kumiko Uchida[1], MD; Akira Madarame[1], MD, PhD; Takashi Morise[1], MD, PhD; Yoshiya Yamauchi[1], MD, PhD; Shin Kono[1], MD, PhD; Sakiko Naito[1], MD, PhD; Takao Itoi[1], MD, PhD

[1]Department of Gastroenterology and Hepatology, Tokyo Medical University, 6-7-1 Nishishinjuku, Shinjuku-ku, Tokyo, Japan

[2]Hayabusa Internal Medicine Home Clinic, Suginami-ku, Tokyo, Japan

[3]Preventive Medicine Center, Tokyo Medical University, Shinjuku-ku, Tokyo, Japan

Corresponding Author:
Hayato Yamaguchi, MD, PhD
Department of Gastroenterology and Hepatology, Tokyo Medical University, 6-7-1 Nishishinjuku, Shinjuku-ku, Tokyo, Japan

Abstract

This viewpoint explores the concept of 4D holographic visualization in gastrointestinal endoscopy and discusses its potential implications from a medical informatics perspective. Current endoscopic practice relies primarily on 2D image displays, requiring operators to mentally reconstruct 3D spatial relationships and temporal changes in a dynamic and deformable anatomical environment. The aim of this viewpoint is to present a conceptual framework that reframes endoscopic visualization as a spatiotemporal information integration problem rather than solely a display problem. We discuss how future visualization systems may integrate spatial and temporal information to support externalized representations of anatomy, endoscope configuration, and procedural dynamics in a shared clinical environment. Potential applications include endoscopic navigation, training, and collaborative procedures. We further discuss key technical and clinical considerations, including information integration, registration stability, real-time responsiveness, workflow compatibility, and uncertainty associated with reconstructed information. This viewpoint highlights spatiotemporal information integration as a central challenge in the development of next-generation visualization systems for gastrointestinal endoscopy.

(JMIR XR Spatial Comput 2026;3:e99446) doi:10.2196/99446

KEYWORDS

gastrointestinal endoscopy; medical informatics; holography; 4D imaging; human-machine interface; spatial cognition

Introduction

Background

Gastrointestinal (GI) endoscopy is essential for the diagnosis and treatment of digestive diseases, with continuous advances in imaging technologies over recent decades [1-3]. Despite these improvements, visualization remains predominantly 2D, requiring clinicians to infer 3D spatial relationships and temporal dynamics from sequential images [4-6].

This limitation is particularly problematic in GI endoscopy due to the inherently dynamic and deformable nature of the GI tract, characterized by peristalsis and respiration and affected by endoscope manipulation. As a result, operators rely heavily on implicit mental reconstruction, increasing cognitive load and contributing to variability in procedural performance, particularly among novice endoscopists [7-9].

These constraints highlight a fundamental mismatch between the spatiotemporal complexity of the target anatomy and the format in which information is presented.

This work reframes endoscopic visualization as an information integration problem rather than a display enhancement problem, emphasizing the need to align visualization paradigms with the spatiotemporal complexity of the target anatomy and its implications for clinical performance and decision-making. The aim of this viewpoint is to present a conceptual framework for 4D holographic visualization in GI endoscopy. Rather than proposing a specific device or implementation pathway, we explore how future visualization systems may integrate and externalize spatial and temporal information to support navigation, training, and collaborative clinical practice. We further discuss the technical and clinical considerations that may influence the development and translation of such systems. This viewpoint is intended for endoscopists, endoscopy trainees,

XSL•FO
RenderX

medical informatics researchers, and developers of next-generation visualization systems.

Why Current Visualization Fails in GI Endoscopy

These limitations are summarized in Table 1, which outlines the key constraints of conventional visualization and their clinical implications. The limitations of conventional visualization in GI endoscopy are fundamentally architectural rather than purely technical. Current systems present information in a format that does not align with the dynamic nature of the target anatomy.

Table . Limitations of conventional endoscopic visualization and potential contributions of 4D holography.

Limitation of conventional visualization	Clinical implication of conventional visualization	Potential contribution of 4D holography to remove limitation
2D display	High cognitive load	Externalized spatial representation
Lack of depth information	Misinterpretation of lesion depth	Phase-informed or computational depth estimation
Discontinuous temporal understanding	Difficulty tracking deformation	Time-continuous visualization
Implicit scope shape estimation	Loop formation misjudgment	Dynamic spatial representation
Individual mental reconstruction	Variability in skill acquisition	Shared spatial awareness and educational scalability

Key limitations include lack of explicit depth representation, discontinuous temporal perception, dependence on mental reconstruction by the operator, and limited understanding of global scope configuration, such as loop formation [4-6].

Existing approaches—including 3D reconstruction, navigation systems, and mixed-reality (MR) visualization—partially address these limitations. However, most approaches rely on indirect estimation or device-centered visualization and do not fully externalize dynamic spatial information into shared clinical space [5,10-13].

4D Holographic Visualization

Concept of 4D Holographic Visualization

The term “holographic” is used conceptually to describe depth-coherent and phase-consistent spatiotemporal representation rather than a strict physical implementation of holography. We define 4D holographic visualization as a system-level paradigm that integrates depth-resolved spatial structure and continuous temporal dynamics into shared real space.

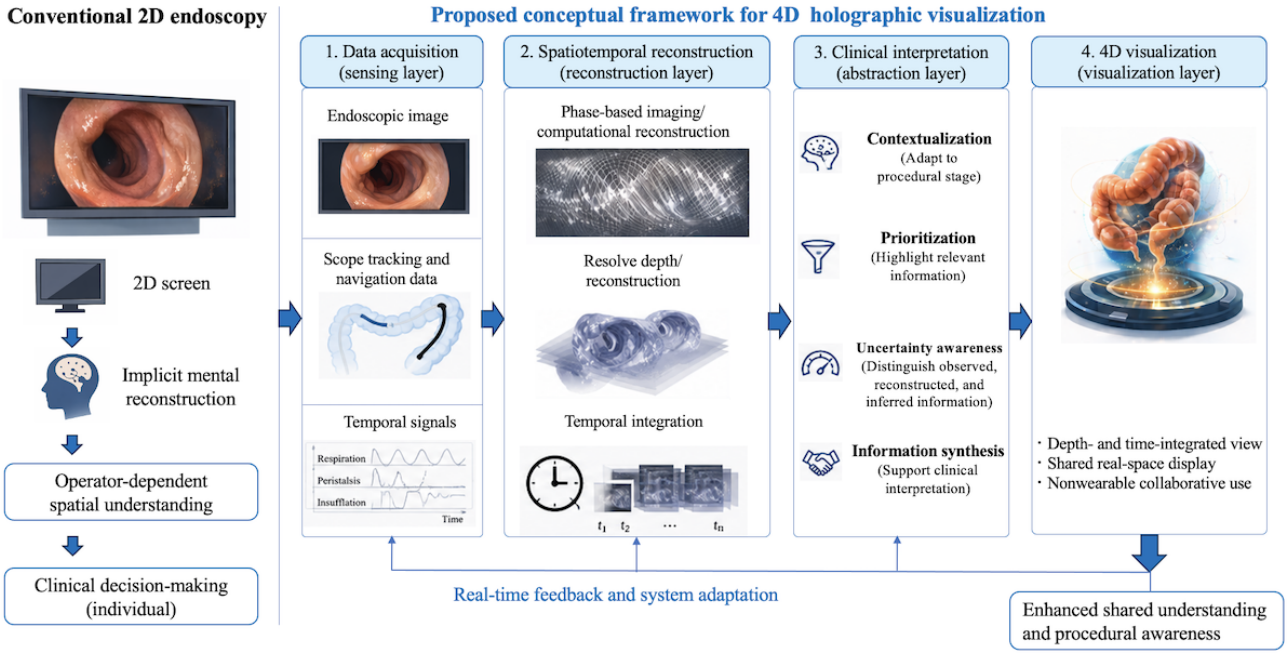
Digital holography enables the acquisition of both amplitude and phase information, forming the basis of quantitative phase imaging [10-12,14-17]. While primarily applied in microscopic imaging, phase-based approaches provide a conceptual foundation for representing dynamic, deformable biological structures.

This paradigm represents a shift from implicit mental reconstruction to explicit system-supported visualization, effectively externalizing spatial cognition in the clinical environment. This conceptual shift toward externalized spatial cognition is illustrated in Figure 1.

Conventional endoscopy relies on operator-dependent mental reconstruction of spatial and temporal relationships from 2D images. The proposed framework conceptualizes 4D visualization as a 4-layer information integration system consisting of sensing, spatiotemporal reconstruction, clinical interpretation, and visualization layers. The clinical interpretation layer emphasizes contextualization, prioritization, uncertainty awareness, and information synthesis to support meaningful integration of complex spatiotemporal information. This illustration is conceptual and does not represent a specific device or implementation.

XSL•FO
RenderX

Figure 1. Conceptual shift from mental reconstruction to integrated spatiotemporal visualization in gastrointestinal endoscopy.



A 4-Layer Information Integration Framework

4D visualization should be understood as an information integration system composed of 4 functional layers: sensing, spatiotemporal reconstruction, clinical interpretation, and visualization. From a medical informatics perspective, these layers define how spatial and temporal information is acquired, reconstructed, interpreted, and externalized for clinical use.

The sensing layer acquires information from multiple sources, including endoscopic images, scope-tracking systems, and temporal signals associated with procedural dynamics. The reconstruction layer transforms these inputs into depth-resolved and temporally coherent representations of anatomy and instrument configuration.

Importantly, reconstructed information alone may not be clinically useful. A clinical interpretation layer is therefore required between reconstruction and visualization. This layer serves to contextualize information according to procedural circumstances, prioritize clinically relevant features, and explicitly distinguish directly measured information from inferred or reconstructed information. Information originating from different sources may carry different levels of reliability and uncertainty, which should be communicated appropriately to users. Such processing may help mitigate information overload and improve the interpretability of complex spatiotemporal data. Rather than displaying all available spatiotemporal information, the purpose of this layer is to transform complex data into clinically actionable representations.

The visualization layer externalizes interpreted information into a shared spatial representation. In this framework, visualization is viewed as the final stage of information integration rather than merely a display modality. The potential clinical value of the system depends not only on the visualization technology itself but also on the effective integration of all preceding layers.

Failures in latency, registration, information interpretation, or workflow compatibility may compromise usability and safety regardless of imaging performance.

From a system-level and medical informatics perspective, representative visualization paradigms differ not only in display modality but also in how spatial and temporal information is integrated and externalized for clinical use. A conceptual comparison is presented in Figure 2, with corresponding system-level characteristics summarized in Table 2.

Representative paradigms are compared based on hypothesized system-level characteristics. These descriptions reflect conceptual tendencies rather than validated performance metrics and are not intended to indicate superiority. These comparisons are conceptual and intended to illustrate system-level differences rather than quantitative performance.

XSL•FO
RenderX

Figure 2. Conceptual comparison of visualization paradigms in gastrointestinal endoscopy. AR: augmented reality; HMD: head-mounted display; MR: mixed reality.

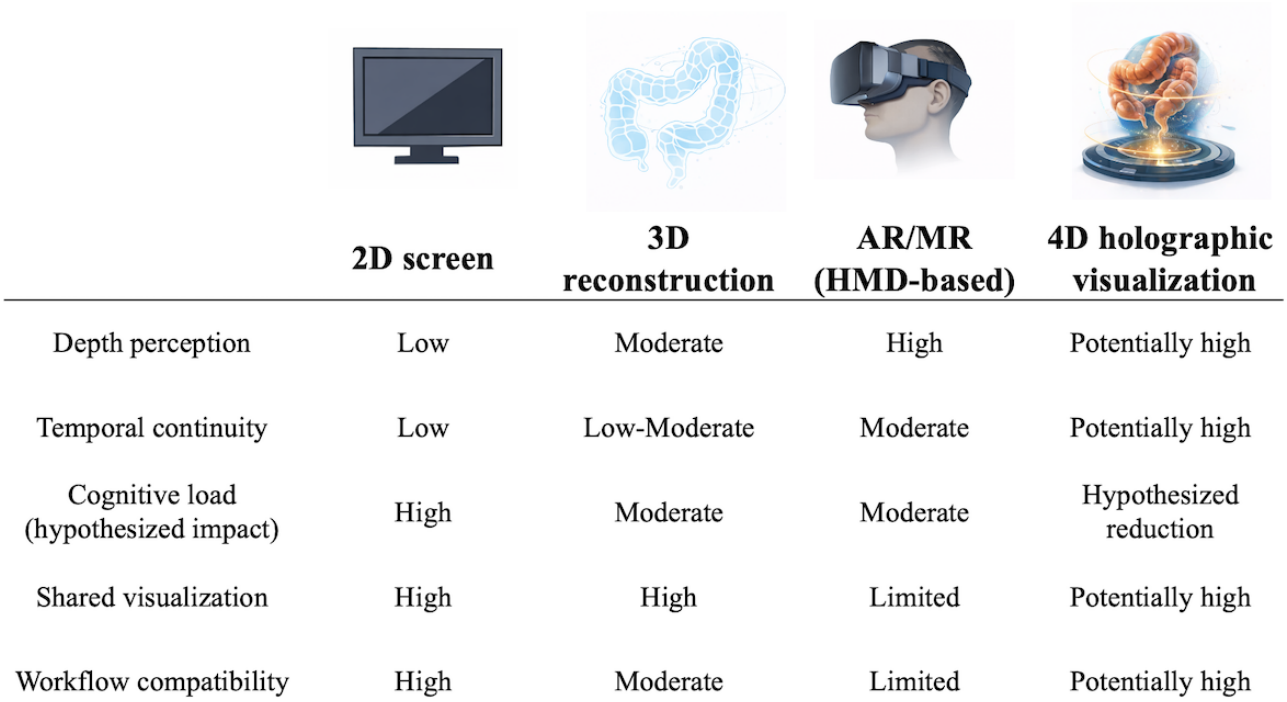


Table . Qualitative comparison of system-level characteristics[a] of visualization paradigms in gastrointestinal endoscopy.

Feature	2D display	3D reconstruction	AR[b]/MR[c] (HMD[d])	4D holography
Spatial depth representation	Implicit, mentally inferred	Partial computational reconstruction	Explicit stereoscopic depth	Phase-resolved spatial representation
Temporal deformation handling	Discrete frame–based	Limited continuity	Moderate, latency dependent	Potential for time-continuous integration
Cognitive load	High	Moderate	Moderate	Hypothesized reduction through spatial externalization
Shared visualization	High	High	Limited (device centered)	High (real-space, nonwearable sharing)
Workflow compatibility	High	Moderate	Limited	Potentially high (workflow dependent)
Primary limitation	No explicit depth/time	Incomplete deformation handling	Workflow and sterility constraints	Real-time implementation challenges

[a]All characteristics represent conceptual hypotheses derived from literature synthesis rather than empirically validated performance metrics.

[b]AR: augmented reality.

[c]MR: mixed reality.

[d]HMD: head-mounted display.

Clinical Relevance and Use Cases of 4D Visualization

Loop Recognition and Navigation

Loop formation during colonoscopy represents a major challenge in spatial understanding. 4D visualization may potentially support real-time recognition of scope configuration, improving procedural efficiency and reducing insertion-related difficulty [18-20]. Previous studies of electromagnetic scope imaging have demonstrated improved recognition of scope configuration and enhanced training performance in selected colonoscopy settings, supporting the potential value of externalized spatial information during endoscopic procedures [18-20].

Support for Novice Endoscopists

By externalizing spatial relationships, 4D visualization may reduce the need for continuous mental reconstruction of anatomy and scope configuration. However, potential benefits are likely to depend on how information is contextualized, prioritized, and integrated into clinically meaningful representations rather than on information quantity alone. Cognitive load has been shown to correlate with procedural performance in endoscopy and simulation-based environments [12,21-24]. However,

XSL•FO
RenderX

increasing the availability of spatial and temporal information does not necessarily reduce cognitive burden. Excessive or poorly organized information may instead increase information overload, distract operators, and impair situational awareness. The potential benefits of 4D visualization are likely to depend not only on information availability but also on effective contextualization, prioritization, and interpretation of information. Therefore, the proposed framework does not assume that more information inherently improves performance. Rather, its potential value lies in transforming complex spatiotemporal data into clinically interpretable information through appropriate abstraction and contextualization.

Shared Spatial Understanding

Unlike head-mounted display–based systems, real-space visualization enables shared perception among multiple operators. This may potentially enhance communication, coordination, and instructional efficiency in team-based procedures [25-28]. Studies from medical informatics and team-based clinical environments have suggested that integrated information displays may improve shared situational awareness and communication efficiency [26].

Implications for Medical Informatics

This framework positions 4D holographic visualization as a medical informatics problem rather than solely an imaging innovation. It involves integration of spatial and temporal data, optimization of human-machine interaction, and alignment with clinical workflows. This paradigm shifts visualization from displaying images to structuring clinical information.

Operationalizing the Clinical Interpretation Layer

The proposed framework becomes clinically meaningful only when reconstructed spatiotemporal information is transformed into clinically interpretable representations that support real-time clinical decision-making. While this viewpoint does not prescribe a specific implementation strategy, we propose several design principles to guide the future development of clinically interpretable 4D visualization systems.

Information presentation should be dynamically adapted to the procedural stage rather than remaining static throughout the examination. For example, during colonoscope insertion, a simplified navigation view may display only luminal direction, scope configuration, and loop formation while temporarily suppressing lesion detection overlays or other secondary information. During withdrawal, the visualization would instead prioritize lesion candidates, blind-spot awareness, and inspection completeness while reducing navigational overlays. Such context-dependent adaptation illustrates how the interpretation layer filters, prioritizes, and simplifies reconstructed information according to the operator's immediate clinical objective rather than displaying all available information simultaneously.

In addition, reconstructed information should remain distinguishable according to its origin. Future systems may integrate directly observed endoscopic images with scope-tracking data, computational reconstruction, preprocedural imaging, and AI-derived predictions. Because these information sources differ in evidential certainty, their provenance and associated uncertainty should be communicated using intuitive visual strategies, such as transparency, confidence indicators, or graphical annotations, enabling clinicians to appropriately interpret reconstructed information.

These design principles position the clinical interpretation layer as the bridge between spatiotemporal reconstruction and clinically actionable visualization, providing an operational and conceptual foundation for future 4D visualization systems.

Future Directions

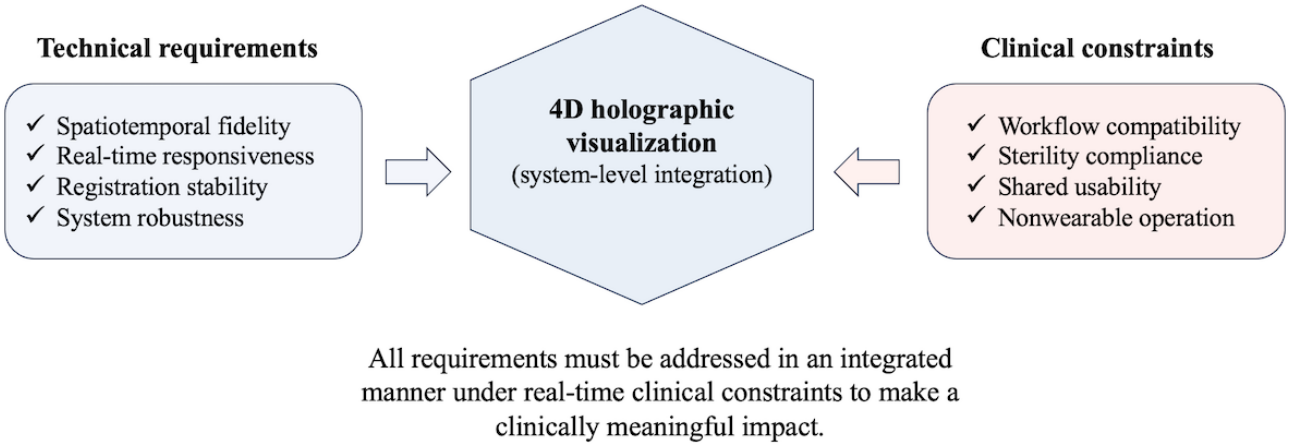
Several challenges must be addressed before clinical translation, including real-time processing, registration stability, robustness to noise, workflow compatibility, and infection control constraints. Registration represents a particularly important challenge in GI endoscopy because the anatomy is highly dynamic and deformable. Differences in luminal distension, peristalsis, respiration, patient positioning, and endoscope manipulation may result in discrepancies between reconstructed models and the actual anatomical state. Consequently, registration errors may affect the accuracy, interpretability, and clinical utility of integrated spatiotemporal representations. Future systems may therefore require uncertainty-aware registration strategies and mechanisms for communicating registration confidence to users. These challenges may be particularly pronounced when integrating endoscopic data with preprocedural imaging modalities such as computed tomography (CT) or CT colonography, where anatomical conditions may differ substantially from those encountered during real-time endoscopy. Future systems may also require mechanisms for representing source-dependent uncertainty and confidence levels in integrated spatiotemporal models. Future systems may incorporate external imaging modalities or additional sensing technologies, but information derived from these sources should remain explicitly distinguishable from directly observed endoscopic findings and should be presented with appropriate confidence and uncertainty information. Integration with AI-based perception systems may further enhance system performance and usability [29-31]. Beyond visualization alone, future development of 4D information integration systems may facilitate interaction with emerging bioelectronic and intelligent therapeutic platforms. Recent advances in bioinspired electronic skins have demonstrated the integration of multimodal sensing, adaptive interfaces, AI-assisted signal interpretation, and human-machine interaction in a single platform [32]. Such developments suggest potential opportunities for future convergence between spatiotemporal visualization systems and next-generation bioelectronic or therapeutic technologies. Although these applications remain speculative, interdisciplinary integration may represent an important direction for future research. The interdependent technical requirements and clinical constraints necessary for implementation are summarized in Figure 3. These components should be addressed in an integrated manner to achieve clinically meaningful impact. Clinically meaningful impact emerges only when technical requirements and clinical constraints are satisfied simultaneously at the system level.

Future empirical studies are required to validate the proposed framework and quantify its impact on clinical performance and

XSL•FO
RenderX

workflow integration. Potential evaluation metrics may include procedural time, error rates, loop formation frequency, cognitive workload (eg, NASA Task Load Index), and training efficiency among novice endoscopists. Importantly, elements of the proposed framework are not entirely novel in isolation. Components such as real-time 3D reconstruction, electromagnetic scope tracking, and MR visualization have already been explored in GI endoscopy. The present framework should therefore be understood as an integration of these existing technological streams into a unified system-level paradigm rather than a completely new standalone technology. From a translational perspective, initial implementation may focus on hybrid systems that combine existing tracking and reconstruction technologies with simplified real-space visualization before achieving fully integrated 4D holographic systems. Such stepwise development may facilitate feasibility testing and clinical adoption. This framework is conceptual and hypothesis generating, and its clinical effectiveness remains to be demonstrated. Potential limitations include system complexity, computational demands, and dependency on stable real-time integration.

Figure 3. System-level requirements for 4D holographic visualization in gastrointestinal endoscopy.



Conclusions

4D holographic visualization represents a system-level reconfiguration of how spatial and temporal information is integrated and presented in GI endoscopy. By externalizing spatiotemporal information into shared real space, this paradigm shifts spatial cognition from the individual clinician to the system level, with implications for procedural performance, training, and collaborative clinical practice. The novelty of this framework lies not in individual technologies but in the explicit system-level integration of spatiotemporal information into shared real-space visualization, which has not been formally conceptualized in GI endoscopy. Although this framework remains conceptual and hypothesis generating, it provides a foundation for future development and evaluation of integrated visualization systems in endoscopy.

Acknowledgments

The authors would like to thank their colleagues and the staff at the Department of Gastroenterology for their valuable assistance and support in conducting this study.

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS KAKENHI grant number JP25K18986).

Authors' Contributions

HY conceived and designed the study, supervised the project, and critically revised the manuscript. HY and MF performed the literature search and analysis. MF, TK, FY, T Ichimiya, T Muramatsu, YK, KU, AM, T Morise, YY, SK, SN, and T Itoi contributed to interpretation of the literature and provided critical intellectual input. All authors drafted or revised the manuscript and approved the final version.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Săftoiu A, Hassan C, Areia M, et al. Role of gastrointestinal endoscopy in the screening of digestive tract cancers in Europe: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy 2020 Apr;52(4):293-304. [doi: 10.1055/a-1104-5245] [Medline: 32052404]
2. Min JK, Kwak MS, Cha JM. Overview of deep learning in gastrointestinal endoscopy. Gut Liver 2019 Jan 11;13(4):388-393. [doi: 10.5009/gnl18384] [Medline: 30630221]
3. Matsuda T, Ono A, Sekiguchi M, Fujii T, Saito Y. Advances in image enhancement in colonoscopy for detection of adenomas. Nat Rev Gastroenterol Hepatol 2017 May;14(5):305-314. [doi: 10.1038/nrgastro.2017.18] [Medline: 28293023]
4. Shah SG, Saunders BP, Brooker JC, Williams CB. Magnetic imaging of colonoscopy: an audit of looping, accuracy and ancillary maneuvers. Gastrointest Endosc 2000 Jul;52(1):1-8. [doi: 10.1067/mge.2000.107296] [Medline: 10882954]
5. Valdastri P, Simi M, Webster RJ III. Advanced technologies for gastrointestinal endoscopy. Annu Rev Biomed Eng 2012;14(1):397-429. [doi: 10.1146/annurev-bioeng-071811-150006] [Medline: 22655598]
6. East JE, Vleugels JL, Roelandt P, et al. Advanced endoscopic imaging: European Society of Gastrointestinal Endoscopy (ESGE) technology review. Endoscopy 2016 Nov;48(11):1029-1045. [doi: 10.1055/s-0042-118087] [Medline: 27711949]
7. Finocchiaro M, Cortegoso Valdivia P, Hernansanz A, et al. Training simulators for gastrointestinal endoscopy: current and future perspectives. Cancers (Basel) 2021 Mar 20;13(6):1427. [doi: 10.3390/cancers13061427] [Medline: 33804773]
8. Li H, Hou X, Lin R, et al. Advanced endoscopic methods in gastrointestinal diseases: a systematic review. Quant Imaging Med Surg 2019 May;9(5):905-920. [doi: 10.21037/qims.2019.05.16] [Medline: 31281783]
9. Luo X, Mori K, Peters TM. Advanced endoscopic navigation: surgical big data, methodology, and applications. Annu Rev Biomed Eng 2018 Jun 4;20(1):221-251. [doi: 10.1146/annurev-bioeng-062117-120917] [Medline: 29505729]
10. Gabor D. A new microscopic principle. Nature 1948 May 15;161(4098):777. [doi: 10.1038/161777a0] [Medline: 18860291]
11. Lee K, Kim K, Jung J, et al. Quantitative phase imaging techniques for the study of cell pathophysiology: from principles to applications. Sensors (Basel) 2013 Mar 28;13(4):4170-4191. [doi: 10.3390/s130404170] [Medline: 23539026]
12. Khan R, Plahouras J, Johnston BC, Scaffidi MA, Grover SC, Walsh CM. Virtual reality simulation training in endoscopy: a Cochrane review and meta-analysis. Endoscopy 2019 Jul;51(7):653-664. [doi: 10.1055/a-0894-4400] [Medline: 31071757]
13. Shah SG, Brooker JC, Williams CB, Thapar C, Saunders BP. Effect of magnetic endoscope imaging on colonoscopy performance: a randomised controlled trial. Lancet 2000 Nov 18;356(9243):1718-1722. [doi: 10.1016/S0140-6736(00)03205-0] [Medline: 11095259]
14. Park Y, Depeursinge C, Popescu G. Quantitative phase imaging in biomedicine. Nature Photon 2018 Oct;12(10):578-589. [doi: 10.1038/s41566-018-0253-x]
15. Balasubramani V, Kujawińska M, Allier C, et al. Roadmap on digital holography-based quantitative phase imaging. J Imaging 2021 Nov 26;7(12):252. [doi: 10.3390/jimaging7120252] [Medline: 34940719]
16. Nguyen TL, Pradeep S, Judson-Torres RL, Reed J, Teitell MA, Zangle TA. Quantitative phase imaging: recent advances and expanding potential in biomedicine. ACS Nano 2022 Aug 23;16(8):11516-11544. [doi: 10.1021/acsnano.1c11507] [Medline: 35916417]
17. Huang Z, Cao L. Quantitative phase imaging based on holography: trends and new perspectives. Light Sci Appl 2024 Jun 27;13(1):145. [doi: 10.1038/s41377-024-01453-x] [Medline: 38937443]
18. Peter S, Reddy NB, Naseemuddin M, Zaibaq JN, McGwin G, Wilcox CM. Outcomes of use of electromagnetic guidance with responsive insertion technology (RIT) during colonoscopy: a prospective randomized controlled trial. Endosc Int Open 2019 Feb;7(2):E225-E231. [doi: 10.1055/a-0754-1879] [Medline: 30705957]
19. Tan X, Yang W, Wichmann D, et al. Magnetic endoscopic imaging as a rational investment for specific colonoscopies: a systematic review and meta-analysis. Expert Rev Gastroenterol Hepatol 2021 Apr 3;15(4):447-458. [doi: 10.1080/17474124.2021.1842192]
20. Ma HM, Gao LP, Wang PF, et al. Efficacy of ScopeGuide-assisted training in enhancing colonoscopy competence and reducing patient discomfort. Surg Laparosc Endosc Percutan Tech 2024 Apr 1;34(2):136-142. [doi: 10.1097/SLE.0000000000001236] [Medline: 38462904]
21. Sewell JL, Bowen JL, Cate OT, O'Sullivan PS, Shah B, Boscardin CK. Learning challenges, teaching strategies, and cognitive load: insights from the experience of seasoned endoscopy teachers. Acad Med 2020 May;95(5):794-802. [doi: 10.1097/ACM.0000000000002946] [Medline: 31425188]
22. Shahidi N, Ou G, Telford J, Enns R. Establishing the learning curve for achieving competency in performing colonoscopy: a systematic review. Gastrointest Endosc 2014 Sep;80(3):410-416. [doi: 10.1016/j.gie.2014.04.056] [Medline: 24973174]
23. Yurko YY, Scerbo MW, Prabhu AS, Acker CE, Stefanidis D. Higher mental workload is associated with poorer laparoscopic performance as measured by the NASA-TLX tool. Simul Healthc 2010 Oct;5(5):267-271. [doi: 10.1097/SIH.0b013e3181e3f329] [Medline: 21330808]
24. Dias RD, Ngo-Howard MC, Boskovski MT, Zenati MA, Yule SJ. Systematic review of measurement tools to assess surgeons' intraoperative cognitive workload. Br J Surg 2018 Apr;105(5):491-501. [doi: 10.1002/bjs.10795] [Medline: 29465749]
25. Ekkelenkamp VE, Koch AD, de Man RA, Kuipers EJ. Training and competence assessment in GI endoscopy: a systematic review. Gut 2016 Apr;65(4):607-615. [doi: 10.1136/gutjnl-2014-307173] [Medline: 25636697]

XSL•FO
RenderX

26. Koch SH, Weir C, Westenskow D, et al. Evaluation of the effect of information integration in displays for ICU nurses on situation awareness and task completion time: A prospective randomized controlled study. Int J Med Inform 2013 Aug;82(8):665-675. [doi: 10.1016/j.ijmedinf.2012.10.002] [Medline: 23357614]
27. Walsh CM, Ling SC, Khanna N, et al. Gastrointestinal Endoscopy Competency Assessment Tool: development of a procedure-specific assessment tool for colonoscopy. Gastrointest Endosc 2014 May;79(5):798-807. [doi: 10.1016/j.gie.2013.10.035] [Medline: 24321390]
28. Singh S, Sedlack RE, Cook DA. Effects of simulation-based training in gastrointestinal endoscopy: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2014 Oct;12(10):1611-1623. [doi: 10.1016/j.cgh.2014.01.037] [Medline: 24509241]
29. Metzger R, Suppa P, Li Z, Vemuri A. Augmented reality navigation systems in endoscopy. Front Gastroenterol 2024;3:1345466. [doi: 10.3389/fgstr.2024.1345466]
30. Mori Y, Kudo SE, Mohmed HEN, et al. Artificial intelligence and upper gastrointestinal endoscopy: current status and future perspective. Dig Endosc 2019 Jul;31(4):378-388. [doi: 10.1111/den.13317] [Medline: 30549317]
31. Chavaz L, Davidovic A, Meling TR, et al. Evaluation of the precision of navigation-assisted endoscopy according to the navigation tool setup and the type of endoscopes. Acta Neurochir (Wien) 2022 Sep;164(9):2375-2383. [doi: 10.1007/s00701-022-05276-w] [Medline: 35764694]
32. Li W, Li J, Ding X, et al. Multi-bioinspired electronic skins with on-demand adhesion and opto-electronic synergistic display capabilities. Innovation (Camb) 2025 May 5;6(5):100877. [doi: 10.1016/j.xinn.2025.100877] [Medline: 40432782]

Abbreviations

AR: augmented reality
CT: computed tomography
GI: gastrointestinal
HCI: human-computer interaction
HMI: human-machine interface
MR: mixed reality

Edited by I Steenstra; submitted 25.Apr.2026; peer-reviewed by T Kanno, W Li; revised version received 04.Aug.2026; accepted 04.Aug.2026; published 27.Aug.2026.

Please cite as:
Yamaguchi H, Fukuzawa M, Kawai T, Yamanishi F, Muramatsu T, Ichimiya T, Kagawa Y, Uchida K, Madarame A, Morise T, Yamauchi Y, Kono S, Naito S, Itoi T
A Conceptual Framework for 4D Holographic Visualization in Gastrointestinal Endoscopy: Toward Externalized Spatiotemporal Cognition
JMIR XR Spatial Comput 2026;3:e99446
URL: https://xr.jmir.org/2026/1/e99446
doi:10.2196/99446

© Hayato Yamaguchi, Masakatsu Fukuzawa, Takashi Kawai, Fumito Yamanishi, Takahiro Muramatsu, Tadashi Ichimiya, Yasuyuki Kagawa, Kumiko Uchida, Akira Madarame, Takashi Morise, Yoshiya Yamauchi, Shin Kono, Sakiko Naito, Takao Itoi. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 27.Aug.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploratory Patterns of Task Difficulty and Visual Feedback in Virtual Reality–Based Upper-Limb Trajectory Training

Yu Shi[1,2], MS; Sophie Dewil[1,2], MS; Zachary Marvin[1,2], MS; Noam Y Harel[3,4,5], MD, PhD; Raviraj Nataraj[1,2], PhD

[1]Department of Biomedical Engineering, Stevens Institute of Technology, 1 Castle Point Terrace, Hoboken, NJ, United States
[2]Movement Control Rehabilitation (MOCORE) Laboratory, Altorfer Complex, Stevens Institute of Technology, Hoboken, NJ, United States
[3]Spinal Cord Damage Research Center, James J. Peters Department of VA Medical Center, New York, NY, United States
[4]Department of Rehabilitation and Human Performance, Icahn School of Medicine at Mount Sinai, New York, NY, United States
[5]Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, United States

Corresponding Author:
Raviraj Nataraj, PhD
Department of Biomedical Engineering, Stevens Institute of Technology, 1 Castle Point Terrace, Hoboken, NJ, United States

Abstract

In this pilot evaluation of a virtual reality platform for upper-extremity motor rehabilitation, 9 neurotypical participants performed a force-modulated tracking task across 4 trajectory difficulty levels with and without augmented visual feedback; tracking accuracy improved in all conditions, with the descriptive visual-feedback advantage more pronounced at higher difficulty.

(JMIR XR Spatial Comput 2026;3:e94745) doi:10.2196/94745

KEYWORDS

virtual reality training; augmented sensory feedback; task difficulty; motor learning; upper extremity function

Introduction

Virtual reality (VR)–based motor training provides a controlled and interactive environment from which to assign training task elements that shape movement practice and motor performance [1]. Such functionality with VR platforms is especially beneficial in designing protocols for motor rehabilitation [2]. This pilot study examined changes in force-based motor performance across two adjustable elements in VR-based training: augmented sensory feedback (ASF) and task difficulty. ASF provides additional real-time information about movement performance to support improved error correction [3]. Visual ASF is well-suited to immersive VR environments and can produce significant performance gains in VR task platforms [3]. The value of augmented feedback may increase with task demands, as greater path curvature and arc length raise spatial control requirements and reliance on external visual guidance [4]. As such, task difficulty is another essential training element, readily controllable in VR, and is central to ensuring skill acquisition in relation to user capability [5].

This pilot study presents a new version of a previously described VR training platform [6] that has participants exert muscular effort with the upper extremity to command a virtual avatar to trace target trajectories. A trajectory-based approach requires spatial accuracy while preserving user autonomy during continuous upper-limb VR training [7]. Further, this study newly explores target-tracking accuracy as a function of increasing training difficulty levels (ie, larger trajectory amplitudes) with and without augmented visual feedback. This pilot study was conducted with a small cohort of neurotypicals to demonstrate initial feasibility; however, the long-term application of this platform is VR-based motor rehabilitation of upper-extremity function after neurological injury [1].

Methods

Ethical Considerations

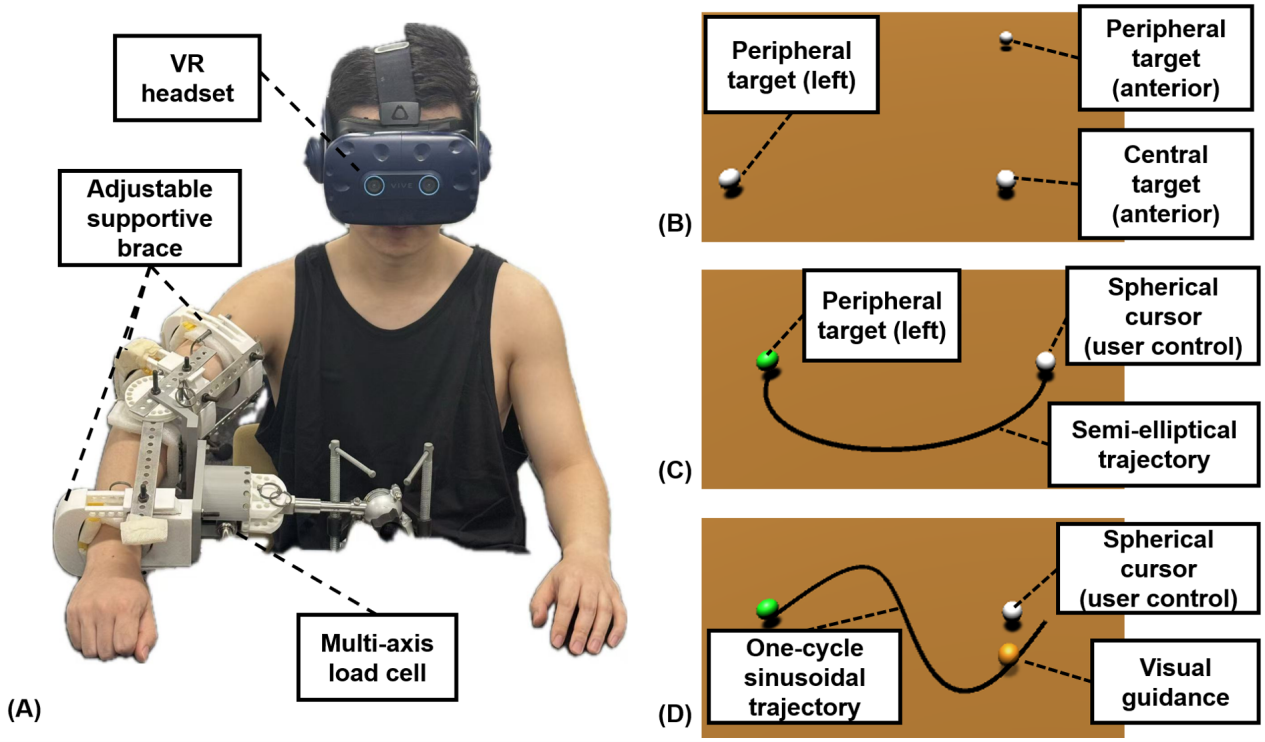
This study was approved by the institutional review board of Stevens Institute of Technology (protocol number 2021 - 036). All participants provided written informed consent before participation.

Participants and Experimental Design

Nine neurotypical, right-side–dominant participants (5 males and 4 females; age range from 18 to 22 years) were enrolled in this study. The sample size reflected an exploratory repeated-measures pilot design rather than powered confirmatory testing. Participants were seated and placed their right arm in a position-adjustable brace used with a previously developed VR platform for training upper-limb function [6]. The virtual task environment was developed in Unity (Unity Software Inc) and displayed through an HTC Vive headset (HTC Corporation). A major adaptation of the prior platform was the addition of a tri-axial load cell (Fibos FA703) mounted to the brace to record 3D quasi-isometric forces. Only anterior-posterior (x) and medial-lateral (y) force components controlled the position of a spherical cursor in the x-y plane of the virtual environment (Figure 1A).

XSL•FO
RenderX

Figure 1. Experimental setup and VR task environment: (A) experimental general setup, (B) spatial arrangement of targets in VR space, (C) VR task environment in pre-/post-training session, (D) VR task environment in training session (with visual augmented sensory feedback and level 2 sinusoid). VR: virtual reality.



The task space contained one central target and two peripheral targets, which were contacted by the cursor after participants followed prescribed trajectories. Distance values were arbitrary Unity task-space units and did not correspond to real-world distances. Peripheral targets were 0.5 units in diameter and positioned 7 units from the central target (Figure 1B). In each trial, participants guided the cursor along the displayed trajectory to contact the active target; after a 1-second pause, the next target and trajectory appeared. Each peripheral target was activated twice in randomized order, followed by a return-to-center movement. Each condition included two pretraining, three training, and two posttraining trials. Pre/post trials used a semielliptical target trajectory (Figure 1C), whereas training trials used a one-cycle sinusoidal trajectory (Figure 1D), allowing assessment of generalization to an untrained but related movement path [8].

Participants completed 8 randomized training conditions (2 ASF conditions × 4 difficulty levels) in a single session, with 5-minute rests between conditions. Each condition took approximately 10 minutes, and no participant reported postsession fatigue. The ASF conditions were no ASF (control case, NF) and visual-only ASF (VF). Visual ASF was presented as a semitransparent orange sphere at the desired cursor location, defined as the projection of the current cursor position onto the prescribed trajectory (Figure 1D). Its position and opacity were continuously updated during movement; it appeared when tracking error exceeded 0.5 units, with opacity increasing linearly to fully opaque at 2.0 units. Difficulty levels L1-L4 used sinusoidal amplitudes of 1, 2, 3, and 4 Unity units, respectively. Motor performance (accuracy) was assessed according to tracking error, computed as the nonnormalized mean Euclidean distance between actual and desired cursor positions across sampled frames. Before calculating pretraining and posttraining mean tracking errors, contact-level tracking-error values were screened separately within each experimental condition using the rmoutliers function in MATLAB R2023b (MathWorks), with values more than 2 SDs from the corresponding condition mean removed. This screening was applied once, at the contact level; all participant-level values, VF – NF differences, and model inputs derive from these screened data. The effect of training on tracking accuracy was calculated as [(pretraining error − posttraining error) / pretraining error] × 100, representing the percentage change in performance from before to after training. As exploratory analysis, the difference in performance was evaluated between VF and NF conditions pooled across all difficulty levels. Further, an exploratory mixed effects model was fitted to examine if participant-level VF – NF differences were a function of difficulty level as a continuous fixed effect and participant modeled as a random intercept.

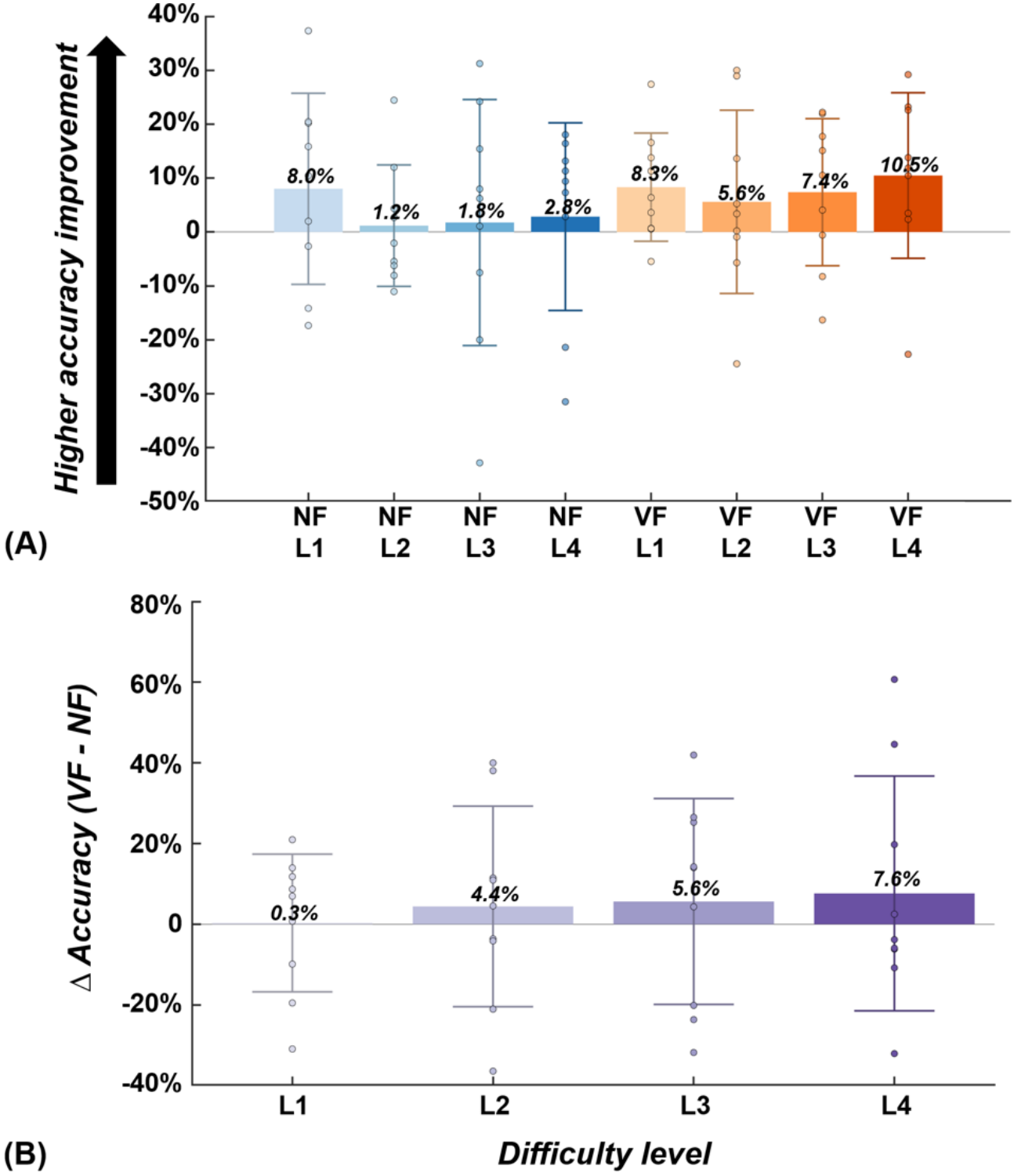
Results

Net percentage change in tracking accuracy was positive for all 8 conditions (Figure 2A). Under NF, the largest mean improvement occurred at the lowest difficulty level, L1, whereas improvements at higher difficulty levels were smaller. In contrast, VF produced positive improvements across all difficulty levels, ranging from approximately 5.6% to 10.5%. Across L1-L4, VF exceeded NF by approximately 4.5% on average. This descriptive difference became more pronounced when only L2-L4 were evaluated, with VF exceeding NF by approximately 5.9% (Figure 2B). An exploratory mixed effects model using participant-level VF – NF differences did not confirm a feedback-by-difficulty effect, with a positive but

XSL•FO
RenderX

uncertain slope of 2.32% per difficulty level, a confidence interval spanning zero (–4.16 to 8.80), and an exploratory *P* value of .47. Thus, the pattern was interpreted as descriptive and hypothesis-generating. Processed participant-level mean tracking-error values for each condition are provided in Table S1 in Multimedia Appendix 1.

Figure 2. (A) Tracking accuracy improvement after training across ASF modes and difficulty levels. Bars show the percentage reduction in posttraining tracking error relative to pretraining error, with positive values indicating improved accuracy. Error bars indicate standard deviations across participants, and points represent individual participants. (B) Participant-level differences in tracking accuracy improvement between visual feedback and no feedback (VF – NF) across difficulty levels. Bars indicate group means, error bars indicate standard deviations across participants, and points represent individual participants. Positive values indicate greater improvement under VF than NF. ASF: augmented sensory feedback; L: difficulty level (1 to 4, increasing difficulty); NF: no augmented feedback; VF: visual augmented feedback.



XSL•FO
RenderX

Discussion

This pilot study examined how two essential VR training design elements (ASF and task difficulty) may interact to influence motor performance during trajectory tracking. Across conditions, tracking accuracy generally improved after training, with greater improvement when visual ASF was present, consistent with our prior VR studies [3,6]. The relatively highest increase with no ASF was observed at L1; however, reliance on lower-difficulty practice can limit progression and underrepresent the control demands required for learning more complex movements [9]. In this pilot study, the benefits of visual ASF were consistently amplified at higher difficulty levels (L2 to L4) and suggest that the benefit of ASF guidance during training increases with motor challenge. Further, our findings indicate the possibility that the difference between visual ASF to no ASF increases proportionately with difficulty level, suggesting visual ASF may either offset fatigue or facilitate greater engagement in the face of greater training challenge. Mechanistically, higher-difficulty trajectories may raise spatial control demands through greater curvature and path-following complexity, where real-time visual guidance becomes an especially useful external reference because the margin for spatial error narrows at higher difficulty. Together, these descriptive patterns offer preliminary rationale for future studies on how visual guidance cues and trajectory difficulty might be jointly tuned in trajectory-training task design [10].

Key study limitations include the small sample size, which was not powered for confirmatory inference, as well as the neurotypical participant sample, single-session design without retention testing, evaluation of only one sensory modality (visual ASF), lack of transfer to activities of daily living, and the absence of formal measures for fatigue or engagement. Thus, the results in this preliminary study are presented as descriptive and treated as hypothesis-generating. Future studies should use adequately powered designs and extend testing to individuals with neurological impairments, with longer-term work potentially exploring adaptive paradigms that adjust parameters such as visual ASF and task difficulty.

Acknowledgments

The authors gratefully acknowledge the funding support that made this research possible. The authors also thank the Movement Control Rehabilitation (MOCORE) Laboratory at Stevens Institute of Technology and all laboratory members for their support and assistance. The authors used ChatGPT 5.0 (OpenAI) to assist to improve language clarity; however, the authors are fully responsible for all research content, interpretations, and related points of accuracy and integrity.

Funding

This project was primarily supported by the National Science Foundation (CAREER award #2238880). The project was further supported by the Charles V. Schaefer School of Engineering and Science and the Provost's Office at Stevens Institute of Technology.

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request. Summary data underlying the reported results are provided in Table S1 in Multimedia Appendix 1.

Authors' Contributions

Conceptualization: RN
Data curation: YS
Formal analysis: YS
Funding acquisition: RN
Investigation: YS
Methodology: YS
Project administration: RN
Resources: RN (lead), SD (supporting), ZM (supporting), NYH (supporting)
Supervision: RN (lead), NYH (supporting)
Writing – original draft: YS
Writing – review & editing: YS (lead), SD (supporting), ZM (supporting), NYH (supporting), RN (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Participant-level mean tracking error values across training phases and conditions.
[DOCX File, 23 KB - xr_v3i1e94745_app1.docx]

References

XSL•FO
RenderX

1. Soleimani M, Ghazisaeedi M, Heydari S. The efficacy of virtual reality for upper limb rehabilitation in stroke patients: a systematic review and meta-analysis. BMC Med Inform Decis Mak 2024 May 24;24(1):135. [doi: 10.1186/s12911-024-02534-y] [Medline: 38790042]
2. Dewil S, Kuptchik S, Liu M, et al. The cognitive basis for virtual reality rehabilitation of upper-extremity motor function after neurotraumas. J Multimodal User Interfaces 2023 Sep;17(3):105-120. [doi: 10.1007/s12193-023-00406-9]
3. Sanford S, Collins B, Liu M, Dewil S, Nataraj R. Investigating features in augmented visual feedback for virtual reality rehabilitation of upper-extremity function through isometric muscle control. Front Virtual Real 2022 Nov;3. [doi: 10.3389/frvir.2022.943693]
4. Sigrist R, Rauter G, Riener R, Wolf P. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. Psychon Bull Rev 2013 Feb;20(1):21-53. [doi: 10.3758/s13423-012-0333-8] [Medline: 23132605]
5. Guadagnoli MA, Lee TD. Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. J Mot Behav 2004 Jun;36(2):212-224. [doi: 10.3200/JMBR.36.2.212-224] [Medline: 15130871]
6. Nataraj R, Liu M, Shi Y, Dewil S, Harel NY. Training-induced modulation of motor task performance and psychophysiological domains through augmented sensory feedback in virtual reality. Front Virtual Real 2026 Mar;7. [doi: 10.3389/frvir.2026.1772411]
7. Zhang L, Guo S, Sun Q. Development and assist-as-needed control of an end-effector upper limb rehabilitation robot. Applied Sciences 2020 Jan;10(19):6684. [doi: 10.3390/app10196684] [Medline: 34306737]
8. Kawano T, Kouzaki M, Hagio S. Generalization in de novo learning of virtual upper limb movements is influenced by motor exploration. Front Sports Act Living 2024;6:1370621. [doi: 10.3389/fspor.2024.1370621] [Medline: 38510523]
9. Thomas A, Paul L, Rasenyalo S, Jones B, Hendricks S. Challenge accepted: a systematic scoping review of the applications of the challenge point framework. J Mot Behav 2025;57(4):444-462. [doi: 10.1080/00222895.2025.2508283] [Medline: 40568842]
10. Kleim JA, Jones TA. Principles of experience-dependent neural plasticity: implications for rehabilitation after brain damage. J Speech Lang Hear Res 2008 Feb;51(1):S225-S239. [doi: 10.1044/1092-4388(2008/018)] [Medline: 18230848]

Abbreviations

ASF: augmented sensory feedback
NF: no ASF
VF: visual-only ASF
VR: virtual reality

Edited by I Steenstra; submitted 13.Mar.2026; peer-reviewed by CW Lau, J Fei; revised version received 30.Jul.2026; accepted 31.Jul.2026; published 21.Aug.2026.

Please cite as:
Shi Y, Dewil S, Marvin Z, Harel NY, Nataraj R
Exploratory Patterns of Task Difficulty and Visual Feedback in Virtual Reality–Based Upper-Limb Trajectory Training
JMIR XR Spatial Comput 2026;3:e94745
URL: https://xr.jmir.org/2026/1/e94745
doi:10.2196/94745

© Yu Shi, Sophie Dewil, Zachary Marvin, Noam Y Harel, Raviraj Nataraj. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 21.Aug.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Novel Virtual Reality Methods for the Treatment of Obsessive-Compulsive Disorder Using a User-Centered Design Approach: Focus Group Study

Michael Colman[1], MSci; Josie F A Millar[1], PhD, DClin; Bhagyashree Patil[2], PhD; Daniel J Finnegan[3], EngD; Mhairi Kristoffersen[1], MSc; Danae Stanton Fraser[1], PhD

[1]Department of Psychology, University of Bath, Claverton Down, Bath, England, United Kingdom

[2]Department of Computer Science, University of Bath, Bath, England, United Kingdom

[3]School of Computer Science and Informatics, Cardiff University, Cardiff, Wales, United Kingdom

Corresponding Author:
Michael Colman, MSci
Department of Psychology, University of Bath, Claverton Down, Bath, England, United Kingdom

Abstract

Background: Obsessive-compulsive disorder (OCD) is a common mental health disorder that can cause significant impairment to occupational and social functioning. While effective treatments exist, outcomes are often hampered by issues outside of the treatment protocol. Virtual reality (VR) presents a potential solution to help support the effective administration of treatment, but its application in the field of OCD has thus far been limited to replicating real-world environments used in exposure exercises, despite the potential to generate environments and stimuli beyond what is possible in reality. Participatory design methods are an effective way to ensure that the development of novel interventions is relevant to the intended end users and is grounded in their expertise and experiences.

Objective: The aim of the workshops conducted in this study was 2-fold. First, to understand the opinions of relevant stakeholders (clinicians and people with lived experience of OCD) on the current state of OCD treatment. Second, to generate novel ideas with the direct input of stakeholders for how VR might effectively be applied to support the treatment of OCD.

Methods: In total, 6 clinicians with experience of treating OCD and 6 people with lived experience of OCD participated in a combination of round-table discussions in large groups, and creative design exercises conducted in smaller groups and individual interviews. All workshops and interviews were audio-recorded, transcribed, and analyzed using collaborative qualitative analysis. Themes were then synthesized via narrative synthesis.

Results: Three themes relating to the first aim were identified: socializing to the model, effective exposure work, and therapeutic alliance (predominantly raised by people with lived experience of OCD). Five themes relating to the second aim were identified, with ideas for VR use covering 5 mechanisms by which it might support treatment: enabling effective exposure work, modeling, illustrating metaphors and the treatment framework, anxiety management, and motivation.

Conclusions: Findings from the first workshops largely restate previous findings emphasizing the importance of clients developing a comprehensive shared understanding of the cognitive model of treatment with their therapist. People with lived experience of OCD also highlighted the importance of developing trust in their therapist's expertise and ability to understand nuances of individual cases. Findings from the second workshop produced a wide variety of ideas for how VR could support the treatment of OCD. Many of these go beyond the concept of simply enabling exposure work as seen in previous literature. Methods to increase client engagement with the cognitive model during treatment were described, and people with lived experience proposed methods of managing anxiety experienced during treatment to facilitate approach behaviors. Avenues for further development of these ideas and areas for further refinement are discussed.

(JMIR XR Spatial Comput 2026;3:e92410) doi:10.2196/92410

KEYWORDS

virtual reality; obsessive-compulsive disorder; user-centered design; participatory research; cognitive behavioral therapy; virtual reality exposure therapy

XSL•FO
RenderX

Introduction

Obsessive-compulsive disorder (OCD) is a mental health disorder with an estimated global lifetime prevalence of 1% - 2% [1,2]. OCD is characterized by obsessions—recurrent and persistent intrusive thoughts, urges, and images that generate significant distress in the individual and are experienced as unwanted [3]. In response to these obsessions, the individual feels compelled to carry out compulsions—repetitive behaviors and/or mental actions intended to suppress or neutralize obsessions and/or reduce distress [3,4]. While the mechanisms remain consistent across cases, the content of obsessions and compulsions can cover diverse topics, from concerns about physical contamination to immoral or socially taboo thoughts [5,6]. OCD can significantly impact individuals' social and occupational functioning [7], and approximately 14% - 26% of people with OCD report attempting suicide at least once in their lifetime [8,9].

Given the heterogeneity of symptoms and other issues, such as common comorbidities, OCD can be a complex disorder to treat [10]. Psychotherapeutic treatment in the form of cognitive behavioral therapy (CBT) with exposure and response prevention (ERP) is widely considered an effective treatment for OCD [11,12]. However, there is also consistent evidence that a significant proportion of people with OCD do not receive CBT, or if they do, it is not technically adequate. For example, a systematic review examining patient adherence to CBT found that 15% (n=877) of eligible patients do not take up CBT at all, while a further 15% (n=139) drop out of treatment [13]. Even when patients do choose to engage in a course of CBT, studies suggest that the inclusion of ERP (a key active component of CBT) varies between 16% and 50% [14,15]. Additionally, when ERP is included, it is sometimes administered incorrectly or inconsistently. For example, the use of imaginal exposure (where the patient is required to imagine facing their feared stimuli) rather than in vivo (ie, where the patient is physically present in facing their fears) [16,17]. In some situations, imaginal exposure has been shown to be less effective than in vivo exposure, as patients might not engage as much with imaginal stimuli and can use cognitive strategies to avoid confronting situations in a manner conducive to effective CBT [18]. As such, there is increasing interest in developing methods to support the effective delivery of treatment for OCD.

Virtual reality (VR) technology has been suggested as one such method. The term "VR" refers to a viewer-centered experience in which the user feels as if they are present in a simulated 3D virtual environment (VE) separate from their physical environment [19,20]. Technologies such as head-mounted displays are used to achieve the sensation of seeing and looking around in a VE, while input devices such as controllers can be used to allow people to navigate and interact with the VE. With exposure work being such a critical part of CBT, VR can provide exposure situations more impactful than imaginal, but without the practical difficulty of in vivo [21]. For those people whose struggles with OCD prevent them from leaving their homes or have comorbid conditions such as depression, VR can also be a tool that enables access to exposure environments that are relevant but currently unrealistic to attempt [22]. Additionally, the level of control VR affords means exposure scenarios can be personalized to be relevant to individual patients' particular difficulties [23], which is pertinent given the emphasis people with OCD have placed on the need for personalized therapy materials [24].

There is substantial evidence that VR can evoke the necessary anxiety to be an effective tool in exposure work with OCD, but the research surrounding its ability to provide long-term treatment outcomes is comparatively underdeveloped [25]. Another important feature of the extant literature around VR and OCD is that all apps developed and tested thus far only add value to current treatment protocols by making exposure environments more accessible and approachable. VR enables the generation of realities with elements and environments that go beyond what is possible in the real world [20]. As such, there is scope to investigate how this capacity might be harnessed to add value to treatment in novel and previously unconsidered ways [26]. However, it remains unclear how best to implement VR effectively within OCD treatment, highlighting the need for the development of novel VR methods to support clinical practice.

Participatory design methods are increasingly being applied to the development of VR apps in therapeutic contexts and may help bridge this gap between conceptual innovation and clinical implementation [27,28]. More broadly, there is a growing call for all therapeutic interventions for OCD to be codeveloped as best practice [29]. The person-based approach (PBA) to intervention development posits that consistently involving end users and relevant stakeholders in the iterative design and development of interventions produces the best development results, maximizing both acceptability and effectiveness [30]. This involvement enables researchers to develop a nuanced understanding of the intended users' experiences, which in turn informs intervention design. Concurrently, it allows intended users to directly contribute to the core content of the interventions to ensure their relevance and impact [31]. The current Medical Research Council (MRC) guidance for complex intervention development highlights the importance of stakeholder involvement during the development phase of new interventions [32]. This study aimed to generate novel ideas for how VR could be used to support or augment the treatment of OCD through participatory design methods. In line with the PBA [30], we aimed to ground these ideas in an understanding of the relevant stakeholders' psychosocial context, including past experiences of treatment, motivations to engage in interventions, and perceived barriers to effective treatment. To achieve these goals, the study adopted a constructivist interpretive approach where stakeholder accounts were used to generate contextualized, actionable design ideas in an exploratory way. The inductive approach of the collaborative qualitative analysis was chosen in line with this constructivist approach. To do this, 2 research questions (RQs) were developed:

1. How do clinicians and people with lived experience of OCD feel about the current state of OCD treatment? (ie, what do they perceive works effectively and what could change?)
2. How do clinicians and people with lived experience of OCD envision using VR in the treatment of OCD?

XSL•FO
RenderX

Methods

Overview

This study has been reported in line with the Journal Article Reporting Standards for Qualitative Primary Research in Psychology (JARS-Qual) [33].

Participants

Two stakeholder groups independently took part in the study: clinical professionals with experience of treating OCD (hereafter referred to as clinicians) and people with lived experience of OCD. Clinicians and people with lived experience of OCD did not participate in the same workshops; all workshops involved only one stakeholder group. However, ideas generated in one group were sometimes shared with the other by the researchers as part of workshop discussions.

Clinicians were recruited through the clinical supervisor's contact network and snowball sampling methods. Following initial participation in Workshop 1, some participants suggested that the research team contact colleagues who might be interested in participating. In total, 6 different clinicians took part in Workshop 1, Workshop 2, or both. They were aged 39 - 53 (mean 47.33, SD 4.84) years, and 4 out of 6 were female. Professional background and experience data were not collected for all participating clinicians due to incomplete demographic forms; however, the inclusion criteria for the study meant that all participating clinicians had more than 10 years' experience of providing CBT treatment to people with OCD.

People with lived experience of OCD were recruited from a group of people with OCD who had expressed their interest in participatory design and taking part in OCD research to the research team. Some of these people had worked together in previous research projects. These people had attended conferences related to OCD and requested to be contacted by the research team if research opportunities arose. All participants recruited in this study had undertaken a course of CBT for OCD and were at a stage in their recovery where they felt psychologically safe to engage in group discussions about OCD and treatment. In total, 6 people with lived experience of OCD took part in one or both of Workshop 1 and/or Workshop 2. They were aged 34 - 50 (mean 42, SD 6.36) years, and 4 out of 6 were female. Data regarding treatment history were not collected for all participants due to incomplete demographic forms (Table 1).

Table . Demographic characteristics of Workshop 1 and 2 participants.

Characteristic	Clinicians (n=6)	People with lived experience of OCD[a] (n=6)
Sex (female), n (%)	4 (66.67)	4 (66.67)
Age (years), mean (SD)	47.33 (4.84)	42 (6.36)
Ethnicity, n (%)		
White British[b]	6 (100)	6 (100)
Treatment background, n (%)		
Clinical psychologist	3 (50)	—[c]
CBT[d] therapist	1 (16.67)	—
N/R[e]	2 (33.33)	—
Treatment experience in years, mean (SD)	18.75 (6.29) (N/R= 2)	—
Age when first seeking treatment for OCD (years), mean (SD)	—	23 (9.31) (N/R=2)

[a]OCD: obsessive-compulsive disorder.

[b]Participants described their own ethnicity rather than selecting a preset option.

[c]Not applicable (used when demographic information applied only to one stakeholder group).

[d]CBT: cognitive behavioral therapy.

[e]N/R: not reported (used when participant demographic data were missing).

A potential concern is that familiarity with the clinical supervisor and/or each other might bias participants toward endorsing each other's or the researchers' views and opinions, rather than expressing their own. However, the topics being discussed and participatory design methods used were different from any of the participants' previous experiences of taking part in OCD-related research. An important part of the process was authentic reflection on the topic at hand, and it was considered that these potential biases were unlikely to impede such reflection. Moreover, engaging with an already established group of stakeholders offered the advantage of a greater level of psychological safety. This is important for encouraging honest reflection and confidence to engage in the creative activities necessary to the methods of the workshops. Additionally, having already worked together as a group on different projects, they had established that they were comfortable working collaboratively in the participatory design atmosphere of the study.

Safety measures were kept in place and made clear to participants throughout the workshops, including the option to mute, turn off the camera, or leave the Microsoft Teams meeting

XSL•FO
RenderX

during workshops, or withdraw from the study between workshops or at any time during a workshop. In all workshops with people with lived experience of OCD, the lead supervisor, a clinical psychologist, was present throughout to provide psychological support and/or to offer follow-up contact if needed.

Ethical Considerations

Ethical approval for conducting the study was obtained from the University of Bath Psychology Research Ethics Committee (Application Number 22 - 107). When participants were first approached about participating, they were presented with an information sheet detailing: the purpose and procedures of the study and how the data they provided would be used for intended research outputs. They were informed of their right to withdraw their participation or data from the study either during or between workshop sessions, with the caveat that the data they contributed during group workshops could not be separated from the wider group due to the audio recording and transcription process. Once participants had read this information, they were presented with an informed consent form to sign if they wished to participate. The details of the Bath Psychology Research Ethics Committee were also provided if they had any concerns or queries. Participant contact details were kept confidential by storing them on the University of Bath's secure network. Once audio data had been recorded and transcribed, any personally identifying data were removed from the transcripts; participants were assigned a code identifier to differentiate contributions, and their contact details and names were deleted for anonymity. Participants were compensated for their time at a rate of £20 (US $26.81) per hour in line with best practice guidance [34].

Design

Two different workshop designs were implemented to address the RQs, hereafter referred to as Workshop 1 and Workshop 2. In total, 8 workshops were conducted: 2 Workshop 1s, followed by 6 Workshop 2s. Each workshop used different methodologies tailored to the stage of the participatory design process [35].

For Workshop 1, a focus group method involving open discussions was used to explore RQ1. The focus group method was appropriate for answering RQ1 as it enabled the collection of rich information on complex topics, while allowing clarification and elaboration as participants built on each other's contributions [36,37]. The aim of Workshop 1 was to provide researchers with a clear understanding of the context of current OCD treatment and to provide contextual scaffolding to inform the conduct of Workshop 2.

Workshop 2 used several creative design methods, including design games or "icebreakers" and the persona-and-scenario exercise [38] to generate ideas for methods to improve or augment OCD treatment, addressing RQ2. These methods are designed to enhance participant creativity by encouraging open exploration unconstrained by practical limitations [39]. This makes them ideal to promote the generation and exploration of novel applications of VR in OCD treatment. The aim of Workshop 2 was to generate novel ideas for how VR could be applied to support the treatment of OCD. As such, the intended output of the study was an analysis of the discourse from the workshops capturing the range of ideas produced along with participants' clarifications and elaborations (refer to the "Results" section).

Both workshops were conducted online via Microsoft Teams, facilitated by members of the research team. The research team consisted of the lead researcher, the clinical supervisor (a clinical psychologist), and a research assistant.

Procedure

Overview

Participants were initially approached by the clinical supervisor through their professional network and provided with an online information sheet and informed consent form, which included contact details for any questions. After returning the form, participants were contacted by the lead researcher to organize a time to meet.

Workshop 1: Focus Group Discussions

Two Workshop 1 sessions were conducted: one with people with lived experience of OCD and one with clinicians. The people with lived experience of OCD focus group lasted 2 hours and included 5 participants, while the clinicians' focus group lasted 90 minutes and included 4 participants. At the start of each session, the lead researcher introduced the research team and the intended purpose and outputs of the workshop. Stakeholder participants then introduced themselves, with this opening segment taking approximately 15 minutes. This was followed by 2 open whole group discussions, separated by a short break.

Discussions were guided by 2 text prompts: "past experiences of receiving/providing treatment for OCD" and "improving the treatment of OCD." Where needed, the lead researcher drew on a preprepared list of questions related to each prompt to stimulate further discussion or clarify participants' views (Multimedia Appendix 1). The first discussion explored participants' perspectives on aspects of treatment they felt worked well, as well as features that they perceived were less effective or overlooked. The second discussion prompt was "improving treatment." In a similar method to the first discussion point, this prompt was focused on highlighting areas of the current treatment protocol which could be improved or receive more attention, as well as introducing ideas outside of the protocol itself. Each discussion lasted approximately 45 minutes for people with lived experience of OCD and 30 minutes for clinicians. The clinical supervisor was present throughout both Workshop 1 sessions, fulfilling slightly different roles in each stakeholder group. In both groups, they helped mediate discussion and provide clarifications between the participants and lead researcher. In the people with lived experience of OCD group, the supervisor also acted as safeguarding lead, monitoring discussions to ensure that participants were not triggered by the content.

Participants were encouraged to contribute using the "raise hand" feature or by typing in the chat. To further support collaboration, a brainstorming platform (Padlet [Nitesh Goel]) was shared with participants and displayed throughout the

XSL•FO
RenderX

session. Participants added their ideas to the Padlet board via virtual "post-its." The research assistant monitored and organized the board, recording verbal contributions and editing entries to maintain legibility and structure for the group discussions.

Following the discussion, the lead researcher delivered a brief educational presentation (approximately 10 minutes) on VR, its applications to OCD thus far, and other mental health problems, and examples of novel therapeutic uses (Multimedia Appendix 2). The first example was Oxford VR's acrophobia treatment app, which uses gamification and an automated therapist agent to treat fear of heights [40]. The second illustrated a "feedforward" mechanism, used in athletic training, in which a cyclist's recorded performance is augmented and represented by an in-VR agent that the cyclist then races [41]. The purpose of the presentation was to broaden participants' thinking, emphasizing VR's capacity not only to replicate reality but also to generate new realities with features beyond what was possible in real life. Participants were invited to reflect on this capacity in advance of Workshop 2. They were asked to consider how VR might change or generate realities rather than focusing on technological constraints, as the next workshop prioritized idea generation over technical refinement.

Afterward, participants were thanked and offered a debrief exercise to help facilitate the transition back into their regular day. Following session closure, the lead researcher contacted participants via email to schedule Workshop 2.

Workshop 2: Creative Design Exercises

Due to the nature of the exercises and the difficulty of rescheduling all participants to meet at the same time, 6 Workshop 2 sessions were conducted. Clinician workshops were run with 1 or 2 participants and lasted 1 hour to accommodate scheduling constraints. People with lived experience of OCD workshops were held with at least 2 participants and lasted 2 hours, with the clinical supervisor present throughout. For clinicians, it was considered safe to participate individually, as their contributions were based on synthesizing professional experience rather than reflecting on personal experiences. By contrast, people with lived experience of OCD workshops involved greater personal reflection, so these sessions required at least 2 participants and the presence of the clinical supervisor to ensure safeguarding.

The workshop structure differed slightly across stakeholder groups, given clinicians' shorter sessions. Both groups began with a brief summary of the discussions and output of Workshop 1 (ie, a list of key treatment features and areas to improve). For participants attending only Workshop 2, a short presentation on VR and its potential use in treatment was also provided. This introductory process lasted approximately 10 minutes in both groups.

People with lived experience of OCD workshops began with an icebreaker and warm-up design game called "Crazy 8's" or "8-Box Blitz" [42]. Participants were asked to respond to a broad prompt ("ways to make OCD treatment better") by filling 8 boxes on a sheet of paper with ideas under a short time limit. The exercise aimed to encourage increasingly unorthodox thinking without allowing time for refinement or consideration of feasibility. Ideas were briefly shared and discussed among participants and researchers. In both people with lived experience of OCD and clinician workshops, participants then used a worksheet (Multimedia Appendix 3) and discussion to create "personas"—characters representing people seeking treatment for OCD, the target audience for the VR intervention. The research assistant took primary responsibility for recording participant contributions on the worksheet, while the lead researcher facilitated deeper persona development by asking clarifying questions.

Following a break, the personas were used as the basis for a scenario exercise. Participants were asked to describe what their persona's treatment might ideally look like in a world where VR technology was more advanced and widely available. A Padlet board with a timeline structure was displayed to support storyboarding of these scenarios. While the structure provided scaffolding to guide idea generation, participants were encouraged to share thoughts and opinions in a free-flowing manner.

In both workshops, once the persona and scenario exercises were completed, the lead researcher summarized the ideas generated during the workshop. Participants were informed that their data would be analyzed and the results would be prepared for publication, which they would be able to access. They were told that one idea from the workshops would be selected for further development, and that the lead researcher might contact them to provide feedback on the prototype. Finally, participants were offered an optional debrief exercise to help transition back into their regular day.

Analysis

All workshops were audio recorded and transcribed. Transcripts were analyzed using collaborative qualitative analysis (CQA) [43]. This 6-phase method, developed by Richards and Hemphill, is broadly rooted in thematic analysis but uses structural elements of the constant comparative method to balance rigor, trustworthiness, and transparency while addressing the challenges of analyzing qualitative data in a team environment. The decision was made to use this approach with multiple analysts due to its structural support for developing qualitative researchers and its integration of the perspectives provided by multiple researchers. In this instance, the team responsible for analysis (henceforth called "the research team") consisted of the lead researcher, a research assistant, and a clinical supervisor.

In phase 1, the research team met to discuss the analysis plan, including a week-by-week plan of team meetings, the data analysis plan, and coding assignments. The lead researcher and research assistant acted as primary and secondary coders, while the clinical supervisor acted as peer debriefer. At this stage, the research team also decided to take an inductive "data-driven" approach, which did not base the analysis on a preexisting theoretical framework. This was due to the exploratory nature of the study and its focus on ideation. However, it was noted that the analysis may often be informed by the cognitive approach to OCD because the majority of the participants'

XSL•FO
RenderX

experiences and opinions were rooted in the cognitive behavioral treatment model.

In phase 2, both coders independently and iteratively open- and axial-coded 3 of the 8 transcripts (38% of the data) over several rounds. Coders used Microsoft Word's comment function for this stage of open-coding, additionally independently writing memos after each round of coding and meeting regularly to discuss relationships between codes and discuss unclear cases. Phase 3 was initiated once the coders both felt comfortable that they were regularly identifying a variety of patterns between codes. The primary coder reviewed meeting notes, memos, and coding of both coders to develop a preliminary codebook of themes and subthemes. For this process, the qualitative data analysis software LiGRE (Logiciels Ex-l-Tec Inc) was used so that themes and subthemes could be adjusted iteratively as necessary for all coders. The preliminary codebook was discussed in a meeting and given to the peer debriefer to review and suggest adjustments. When this was complete, the preliminary codebook was pilot-tested in phase 4 by having the 2 coders apply it to previously uncoded transcripts. Discrepancies and questions about the structure of the codebook were iteratively discussed in meetings between the coders, with modifications being made as necessary and then applied to new transcripts again by both coders independently. When developing the codebook regarding how stakeholders felt VR could be applied to the treatment of OCD, a decision was made to separate specific ideas into subthemes, while the proposed mechanism by which ideas might support treatment was used to distinguish themes. When both coders felt confident about the codebook, the peer debriefer reviewed it again. In phase 5, the coders divided the remaining transcripts and independently coded them according to the updated version of the codebook. Although split coding was used, meetings between coders were still conducted regularly to discuss excerpts where coders felt unsure. A final review was then conducted where both coders reviewed all coded transcripts to check consistency, resolve any uncertainties or disagreements between coders, and decide how to handle outliers. Finally, in phase 6, the research team developed the codebook into a thematic structure to describe the perspectives of the participants, which can be used in the write-up of the analysis.

Following this, the lead researcher produced a summary of themes from Workshop 1 and a narrative synthesis of themes identified in the data from Workshop 2. From Workshop 1, 3 major themes were developed: (1) the importance of developing a shared understanding of the cognitive treatment model for OCD with clients, (2) the importance of high-quality and balanced exposure work and behavioral experiments, and (3) the importance of a strong therapeutic alliance between the client and clinicians. For Workshop 2 which focused on novel applications of VR to the treatment of OCD, 5 major themes were developed: (1) VR to support or in some cases enable effective exposure work, (2) VR as a modeling tool, (3) VR to visually illustrate aspects of the treatment model to clients, (4) VR as an "anxiety management" tool, and (5) VR as a motivational tool. For a full table of themes, subthemes, and codes from each stakeholder group, refer to Tables 2 and 3. The results were organized by RQs, themes, and subthemes, rather than by workshops or stakeholder group. This was because, while the workshops had distinct aims, themes relevant for both RQs were identified across both workshops and from both stakeholder groups.

XSL•FO
RenderX

Table . Themes, subthemes, and code data for research question 1[a].

Themes and subthemes	Codes
Socializing to the model	
__[b]	• 24 codes • 3 clinicians (19) • 3 people with lived experience of OCD[c] (5)
Exposure work and behavioral experiments	
Effective timing of exposure work and behavioral experiments	• 17 codes • 3 clinicians (8) • 3 people with lived experience of OCD (9)
Keeping the quality of work consistent and model-focused	• 16 codes • 2 clinicians (8) • 4 people with lived experience of OCD (8)
Therapeutic alliance	
Establishing and maintaining a compassionate stance	• 18 codes • 2 clinicians (5) • 4 people with lived experience of OCD (13)
The therapist's theoretical understanding of OCD	• 17 codes • 6 people with lived experience of OCD (17)
The therapist's ability to understand nuances of an individual case	• 20 codes • 1 clinician (1) • 5 people with lived experience of OCD (19)

[a]For each subtheme, the total number of codes identified is recorded. Then, the number of members of each stakeholder group who mentioned these codes is recorded. Finally, the number of codes coming from each stakeholder group is also recorded next to them (in brackets).

[b]Not available.

[c]OCD: obsessive-compulsive disorder.

XSL•FO
RenderX

Table . Themes, subthemes, and code data for research question 2[a].

Themes and subthemes	Codes
Supporting and enabling effective exposure work	
Enabling exposure for impossible or imaginal scenarios	• 27 codes • 4 clinicians (20) • 2 people with lived experience of OCD[b] (7)
VR[c] as a “stepping stone” for in vivo exposure and supporting work outside the office	• 19 codes • 5 clinicians (12) • 4 people with lived experience of OCD (7)
Modeling behavior and experience	
__[d]	• 12 codes • 4 clinicians (6) • 2 people with lived experience of OCD (6)
Illustrating metaphors and the treatment framework	
Representing an externalized OCD	• 33 codes • 3 clinicians (26) • 3 people with lived experience of OCD (7)
Keeping the treatment framework present	• 28 codes • 5 clinicians (24) • 3 people with lived experience of OCD (4)
Anxiety management	
When overwhelmed during exposure	• 18 codes • 2 clinicians (2) • 4 people with lived experience of OCD (16)
For grounding purposes outside exposure	• 20 codes • 1 clinician (5) • 4 people with lived experience of OCD (15)
Motivation	
—	• 36 codes • 4 clinicians (19) • 2 people with lived experience of OCD (17)

[a]For each subtheme, the total number of codes identified is recorded. Then, the number of members of each stakeholder group who mentioned these codes is recorded. Finally, the number of codes coming from each stakeholder group is also recorded next to them (in brackets).

[b]OCD: obsessive-compulsive disorder.

[c]VR: virtual reality.

[d]Not available.

Reflexivity

The research team consisted of the lead researcher (a PhD student researching OCD and VR), a research assistant (Master’s graduate in applied clinical psychology), and the clinical supervisor (a clinical psychologist specializing in CBT for OCD). The lead researcher had experience working within cross-disciplinary VR research but no prior clinical experience of OCD and no previous experience conducting coproduction research. The research assistant had previous experience conducting OCD-related research, including studies using coproduction methods and the PBA. The clinical supervisor brought 18 years of clinical experience treating OCD and previous research experience developing OCD interventions using coproduction methods.

The lead researcher conducted the workshops and acted as the primary coder and author of the analysis; the research assistant supported the conduct of the workshops and acted as the second coder; and the clinical supervisor provided pastoral support during workshops with people with lived experience of OCD and acted as the peer debriefer during analysis. While the clinical supervisor was present during these workshops, they did not actively contribute to idea generation and only acted in a safeguarding capacity. During the analysis procedure, the clinical supervisor also did not make any unilateral coding decisions, but provided clarification on clinical terminology and frameworks to the coders.

The study was informed by the PBA and conducted by a team with interests in OCD treatment, intervention development, and

XSL•FO
RenderX

VR. As such, the authors acknowledge that participants' accounts were interpreted through clinical, technological, and intervention-development perspectives, which may have influenced the construction of themes. The clinical supervisor had existing professional relationships with most participants, which may have influenced what participants felt comfortable sharing, although these relationships may also have facilitated trust and psychological safety during workshops. To promote reflexivity throughout the study, both coders maintained reflexive journals, recorded analytic memos, and met regularly to discuss assumptions and interpretations.

Results

RQ1: How Do Clinicians and People With Lived Experience Feel About the Current State of OCD Treatment?

Three themes were developed that represent how both clinicians and people with lived experience of OCD perceived the current state of psychological treatment for OCD and its most important features to provide scaffolding for idea generation in Workshop 2. These three themes are (1) developing a shared understanding of the model; (2) quality exposure work and behavioral experiments, which comprises 2 subthemes (effective timing of different exposure exercises and keeping the quality of exposure work consistent and model-focused); and (3) therapeutic alliance, which comprises 3 subthemes (establishing and maintaining a compassionate stance; the importance of the therapist's theoretical understanding of OCD; and the importance of the therapist's ability to understand individual patients' cases). The importance of these 3 themes is largely well-established in the literature surrounding CBT for OCD [44-46], but they are briefly summarized here to contextualize the findings from Workshop 2.

Both stakeholder groups appreciated the importance of therapists and their clients properly understanding and applying the cognitive model to all aspects of treatment. This goes beyond the therapist understanding the model and ensuring it is applied to also ensuring the client has understood and "bought in" to the model, its necessity, and the effort it takes.

> *Getting the buy-in is absolutely crucial, so how you set it up, the understanding, by doing your Theory A, Theory B, so by the time you're doing the active elements, which are absolutely crucial […] they've already bought into the need to do that.* [C2]

Similarly, both stakeholder groups felt that effective exposure work was a key element of the success of treatment. They emphasized the importance of keeping the exposure work's quality consistent and focused on the cognitive model. This was raised as an area of difficulty by both groups, highlighting that therapists are not always able to regularly check in on their clients to support the treatment work they do outside the office and "squeeze the learning out of it" (C1). Effective exposure work was also characterized by the stakeholders as being a balance between scheduled and unscheduled exercises, and clinicians and people with lived experience of OCD highlighted that exercises based on an adaptable plan allowed for the client to move at a pace that best allowed for applying learning and growing confidence. However, the benefit of unplanned exercises was also highlighted, with both groups suggesting these might be more effective at dismantling OCD beliefs due to their unexpected and uncertain nature.

> *When I knew what was coming, it was easier, if that's even the right word, than what I didn't know was coming, and those were the best ones for me to get better.* [P4]

Finally, both groups highlighted the importance of a strong therapeutic alliance between client and clinician, although people with lived experience of OCD spoke more on this topic and suggested a different area to improve than clinicians. Both groups highlighted the importance of maintaining a compassionate stance to establish trust that the therapist will support clients to engage fully with treatment. Additionally, people with lived experience of OCD highlighted how the therapist's demeanor can be critical in allowing clients to fully explain the breadth and nature of their concerns, particularly concerning taboo obsessions.

> *People don't always admit everything that they perhaps might admit- that more sort of perhaps less taboo aspects of OCD […] and they'll leave some of it out that they don't wanna talk about. So I guess it's all about getting, making the person, the patient, feel comfortable enough to be okay with doing that.* [P5]

Otherwise, exclusively people with lived experience of OCD highlighted how variance in the quality of treatment in their experience was often caused by variance in the therapist's expertise and ability in 2 areas: their theoretical understanding of OCD and their understanding of an individual's case. Strong theoretical understanding of OCD was said to increase clients' confidence in their therapist's ability to help them, and several people with lived experience of OCD recounted negative experiences of treatment stemming from a lack of theoretical understanding. The therapist's understanding of a patient's idiosyncrasies was an often-reported problem by people with lived experience of OCD, who felt that this understanding allowed for greater adaptability of the CBT protocol. This adaptability was highlighted as an area for improvement and was, in turn, thought to help motivate clients to engage with therapy and help them navigate situations where they felt the CBT protocol was not working for them.

RQ2: How Do Clinicians and People With Lived Experience of OCD Envision Using VR in the Treatment of OCD?

Overview

Five themes were developed encompassing the ideas clinicians and people with lived experience of OCD generated regarding how to apply VR to the treatment of OCD: (1) supporting and enabling effective exposure work, (2) modeling typical behavior and "self"-modeling, (3) illustrating metaphors and the treatment model, (4) "anxiety management," and (5) motivation. Themes 1, 3, and 4 each have 2 subthemes. These themes were not separated based on stakeholder group, as several were independently coded in both groups. Instead, to best examine

XSL•FO
RenderX

differences or similarities of perspective, data relating to each theme were grouped together (with specific references to stakeholder group [clinician=Cx; people with lived experience of OCD=Px] noted where relevant).

Theme 1: Supporting and Enabling Effective Exposure Work

Overview

This theme highlights how clinicians and people with lived experience of OCD felt VR could benefit treatment simply by increasing the possibility and accessibility of exposure work for clients. The ideas contained in this theme extend upon existing literature suggesting VR can increase access to exposure scenarios by demonstrating OCD-specific situations which could be especially well-addressed by the reality generation capabilities of VR. The ways in which VR could be applied for this purpose are delineated in 2 subthemes.

Enabling Exposure for Impossible or Imaginal Scenarios

Clinicians suggested that an obvious use for VR's reality-generation capabilities is as an effective replacement for imaginal exposure. Imaginal exposure is an exposure exercise where the client uses mental imagery to expose themselves to anxiogenic situations in their imagination. Clinicians highlighted that some cases can be extremely difficult to treat with effective in vivo (in-person) exposure work, particularly with symptom dimensions of OCD where the necessary exposure scenarios are not possible to facilitate.

Clinicians proposed 2 types of impossible scenarios that could be enabled through VR. First, scenarios that are practically too difficult to facilitate, for example, going to a busy swimming pool with children around (to address obsessions related to pedophilia). Second, scenarios that are physically impossible to facilitate, such as causing harm to come to loved ones or moral obsessions about the afterlife. In both types of scenarios, clinicians proposed that VR would essentially act as a more visceral replacement for imaginal exposure, creating impactful scenarios in instances where in vivo is not possible, thus enabling effective exposure work with clients.

> *I tried just the stunt walking on the plank thing, and […] rationally, you know that you're on the floor, but it doesn't feel it and I think that is the powerful thing about VR is that it does capture the emotion in a way that I think most rational people wouldn't think is possible and in physical sensations, you know, get right into that. Even though I can say 'I know I'm on the floor, I can walk, whichever,' I know that it's still hard to do. So, there is something about it which is more powerful than imaginal exposure, I think, and I'm slightly reluctant to say that, but...* [C6]

There was debate, from both clinicians and people with lived experience of OCD, over how far this concept of generating impossible or nonreal scenarios could be usefully implemented. For example, the ideas of a "worst case scenario" (C3) exposure or visually demonstrating the spread of contamination as imagined by a person with OCD in an exposure exercise were suggested. Arguably, both could be useful exposures: the former allows clients to challenge beliefs related to coping with feared outcomes. The latter increases the intensity of the exposure, potentially making the work done during it more impactful ("I would want to face absolute destruction and know that I could still go in and make dinner" P7). However, both ideas received criticism: learning to tolerate doubt and uncertainty of potential outcomes is an important part of treatment. Concerns were raised that using VR to show something that you normally cannot see (such as the spread of germs) might do more harm than good.

> *I'm not sure if that- in a way I'm, I just think perhaps it, that would feed into someone's OCD in that you're kind of, you know, making it visual, something that- that wasn't before? Umm, that's just my personal kind of thing, I don't know, I just feel like that might be utterly terrifying, and like worse than reality in some ways.* [P5]

A further benefit raised by clinicians and people with lived experience of OCD was the potential for increased personalization of exercises to the client. It was felt that personalization would increase ecological validity and impact of exposure work, allow for transferring learning to different relevant environments, and increase motivation as client goals could be worked into the exercises. Both groups acknowledged the difficulty of balancing between widely applicable VR apps and the idiosyncrasy of OCD presentations.

> *I mean, you could model the place as well, couldn't you, so then you're obviously changing the location, so she could have an office, lunch hall, she could have her dad's house so that the stuff that she's learning can be applied to different areas as well.* [P7]

VR as a "Stepping Stone" for In Vivo Exposure and Supporting Work Outside the Office

Clinicians and people with lived experience of OCD consistently felt that VR, through its ability to induce a sense of presence, could support treatment as part of a hierarchy as an introductory step toward in vivo exposure. Both groups mentioned examples of people being "stuck" when it came to exposure work and felt that this "stepping stone" use of VR could be the answer in a lot of cases.

> *For me the kind of beauty of VR would be on the one hand, my brain is absolutely in it. I've touched it. I've done this. I've done that. I've got that whole feeling. But there's a little part of my brain can go 'OK, I can try these things out. I can try and tolerate the anxiety I'm feeling' without having that whole fear of that undoing.* [P1]

Within the bounds of the scenario exercise (a future where VR is more commonplace and available than it is today), clinicians also felt VR could be used to support exposure work done outside the office, for example, through remote work or homework. They proposed the control VR allows over the environment could ensure exposure exercises were run in a consistent manner, whether the therapist was present or not. For example, clients could do effective work in a session, the therapist could note what worked best and then provide a structured recreation of that via VR to be taken away for

XSL•FO
RenderX

homework. This version of homework would remove extraneous variables that might affect client homework otherwise.

Theme 2: Modeling Typical Behavior and "Self"-Modeling

This theme focuses on how people with lived experience of OCD and clinicians felt VR could be used to model typical behavior and thought patterns to clients with OCD whose experience may mean they are unsure what this looks like. C1 highlighted that some clients do not know how to behave in the absence of OCD due to the prolonged course of their symptoms and the isolation from others. Thus, VR could provide a visual demonstration of a character modeling behaviors (eg, having a shower) in a typical or antiobsessional way. P1 suggested this could also be applied in a new way beyond the physical demonstration to visually model typical mental processes and highlight how intrusive thoughts appear for everyone, but that the interpretations of those thoughts differ between people with and without OCD:

> *I was trying to work out how you could really show the person who's doing the VR that the person who's in the VR doing the activity wasn't having the same thoughts, but so I was thinking like word bubbles, but then also kind of possibly like heart rate down the side of just like average whereas you know like I would know mine would be going 300 beats a minute and theirs is just like 60.* [P1]

Clinicians and people with lived experience of OCD felt that this modeling strategy could be augmented by having the character they watch in the VR actually be a version of themselves, building upon the idea of the feedforward mechanic shown in the educational presentation at the end of Workshop 1 [41]. Both groups felt this feedforward idea could be adapted from its use in an athletic context in the presentation into a form of "self"-modeling that could act as a form of rehearsal for the client or act as an approach supporting behavior, something which encourages clients to attempt to engage in the exposure themselves. People with lived experience of OCD felt this could help support generalizing learning to different situations as well, as the avatar of the client could be shown doing exposure exercises in many different environments.

Theme 3: Illustrating Metaphors and the Treatment Model

Overview

This theme reflects how clinicians and people with lived experience of OCD felt VR's visualization capabilities could be used for the purpose of better integrating metaphor into treatment and supporting clients' application of the cognitive model to their treatment experience. Metaphor was frequently mentioned in both Workshop 1 and 2 as a technique therapists use to help clients understand and address their OCD, and VR could act as a more impactful delivery system for those metaphors. These ideas are delineated into 2 subthemes: the first is externalizing and visually representing OCD as a metaphorical character in the VR with the client—this has some research support already in the field of anxiety disorders, so stakeholders' ideas may be an adaptation of this [47]. The second is using VR to support clients' understanding of the treatment model and actively apply it throughout their treatment using novel context-relevant cues to keep the model present in clients' minds.

Representing an Externalized OCD

Clinicians felt that rendering OCD as a visual entity in VR separate from the client could help better demonstrate the cognitive processes (eg, lack of disconfirmation, the effects of safety-seeking behaviors, and neutralizing) at play in the maintenance and treatment of OCD. This could be done as a stand-alone activity where the therapist explains treatment concepts with a metaphor for OCD while the patient watches a visualization of this metaphor in VR. For example, OCD is sometimes explained to clients with the "Anxiety Monster" metaphor. This metaphor represents the process by which OCD begins as an innocuous response to normal anxiety, but over time, as the client interprets it as a threat and seeks safety rather than confronting it, it becomes a more dominant force that grows and controls them more.

Alternatively, clinicians suggested having the externalized OCD be diegetically present during VR exposure work and reacting to the client's actions to reinforce the learning clients should take from the exposure. For example, if a client successfully carried out an exposure without responding, the OCD representation might shrink in size, or express upset and metaphorically "lose power," which would be shown to the client in real time.

> *When you start starving that monster, when you start saying no, the monster starts losing weight, getting smaller and so eventually his voice is so small it no longer controls you. So you can do some kind of virtual reality imagery around this […] It's all about the person practically playing out saying "no, I'm not gonna give you that."* [C3]

There was significant discussion from both clinicians and people with lived experience of OCD over the most appropriate form for the externalized OCD to take and the nuances of its design. For example, clinicians highlighted The Anxiety Monster and The OCD Bully as 2 metaphors they regularly used in their current treatment of OCD to illustrate the effect of counterproductive strategies such as safety-seeking behaviors on the maintenance of OCD. However, some felt that a visualized OCD Bully might be too visceral an image as many people have previously experienced bullying. Additionally, some felt that any sort of Anxiety Monster should not be too "evil" but rather "annoying," having developed from a relatively harmless symbol of anxiety. Conversely, P7 felt that the idea of OCD as a "monster" that shrinks was not helpful, as it did not effectively acknowledge that the origin of the thoughts manifesting as OCD has come from a place of normal anxiety that at one point acted as a form of self-preservation.

> *I kind of always see it as, you know, if you're challenging these things, this thing grows because it's not as frightened anymore, you know, rather than it be a scary, horrible monster, which I know will probably work for a lot of people, I prefer the*

XSL•FO
RenderX

> *visualization where things grow the less you do it because you're enriching it.* [P7]

Similar to this idea, C1 suggested visualizing an externalized "self"; an image of the client that exists in the VR alongside a representation of the OCD. When the client is doing an exposure, their virtual "self" supports and reminds them why they are doing exercises they find difficult to tackle OCD, and as compulsions are resisted, their virtual "self" grows while the OCD representation shrinks.

Additionally, people with lived experience of OCD suggested that externalizing OCD might help make it feel more contained and manageable when addressing it in treatment, although this is likely to differ depending on the individual and demonstrates the importance of accounting for individual preferences. Both groups suggested it might help clients in thinking about cognitive aspects of the treatment such as Theory A and Theory B during exposure. The externalized OCD might support the process of examining their thoughts and interpretations during exposure and highlight "OCD interpretations" of intrusive thoughts.

Supporting the Application of the Cognitive Model During Treatment

Clinicians and people with lived experience of OCD felt that several aspects of the cognitive treatment model could be supported with the use of VR. Given the importance of understanding and applying the treatment model to everything clients do, VR could effectively provide this in a variety of ways.

The predominant method suggested to help clients apply the model to their work in therapy was to provide prompts during VR exposure exercises encouraging reflection on their cognitions. The prompts could be simple pop-up reminders, either as text like "It's the way you're responding to the thought, not the thought that's the problem" (P5) or diegetically embedded in the scenario, for example, a character in the scenario talks to the client. However, clinicians felt that an interactive system active during clients' exposure exercises would support them in actively reflecting on their experience and applying it to their learning about their cognitions and reduce the potential for them to just passively experience and endure the exposure. The VR app could periodically interject prompts in the form of questions for the clients that they can respond to as they are doing the exercise about how they are feeling and behaving.

> *So if the programme is saying "are you doing anything to keep yourself safe in any way?" for the person to then be able to say, "well, actually yes, I'm avoiding looking actually at the child, I'm looking slightly next to it." So then that then prompting something about also "what do you need to do now? You need to get closer" or so that there's, it's much more of a sort of two way interaction and for them to be able to say I am doing this, theory B would tell me that and this is what I'm learning, not just to be told what you're learning.* [C6]

Clinicians noted that this process would require the supervision of a therapist to monitor how the client was reacting to the exposure in terms of anxiety levels and reassurance or avoidance strategies, as these can be difficult for the client to note themselves. These prompts could be random or linked to biofeedback, with the former encouraging consistent application of the model and the latter ensuring clients do not forget to apply the model when overwhelmed.

C5 suggested using VR as a tool for metacognitive work as part of the more cognitive aspects of the treatment model. Metacognitive work involves exercises designed to challenge clients' currently held maladaptive beliefs about how the world works, for example, an inflated sense of responsibility for adverse outcomes or issues of magical thinking. They suggested VR's ability to facilitate impossible situations could be used in a similar fashion to the "script-writing" technique. This exercise is designed to translate the client's beliefs and thoughts into a medium that allows them to then be examined more closely and objectively; for example, the client writes about their OCD as if they were writing a script for a movie. This allows cognitions and beliefs to be examined, questioned, and played with by the therapist and client together to challenge their logic and reduce their power over the client.

> *It could be like […] "OK, alright, here's an avatar of my brother, how should he die today? Meteor, or maybe eaten by a dragon […] let's have a choice" and something like that then, rather than treating the thoughts as really serious and something you've got to deal with is like the most important thing actually just going, "it's just a shitty thought that you don't have to pay any attention to" and in fact, let's mess about with the thought."* [C5]

Theme 4: "Anxiety Management": Facilitating Approach and Building a Toolkit

Overview

This theme was primarily discussed by people with lived experience of OCD and showcases ideas not considered by the researchers or clinicians about how VR could be used as a tool to help clients manage the extreme anxiety they feel during treatment in ways conducive to their treatment. The intention behind this would be to support continued engagement in treatment and can be explained in the 2 subthemes. First, to use VR as an "anxiety management" tool within exposure exercises to support clients through feelings of overwhelming anxiety that might prevent them from engaging fully in the exposure. The second refers to the use of "anxiety management" tools and techniques outside of exposure work as part of a wider toolkit clients can use alongside their CBT.

When Overwhelmed During Exposure

Extreme anxiety is commonly experienced in exposure exercises. People with lived experience of OCD suggested that VR could be used to support clients if that anxiety ever became extreme to the point of being overwhelming and preventing them from engaging in the exercise properly. They proposed using motivational prompts (such as voice messages or images from loved ones) or having access to a virtual "safe space" to

XSL•FO
RenderX

help manage anxiety and allow continued effective engagement in the exposure. It was raised that these tools could be used both to complement in vivo exposure as well as part of a VR exposure.

> *When I'm in the middle of an exposure and my anxiety's sky high […] I totally forget what to do, so it might be helpful to have like prompts saying 'do you want to see a face of a loved one? Do you want to go to your safe place?' and then you just press a button to do that. Because […] I just freeze.* [P1]

However, people with lived experience of OCD acknowledged that the use of these tools could constitute avoidance strategies if not implemented properly and perhaps hamper progress.

> *There's two aspects to it, isn't there? Like if, if you were doing that to remind yourself why you were continuing with the exposure, I think that would be like in line with ACT, but then if you were doing it to reassure yourself and to avoid the exposure, then you'd be negating what you've done […] so it's very important to like to do it in the ACT way rather than the avoidance way.* [P4]

Building a Toolkit for Anxiety Outside Exposure

Primarily, people with lived experience of OCD felt that VR could provide clients with tools they could use outside exposures; both to ground them immediately afterward, and to develop anxiety management skills for wider recovery.

People with lived experience of OCD suggested several examples of tools that could be used in both instances, including "safe" or meditation spaces (optionally cocreated with the patient), guided imagery, grounding exercises (such as the 5-4-3-2-1 anxiety coping exercise), or even readily available VR games. For immediate postexposure purposes, there was an emphasis on these activities being quick and guided, due to the energy it takes to do exposure exercises. It could even automatically move the client from exposure to grounding, removing the impetus from the client to come up with a strategy after a difficult session.

> *I think there should also be the option to be to be able to access it away from exposure time, because that's a skill, isn't it, mindfulness and grounding. So if you had the option to visit those techniques and practise them at other times, even when it's not exposure time, I guess you'll get better at them?* [P7]

Building on the idea of a "safe space," C1 proposed externalizing a representation of the client's self-image as an avatar in the space called a "compassionate self." The client could customize this "compassionate self" however they wanted, and it would then walk them through the postexposure routine a therapist would normally do. C1 felt this could give clients a powerful image of self-compassion to hold onto when not in the therapy room.

Theme 5: Motivation

Clinicians and people with lived experience of OCD felt that VR could be used as a motivational tool by demonstrating to clients what their lives could look like when OCD disrupts them less severely. They proposed this might help clients engage in the difficult parts of therapy by making their goals more consistently apparent and reminding them why the difficulty of treatment is worth it.

One proposed mechanism for this (both within and between sessions) was to immerse people in environments and scenarios that they aspired to engage in after therapy. Both groups felt this could help clients build goals and reclaim parts of their life they felt were "lost" due to OCD. A clinician even suggested an "unlocking" mechanic, where the client symbolically "unlocked" and reclaimed parts of their life by successfully completing exposure exercises relevant to their goals.

> *You know, we obviously get people to set goals and think of an imagined future that is better than the one that they have, but we could capitalize more on that. Potentially you could have non-OCD land; you could just go and visit that.* [C2]

Both groups also discussed the idea of visualizing clients' progress and therapy journey. Clinicians felt this could act as a good motivational tool if clients could predict how they would fare in exercises and then watch their actual performance back. People with lived experience of OCD additionally felt this might help clients break their treatment down into understandable steps of how they will address their OCD. However, they also highlighted the potential for this to backfire and affect self-esteem if clients felt they were not achieving their goals.

Alternatively, similar to the "self"-modeling mentioned in Theme 3, clinicians and people with lived experience of OCD proposed using an avatar of the client as a motivational tool. This was also likely inspired by the presentation of the feedforward mechanic [41] at the end of Workshop 1, but adapted to the context of motivation to engage in treatment for OCD. One idea was to have this avatar demonstrate behaviors the client aspired to do (in a highly personalized format for increased emotional weight) in a way they currently felt unable to with their OCD, perhaps even with visual thought bubbles demonstrating their mental processes as well. Another was to have the avatar be diegetically present during exposure exercises, either as a visual reminder in the background of the client's goals, or alternatively as a supportive figure giving the client direct encouragement.

> *Is she able to kind of just watch herself going around and making food without any repetitive motions or anything like that, you know, and maybe her son is just sitting there coloring and he's not given up waiting for dinner because it takes her forever or something like that?* [P7]

Finally, P7 suggested gamification of treatment content may be useful for some people. For example, they felt that competing against your past performances in an exposure exercise could be quite a good motivator to push harder and could be a reward for some people. However, they expressed the possibility that "losing" could damage confidence and motivation, as well as the fact that gamification likely would not universally motivate everyone.

XSL•FO
RenderX

Discussion

Principal Findings

This study aimed to generate novel ideas for the application of VR to the treatment of OCD from relevant stakeholders who might benefit from the use of such ideas in treatment. To achieve this, we took a 2-phased approach using participatory design workshops to first understand stakeholder perceptions of the current state of OCD treatment, then, second, generate ideas for novel applications of VR to support treatment. The PBA for intervention development [30] posits that the generation of new ideas should be rooted in an understanding of the context of the issue. By discussing the current state of treatment with stakeholders in the first phase of workshops, this context was developed and shared between all researchers and participants. This then supported the generation of some new ideas related to VR and the refinement and application of others to the context of OCD in the second phase of the workshops.

The first phase of workshops investigated stakeholder views of the current state of OCD treatment and identified several important features of successful OCD treatment. These themes were (1) developing a shared understanding of the model, (2) quality exposure work and behavioral experiments, and (3) therapeutic alliance. The findings from Workshop 1 are largely in line with previous literature about high-quality treatment for OCD [44-46]. Of particular note from the workshops was the emphasis on the need for the client to understand how the CBT treatment model works and to buy into its effectiveness, and for treatment to be adaptable and personally relevant to clients.

Workshop 1 provided researchers and participants with a contextual understanding of the perceived current state of OCD treatment, including key features to keep and areas to develop upon. Workshop 2 produced a wide range of ideas for how VR could support the treatment of OCD. These could be delineated into 5 themes based on the proposed mechanisms for how they would support treatment: (1) enabling exposure work, (2) modeling, (3) illustrating the treatment model, (4) “anxiety management,” and (5) motivation. It is worth noting that the novelty of the ideas proposed in Workshop 2 varies. For example, the idea of a self-modeling avatar is likely not totally organic, but based around the idea of feedforward presented as an example at the end of Workshop 1. It could be argued that this presentation might have primed workshop participants to consider this particular idea for development. Additionally, the idea of enabling exposure exercises for impossible scenarios is not entirely novel, but rather an extension of previous VR research in OCD where it is used to provide access to normally inaccessible situations [18,21]. As such, the claim that all the ideas generated from Workshop 2 are entirely novel should be tempered with the understanding that some of these ideas likely have origins in the participants’ background and the researcher’s influence.

Nonetheless, the findings from Workshop 2 suggest that there is a much broader scope for the application of VR to the particular context of the treatment of OCD than has previously been explored [25]. While both stakeholder groups contributed to all themes, ideas from clinicians predominantly centered around how VR could support consistent and high-quality application of the cognitive model. Perhaps expectedly, there were several ideas involving VR in exposure work, for example in enabling exposures for scenarios which are normally impossible to facilitate and thus helping ensure the quality of exposure conducted in CBT [16,17]. However, despite previous studies proposing VR simply as an effective one-for-one in vivo replacement [25,47,48], the majority of ideas presented in Workshop 2 went beyond this. For example, by making cognitive concepts of the model accessible to clients and interactive during exposure work through integrated real-time metaphorical and theoretical framework prompts and self-modeling tools. Several of the proposed ideas already have some research support in other areas, so are not entirely novel but have been adapted to the field of OCD by the stakeholders. For example, several participants proposed adapting the idea of feedforward for athletics training [41] that was presented at the end of the first workshop into a self-model for exposure exercises. Additionally, some suggested the idea of the creation and destruction of an anxiety avatar, which has seen some interest already in other mental health disorders such as anxiety [47]. It is possible that the use of VR can help clients engage in metaphors in a more tangible, multisensory way, thus making it easier to continue applying and learning from the cognitive model [49]. VR could also augment the cognitive element of CBT through visualized metacognitive strategies to address clients’ thought processes more directly. Recent reviews suggest that metacognitive therapy using methods like the ones described in the “Supporting the Application of the Cognitive Model During Treatment” section may be as effective as gold-standard ERP in reducing OCD symptoms [50]. As discussed in the “Results” section, VR could support the administration of these exercises by allowing visualization and interaction with these exercises and concepts, highlighting a completely untapped avenue for VR to enhance treatment.

People with lived experience of OCD made suggestions that aligned with those of clinicians; however, they predominantly focused on using VR to support engagement with treatment. This was proposed through the use of “anxiety management” tools to prevent overwhelming anxiety and disengaging from effective exposure work, and through motivational measures to support engagement with the treatment as a whole. The purpose of exposure exercises and behavioral experiments is to elicit similar feelings to those clients experience in situations they find difficult in their daily lives. By confronting these feelings and reflecting on the cognitive model, they can change their beliefs and behavior around them [51]. The focus of people with lived experience of OCD on these ideas surrounding “anxiety management” suggests that there may be value in investigating VR’s ability to address this overwhelming anxiety as part of a treatment protocol for OCD. This idea is hitherto uninvestigated in the field of VR for OCD; however, there has been some evidence of success with this idea in generalized anxiety disorder [52,53]. The primary mechanism of these instances is the idea of a relaxing “safe” space, so the other ideas mentioned by stakeholders, which go beyond this, may provide plentiful opportunities for further research.

XSL•FO
RenderX

Limitations

Several limitations of the study should be considered when interpreting these findings. First, the stakeholder sample was relatively small and homogenous. Both the clinicians and people with lived experience of OCD who participated in the study were recruited from established OCD research and involvement groups, or from the clinical supervisor's contact networks. Consequently, clinicians were highly experienced in treating OCD, while people with lived experience of OCD had previous experience participating in OCD-related research and were comfortable discussing their experiences in group settings. The findings may therefore reflect the perspectives of a particularly engaged and research-active group of stakeholders and may not fully represent the views of individuals with OCD who have had less contact with treatment, research, or coproduction activities. This is relevant to bear in mind, as several ideas proposed in the workshops saw immediate disagreement between stakeholders, so it is possible that people with lived experience of OCD who are not comfortable engaging in research or treatment might object to other proposed ideas. Additionally, all participants identified as White British, and approximately two-thirds of participants in both stakeholder groups were female. Given evidence that OCD presentation, help-seeking behaviors, and treatment experiences can vary across demographic groups [54-56], alternative priorities and perspectives may have emerged from a more diverse sample.

Second, the small size of the sample in the study may raise concerns over whether the array of ideas in this study can be considered exhaustive. The study was exploratory in nature, with a focus on generating a breadth of stakeholder-informed ideas for future intervention development, rather than providing a comprehensive account of stakeholder experiences. Data saturation was therefore not specified as a stopping criterion and was not formally assessed. The CQA approach taken in this study is based on thematic analysis, where there has been some debate over the utility of trying to achieve data saturation [57]. While a broad range of perspectives and design ideas were generated across workshops, additional stakeholders may have contributed further ideas or refinements. The findings should therefore be viewed as an initial set of stakeholder-informed design concepts rather than an exhaustive account of all potential applications of VR within OCD treatment. Logistical constraints meant that Workshop 2 sessions with participants were often limited to 60‑ to 90-minute meetings, which limited the development of some of the ideas. Ideally, according to the PBA [30], there would be a process of idea refinement following the initial generation of ideas. Due to the time constraints of the project, we could not engage in this stage of the approach. As such, while many ideas from Workshop 2 contain sufficient detail to begin prototyping, there are some which lack detail on how they might specifically be implemented or how they would work in practice. In these cases, the ideas are better categorized as design concepts rather than fully formed methods ready for development and implementation. However, given the primary aim of the workshops was to generate novel ideas, the goal has still been achieved, albeit with some areas for elaboration left outstanding. Future researchers could continue the process of iterative development with more design workshops and larger, more diverse samples. However, these could be extended to include iterations focused on refining existing concepts to translate more of these ideas into actionable development prospects.

In a similar vein, the thematic structure of the analysis could be argued to be a limitation from the perspective of clinical psychologists looking to develop specific interventions from the ideas in this study. While this study gives readers an overview of the potential for VR to support treatment, it does not provide explicit, fully refined ideas that can be taken into further development. For example, the idea of the "self" avatar spans several different themes, including modeling, motivation, and "anxiety management," and there is no single description of how to implement it into treatment. Researchers seeking to test and implement a specific idea would need to first carefully check each theme and subtheme of the analysis to understand the proposed mechanisms through which that idea could support treatment. Further development of any of the ideas contained in this study should involve further elaboration and consultation with relevant stakeholders in the design process.

Implications for Future Research

The major implication of this study for future research is that it provides a wealth of stakeholder-supported novel VR methods which could be developed and applied to support or improve the implementation of CBT for OCD. Furthermore, due to the participatory design nature of the study, these methods are targeted toward addressing issues raised by people with the most experience of OCD who best understand the difficulties of its treatment. As such, this workshop study provides several potential treatment tool mechanisms for other researchers to develop and investigate. For example, enabling the conduct of exposure exercises which are normally impossible and augmenting the learning obtained from exposure exercises by more directly including aspects of the cognitive treatment model seem to be the most endorsed and well-developed concepts from the workshops. These are also the most geared toward supporting the high-quality administration of the cognitive treatment model. However, there was also significant support for ideas focused less directly on the model, such as visualizing client goals and supporting the management of anxiety during exposure exercises to maintain effective engagement. Further participatory design or development projects could select, expand, and refine ideas raised in these workshops to produce applications supported by stakeholder design throughout the process.

The second avenue for future investigation would be to run similar workshops with other groups of clinicians and people with lived experience of OCD to include a greater breadth of experiences in design ideation. As was raised several times throughout the workshops, OCD can be a highly idiosyncratic experience for individuals, and some ideas that were raised were not seen as universally applicable. Indeed, there is evidence that cultural and social differences between people affect experiences of using services for OCD [56]. Replications of this study with different ethnic and social groups could help understand how the ideas proposed in this workshop would be perceived in different care settings.

XSL•FO
RenderX

This study demonstrates the power of participatory design in applying VR to the treatment of OCD. With the ideas raised by participants, there are a large number of avenues and novel methods for researchers to test in applying VR to support and enhance the treatment of OCD. It is envisaged that refining and applying these ideas, the theoretically effective treatments already available might be more consistently applied at a high quality, with support from such VR apps.

Acknowledgments

We are grateful to Sophie Minns for her assistance in facilitating several of the workshops.

The authors attest that there was no use of generative AI technology in the generation of any text, figures, or other informational content of this manuscript.

JFAM is a co-corresponding author, and they can be contacted at the following email address: jfam20@bath.ac.uk.

Funding

This original research was part of a self-funded PhD by the author MC. Funding to compensate participants' time was provided by the University of Bath's Department of Psychology Public Engagement Fund.

Data Availability

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

Authors' Contributions

Conceptualization: MC, JFAM, BP, DJF, DSF
Data curation: MC, MK
Formal analysis: MC, MK
Funding acquisition: MC
Investigation: MC, JFAM, MK
Methodology: MC, JFAM, DJF
Project administration: MC
Resources: MC, JFAM, MK
Supervision: JFAM, BP, DJF, DSF
Writing – original draft: MC
Writing – review & editing: MC, JFAM, BP, DJF, MK, DSF

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional prompts and follow-up questions used in Workshop 1.
[DOCX File, 15 KB - xr_v3i1e92410_app1.docx]

Multimedia Appendix 2
Brief educational presentation on virtual reality.
[DOCX File, 4585 KB - xr_v3i1e92410_app2.docx]

Multimedia Appendix 3
Persona development worksheet.
[DOCX File, 14 KB - xr_v3i1e92410_app3.docx]

References

1. Fawcett EJ, Power H, Fawcett JM. Women are at greater risk of OCD than men: a meta-analytic review of OCD prevalence worldwide. J Clin Psychiatry 2020 Jun 23;81(4):19r13085. [doi: 10.4088/JCP.19r13085] [Medline: 32603559]
2. Ruscio AM, Stein DJ, Chiu WT, Kessler RC. The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. Mol Psychiatry 2010 Jan;15(1):53-63. [doi: 10.1038/mp.2008.94] [Medline: 18725912]
3. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Publishing; 2013. [doi: 10.1176/appi.books.9780890425596]

XSL•FO
RenderX

4. Wilson KA, Chambless DL. Inflated perceptions of responsibility and obsessive-compulsive symptoms. Behav Res Ther 1999 Apr;37(4):325-335. [doi: 10.1016/s0005-7967(98)00146-6] [Medline: 10204278]
5. Purdon C. Obsessive-compulsive disorder. In: Wenzel A, editor. Handbook of Cognitive Behavioral Therapy: Applications: American Psychological Association; 2021:67-97. [doi: 10.1037/0000219-003]
6. Williams MT, Wetterneck CT. Sexual Obsessions in Obsessive-Compulsive Disorder: A Step-by-Step, Definitive Guide to Understanding, Diagnosis, and Treatment: Oxford University Press; 2019. [doi: 10.1093/med-psych/9780190624798.001.0001]
7. Huppert JD, Simpson HB, Nissenson KJ, Liebowitz MR, Foa EB. Quality of life and functional impairment in obsessive-compulsive disorder: a comparison of patients with and without comorbidity, patients in remission, and healthy controls. Depress Anxiety 2009;26(1):39-45. [doi: 10.1002/da.20506] [Medline: 18800368]
8. Albert U, De Ronchi D, Maina G, Pompili M. Suicide risk in obsessive-compulsive disorder and exploration of risk factors: a systematic review. Curr Neuropharmacol 2019;17(8):681-696. [doi: 10.2174/1570159X16666180620155941] [Medline: 29929465]
9. Torres AR, Prince MJ, Bebbington PE, et al. Obsessive-compulsive disorder: prevalence, comorbidity, impact, and help-seeking in the British National Psychiatric Morbidity Survey of 2000. Am J Psychiatry 2006 Nov;163(11):1978-1985. [doi: 10.1176/ajp.2006.163.11.1978] [Medline: 17074950]
10. Semeniuc S, Sterie MC, Soponaru C, Butnaru S, Gavrilovici O. Therapists' problematic experiences when working with obsessive-compulsive disorder: a qualitative investigation of schema modes, mode cycles, and strategies to return to healthy adult mode. Front Psychiatry 2023;14:1157553. [doi: 10.3389/fpsyt.2023.1157553] [Medline: 38161727]
11. Öst LG, Havnen A, Hansen B, Kvale G. Cognitive behavioral treatments of obsessive-compulsive disorder. A systematic review and meta-analysis of studies published 1993-2014. Clin Psychol Rev 2015 Aug;40:156-169. [doi: 10.1016/j.cpr.2015.06.003] [Medline: 26117062]
12. Öst LG, Enebrink P, Finnes A, et al. Cognitive behavior therapy for obsessive-compulsive disorder in routine clinical care: a systematic review and meta-analysis. Behav Res Ther 2022 Dec;159:104170. [doi: 10.1016/j.brat.2022.104170] [Medline: 36302283]
13. Leeuwerik T, Cavanagh K, Strauss C. Patient adherence to cognitive behavioural therapy for obsessive-compulsive disorder: a systematic review and meta-analysis. J Anxiety Disord 2019 Dec;68(102135):102135. [doi: 10.1016/j.janxdis.2019.102135] [Medline: 31704633]
14. Stobie B, Taylor T, Quigley A, Ewing S, Salkovskis PM. "Contents May Vary": a pilot study of treatment histories of OCD patients. Behav Cogn Psychother 2007 May;35(3):273-282. [doi: 10.1017/S135246580700358X]
15. Külz AK, Hassenpflug K, Riemann D, Linster HW, Dornberg M, Voderholzer U. Psychotherapeutic care in OCD outpatients--results from an anonymous therapist survey. Psychother Psychosom Med Psychol 2010 Jun;60(6):194-201. [doi: 10.1055/s-0029-1202837] [Medline: 19277945]
16. Gillihan SJ, Williams MT, Malcoun E, Yadin E, Foa EB. Common pitfalls in Exposure and Response Prevention (EX/RP) for OCD. J Obsessive Compuls Relat Disord 2012 Oct 1;1(4):251-257. [doi: 10.1016/j.jocrd.2012.05.002] [Medline: 22924159]
17. Moritz S, Külz A, Voderholzer U, Hillebrand T, McKay D, Jelinek L. "Phobie à deux" and other reasons why clinicians do not apply exposure with response prevention in patients with obsessive-compulsive disorder. Cogn Behav Ther 2019 Mar;48(2):162-176. [doi: 10.1080/16506073.2018.1494750] [Medline: 30064295]
18. Bush J. Viability of virtual reality exposure therapy as a treatment alternative. Comput Human Behav 2008 May;24(3):1032-1040. [doi: 10.1016/j.chb.2007.03.006]
19. Rauschnabel PA, Felix R, Hinsch C, Shahab H, Alt F. What is XR? Towards a framework for augmented and virtual reality. Comput Human Behav 2022 Aug;133:107289. [doi: 10.1016/j.chb.2022.107289]
20. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. Front Robot AI 2016;3. [doi: 10.3389/frobt.2016.00074]
21. Gega L. The virtues of virtual reality in exposure therapy. Br J Psychiatry 2017 Apr;210(4):245-246. [doi: 10.1192/bjp.bp.116.193300] [Medline: 28373222]
22. Leeuwerik T, Caradonna G, Cavanagh K, et al. A thematic analysis of barriers and facilitators to participant engagement in group exposure and response prevention therapy for obsessive–compulsive disorder. Psychol Psychother 2023 Mar;96(1):129-147. [doi: 10.1111/papt.12430] [Medline: 36302721]
23. Rizzo AS, Koenig ST. Is clinical virtual reality ready for primetime? Neuropsychology 2017 Nov;31(8):877-899. [doi: 10.1037/neu0000405] [Medline: 29376669]
24. Knopp-Hoffer J, Knowles S, Bower P, Lovell K, Bee PE. "One man's medicine is another man's poison": a qualitative study of user perspectives on low intensity interventions for Obsessive-Compulsive Disorder (OCD). BMC Health Serv Res 2016 May 18;16(1):188. [doi: 10.1186/s12913-016-1433-3] [Medline: 27194033]
25. Colman M, Millar J, Patil B, et al. A systematic review and narrative synthesis of the use and effectiveness of extended reality technology in the assessment, treatment and study of obsessive compulsive disorder. J Obsessive Compuls Relat Disord 2024 Jul;42:100893. [doi: 10.1016/j.jocrd.2024.100893]

XSL•FO
RenderX

26. Cooper D, Champion SM, Stavropoulos L, Grisham JR. How technology can enhance treatment: a scoping review of clinical interventions for anxiety and obsessive-compulsive spectrum disorders. Br J Clin Psychol 2022 Jan;61 Suppl 1(S1):8-30. [doi: 10.1111/bjc.12279] [Medline: 33570762]
27. Birckhead B, Khalil C, Liu X, et al. Recommendations for methodology of virtual reality clinical trials in health care by an International Working Group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973. [doi: 10.2196/11973] [Medline: 30702436]
28. Teng MQ, Gordon E. Therapeutic virtual reality in prison: participatory design with incarcerated women. New Media Soc 2021 Aug;23(8):2210-2229. [doi: 10.1177/1461444821993131]
29. Millar JFA, Higson-Sweeney N, Jenkins TA, Waites EF, Minns S. “Are we genuinely going to have our voices heard?” The experience of co-producing a blended intervention to prevent relapse in obsessive-compulsive disorder: a qualitative study on the perspectives of experts by lived experience. BMC Psychiatry 2024 Dec 18;24(1):906. [doi: 10.1186/s12888-024-06355-1] [Medline: 39696127]
30. 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 Jan 30;17(1):e30. [doi: 10.2196/jmir.4055] [Medline: 25639757]
31. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237. [doi: 10.1016/j.pmrj.2018.07.004]
32. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. Int J Nurs Stud 2024 Jun;154:104705. [doi: 10.1016/j.ijnurstu.2024.104705] [Medline: 38564982]
33. 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] [Medline: 29345485]
34. Briefing notes for researchers: public involvement in NHS, health and social care research. : National Institute for Health and Care Research; 2021 URL: https://www.nihr.ac.uk/briefing-notes-researchers-public-involvement-nhs-health-and-social-care-research [accessed 2023-05-02]
35. Hagen P, Collin P, Metcalf A, Nicholas M, Rahilly K, Swainston N. Participatory design of evidence-based online youth mental health promotion, intervention and treatment. : Young and Well Cooperative Research Centre; 2012 URL: https://www.westernsydney.edu.au/young-and-resilient/intergener8/our-work/shared-media/young-and-well-crc-im-pd-guide.pdf [accessed 2023-02-03]
36. Acocella I. The focus groups in social research: advantages and disadvantages. Qual Quant 2012 Jun;46(4):1125-1136. [doi: 10.1007/s11135-011-9600-4]
37. Queirós A, Faria D, Almeida F. Strengths and limitations of qualitative and quantitative research methods. Eur J Educ Stud 2017:9 [FREE Full text] [doi: 10.46827/ejes.v0i0.1017]
38. Ozkaynak M, Sircar CM, Frye O, Valdez RS. A systematic review of design workshops for health information technologies. Informatics 2021;8(2):34. [doi: 10.3390/informatics8020034]
39. Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mHealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579. [doi: 10.2196/11579] [Medline: 30552080]
40. Freeman D, Haselton P, Freeman J, et al. Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. Lancet Psychiatry 2018 Aug;5(8):625-632. [doi: 10.1016/S2215-0366(18)30226-8] [Medline: 30007519]
41. Barathi SC, Finnegan DJ, Farrow M, et al. Interactive feedforward for improving performance and maintaining intrinsic motivation in VR exergaming. 2018 Apr 21 Presented at: CHI ’18; Apr 21-26, 2018. [doi: 10.1145/3173574.3173982]
42. Maguire M. A framework for user-requirements analysis and development of creative design concepts. : Springer International Publishing; 2022 Presented at: HCI International 2022 - Late Breaking Papers Design, User Experience and Interaction HCII 2022 Lecture Notes in Computer Science; May 26 to Jul 1, 2022 p. 97-113. [doi: 10.1007/978-3-031-17615-9_7]
43. Richards KAR, Hemphill MA. A practical guide to collaborative qualitative data analysis. J Teach Phys Educ 2018 Apr;37(2):225-231. [doi: 10.1123/jtpe.2017-0084]
44. Waller G. Evidence-based treatment and therapist drift. Behav Res Ther 2009 Feb;47(2):119-127. [doi: 10.1016/j.brat.2008.10.018] [Medline: 19036354]
45. Wade TD, Waller G. Ten generic competences to improve outcomes of cognitive behaviour therapy: evidence, postulated processes, and clinical implications. Behav Res Ther 2025 Oct;193:104826. [doi: 10.1016/j.brat.2025.104826] [Medline: 40749419]
46. Waller G, Turner H. Therapist drift redux: why well-meaning clinicians fail to deliver evidence-based therapy, and how to get back on track. Behav Res Ther 2016 Feb;77:129-137. [doi: 10.1016/j.brat.2015.12.005] [Medline: 26752326]
47. Pimentel D, Kalyanaraman S. Customizing your demons: anxiety reduction via anthropomorphizing and destroying an “Anxiety Avatar”. Front Psychol 2020;11:566682. [doi: 10.3389/fpsyg.2020.566682] [Medline: 33343445]

XSL•FO
RenderX

48. Cullen AJ, Dowling NL, Segrave R, Carter A, Yücel M. Exposure therapy in a virtual environment: validation in obsessive compulsive disorder. J Anxiety Disord 2021 May;80:102404. [doi: 10.1016/j.janxdis.2021.102404] [Medline: 33894550]
49. Calancea A. Metaphors as cognitive and therapeutic tools: an interdisciplinary analysis. 2025 Presented at: Scientific Conference with International Participation "The contemporary issues of the socio-humanistic sciences"; Mar 13-14, 2025; Chi inău, Moldova p. 85-93. [doi: 10.54481/pcss2025.08]
50. Exner C, Kleiman A, Haberkamp A, Hansmeier J, Milde C, Glombiewski JA. Metacognitive therapy versus exposure and response prevention for obsessive-compulsive disorder - a non-inferiority randomized controlled trial. J Anxiety Disord 2024 Jun;104:102873. [doi: 10.1016/j.janxdis.2024.102873] [Medline: 38729024]
51. Bream V, Challacombe F, Palmer A, Salkovskis P. Cognitive Behaviour Therapy for OCD, 1st edition: Oxford University Press; 2017. [doi: 10.1093/med:psych/9780198703268.001.0001]
52. Gorini A, Riva G. The potential of virtual reality as anxiety management tool: a randomized controlled study in a sample of patients affected by generalized anxiety disorder. Trials 2008 May 5;9(25):25. [doi: 10.1186/1745-6215-9-25] [Medline: 18457580]
53. Wiederhold BK, Bouchard S. Generalized anxiety disorder and obsessive–compulsive disorder: efficacy and the development of virtual environments. In: Wiederhold BK, Bouchard S, editors. Advances in Virtual Reality and Anxiety Disorders: Springer US; 2014:235-257. [doi: 10.1007/978-1-4899-8023-6_11]
54. Mathes BM, Morabito DM, Schmidt NB. Epidemiological and clinical gender differences in OCD. Curr Psychiatry Rep 2019 Apr 23;21(5):36. [doi: 10.1007/s11920-019-1015-2] [Medline: 31016410]
55. Stengler K, Olbrich S, Heider D, Dietrich S, Riedel-Heller S, Jahn I. Mental health treatment seeking among patients with OCD: impact of age of onset. Soc Psychiatry Psychiatr Epidemiol 2013 May;48(5):813-819. [doi: 10.1007/s00127-012-0544-3] [Medline: 22763495]
56. Kolvenbach S, Fernández de la Cruz L, Mataix-Cols D, Patel N, Jassi A. Perceived treatment barriers and experiences in the use of services for obsessive-compulsive disorder across different ethnic groups: a thematic analysis. Child Adolesc Ment Health 2018 May;23(2):99-106. [doi: 10.1111/camh.12197] [Medline: 32677337]
57. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exerc Health 2021 Mar 4;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]

Abbreviations

CBT: cognitive behavioral therapy
CQA: collaborative qualitative analysis
ERP: exposure and response prevention
JARS-Qual: Journal Article Reporting Standards for Qualitative Primary Research in Psychology
MRC: Medical Research Council
OCD: obsessive-compulsive disorder
PBA: person-based approach
RQ: research question
VE: virtual environment
VR: virtual reality

Edited by I Steenstra; submitted 29.Jan.2026; peer-reviewed by K Kaur; revised version received 03.Jul.2026; accepted 03.Jul.2026; published 20.Aug.2026.

Please cite as:
Colman M, Millar JFA, Patil B, Finnegan DJ, Kristoffersen M, Stanton Fraser D
Novel Virtual Reality Methods for the Treatment of Obsessive-Compulsive Disorder Using a User-Centered Design Approach: Focus Group Study
JMIR XR Spatial Comput 2026;3:e92410
URL: https://xr.jmir.org/2026/1/e92410
doi:10.2196/92410

© Michael Colman, Josie F A Millar, Bhagyashree Patil, Daniel J Finnegan, Mhairi Kristoffersen, Danae Stanton Fraser. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 20.Aug.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Metaverse-Based Psychiatric Consultation for Youths With Mental Health Conditions: Qualitative Descriptive Feasibility Study

Mio Ishii[1], MD, PhD; Junichi Fujita[1,2], MD, PhD; Nao Toyohara[1,2], MD; Keiko Ide[1], MD, PhD; Tomoko Moroga[3], RN; Mizuho Takayama[3], BA; Takeshi Asami[1], MD, PhD; Tomoyuki Miyazaki[2], MD, PhD

[1]Department of Psychiatry, Yokohama City University School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama, Japan

[2]Center for Promotion of Research and Industry-Academic Collaboration, Yokohama City University, Yokohama, Japan

[3]Department of Child Psychiatry, Yokohama City University Hospital, Yokohama, Kanagawa, Japan

Corresponding Author:
Mio Ishii, MD, PhD
Department of Psychiatry, Yokohama City University School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama, Japan

Abstract

Background: Youth mental health is a global public health priority, with rising rates of anxiety and depression, particularly after the COVID-19 pandemic. Despite the early onset and substantial burden of mental disorders in this age group, young people are less likely than adults to seek professional help and face barriers such as workforce shortages, stigma, and low mental health literacy. Although efforts such as outreach initiatives and school-based programs have been implemented, innovative and scalable solutions remain limited. Digital technologies, including the metaverse, may offer flexible and stigma-reducing approaches to mental health care; however, evidence regarding their real-world feasibility and acceptability is scarce.

Objective: This study investigated the feasibility and user experience of metaverse-based psychiatric consultations for young people with mental health conditions.

Methods: We conducted a qualitative descriptive feasibility study at a single academic institution in Yokohama, Japan, between July and November 2023. A total of 26 participants aged 16 to 25 years (mean age 19.9, SD 2.4; 15 male, 7 female, 3 nonbinary, 1 no response) who self-identified as having mental health concerns were recruited from local psychiatric clinics, schools, universities, and social media. Reported concerns included anxiety, depressive symptoms, and autism spectrum disorder traits. Participants completed a 30- to 40-minute one-on-one metaverse-based consultation with a psychiatrist using avatars in a virtual reality environment, followed by semistructured interviews exploring feasibility, usability, and user perceptions. Data were analyzed using thematic analysis, and data collection continued until no substantially new themes emerged. Reporting followed the APA Journal Article Reporting Standards for qualitative research.

Results: All participants completed the study without any adverse psychological events. Five participants experienced minor, transient physical discomfort (eg, headaches and virtual reality–related sickness), which resolved without medical intervention. Thematic analysis identified 3 primary domains: perceived psychological safety through avatar-mediated interaction, enhanced spatial presence facilitating rapport, and increased autonomy within the virtual environment. Metaverse consultations were perceived as particularly beneficial for individuals experiencing interpersonal anxiety, sensory sensitivities (including autism spectrum disorder traits), difficulty leaving home due to psychiatric conditions, psychological resistance to traditional psychiatric settings, or discomfort with physical self-presentation.

Conclusions: This qualitative descriptive feasibility study provides preliminary evidence that metaverse-based psychiatric consultations are a feasible and acceptable approach for supporting young people with mental health conditions. Unlike conventional telepsychiatry based on videoconferencing, the use of avatar-mediated interaction and immersive virtual environments may reduce psychological barriers related to self-presentation, stigma, and interpersonal anxiety for specific subgroups of youth. These findings suggest that metaverse-based consultations can be effectively integrated into clinical pathways as a complementary, “low-threshold” access point within stepped or hybrid care models, ultimately bridging the gap between initial help-seeking and formal psychiatric treatment.

(JMIR XR Spatial Comput 2026;3:e83688) doi:10.2196/83688

KEYWORDS

mental health; metaverse; youth mental health; virtual reality; telepsychiatry

XSL•FO
RenderX

Introduction

Youth mental health is a critical global issue that remains complex and challenging. Reports indicate that 1 in 5 young people experiences anxiety, and 1 in 4 faces depression due to factors such as poverty and shifts in lifestyle patterns [1,2]. This prevalence doubled before and after the COVID-19 pandemic [3]. With more than 25% of mental disorders beginning before the age of 25 years [4] and the significant socioeconomic impact of these conditions [5], the need for effective mental health care for young people is clear. Despite this, young people are less likely to seek professional support for mental health problems than older age groups and often struggle to access timely and high-quality mental health services (MHSs) [6-8]. Several factors contribute to this gap between demand and supply, including physical barriers such as the shortage and geographic maldistribution of child and adolescent psychiatrists, as well as psychological barriers such as stigma and low mental health literacy [9-11]. Although initiatives, including school-based psychoeducation, peer training, and outreach programs, have been implemented to mitigate these barriers, a breakthrough solution remains elusive [6,12-14], underscoring the need for further research.

As an innovative response to these challenges, information and communication technology (ICT) has become increasingly central to youth mental health care. McGorry et al [9] have called for a redesign of youth MHSs and conceptualized youth psychiatry as a distinct discipline, emphasizing early intervention, co-designed youth-friendly services, integration of digital technology with human support, and an extended service boundary up to age 25. Consistent with this framework, ICT-based services are expected to enhance accessibility, reduce stigma, and provide flexible, developmentally sensitive support for late adolescents and young adults, whose daily communication and help-seeking increasingly occur online [9,15]. However, recent systematic reviews caution that the real-world impact of digital youth MHSs may be constrained by usability burdens, uneven engagement, and variable acceptability across diverse youth populations [16]. Accordingly, feasibility studies in novel, youth-centered digital environments are a critical step toward identifying formats that young people can realistically and sustainably use.

One emerging ICT modality is the metaverse, broadly defined as a persistent, shared virtual space in which users interact through avatars within immersive (eg, virtual reality [VR]/mixed reality) or nonimmersive environments [17,18]. Unlike stand-alone VR or augmented reality applications that provide isolated, session-based experiences, metaverse environments are typically continuous, socially interactive spaces in which users can maintain identity, customize self-representation, and engage in repeated or ongoing interactions over time. These environments may include synchronous communication, spatialized presence, and customizable virtual settings, allowing users to experience a sense of copresence and embodiment beyond conventional video-based teleconferencing [19].

In youth mental health contexts, these features may be particularly relevant. Recent reviews of immersive and metaverse-based digital therapies suggest potential experiential advantages over conventional telepsychiatry; however, robust empirical evidence remains limited and heterogeneous [18]. Avatar-mediated interaction and the option to engage from a private physical location may lower the psychological threshold for help-seeking and support continued participation, as suggested by early metaverse-based or VR-enabled support initiatives [17,18,20]. In Japan, for example, emerging initiatives in some cities use VR-based support services for individuals with social withdrawal (Hikikomori) [20], exploring approaches to reduce stigma and facilitate access to support. The metaverse could extend these efforts internationally by offering an alternative environment in which young people can seek help discreetly and interactively while maintaining a sense of autonomy.

However, empirical evidence on the feasibility, acceptability, and ethical implementation of metaverse-based MHSs, such as psychiatric consultations for young people, remains limited, particularly in real-world clinical contexts. This study, therefore, aims to assess the feasibility of metaverse-based psychiatric consultations tailored for young people with mental health conditions, with a focus on user experience, usability, and implementation considerations. The findings will guide the development and implementation of future youth MHS, addressing key barriers to access and engagement.

Methods

Research Design Overview

This study employed a qualitative descriptive study to assess the feasibility of metaverse-based psychiatric consultations for youths with mental health conditions. Semistructured individual interviews were conducted following the consultations to explore participants' experiences, perceived usability, and acceptability of the intervention.

Study Participants or Data Sources

Participants were young people aged 16 to 25 years who self-identified as experiencing mental health conditions. The research team consisted of psychiatrists and mental health professionals with clinical experience in youth mental health care, as well as researchers with expertise in digital mental health and qualitative research. These backgrounds informed the study design, particularly the focus on feasibility, acceptability, and ethical considerations of metaverse-based psychiatric consultations.

Some participants were recruited from clinical settings where members of the research team were involved in their care. In such cases, the clinicians who had an existing therapeutic relationship with a participant did not conduct the postconsultation interviews and were not involved in the qualitative data analysis for that participant. This role separation was implemented to minimize potential power imbalances, social desirability bias, and undue influence on participants' responses.

For participants recruited through educational institutions or social media, no prior relationship with the research team existed before study participation. All participants were informed that

XSL•FO
RenderX

participation was voluntary, would not affect their clinical care or educational standing, and that they could withdraw at any time without consequences. These measures were taken to ensure ethical integrity and to support open and honest expression during the research process.

Participant Recruitment

Recruitment Strategy

Participants were recruited between July and November 2023 through multiple channels, including flyer distribution at Yokohama University Hospital, nearby clinics, local universities, high schools, community support organizations, and announcements on social media platforms. Recruitment materials briefly described the study as an opportunity to experience a metaverse-based psychiatric consultation and to share feedback about the experience. Interested individuals contacted the research team directly and were screened for eligibility.

All participants received written and verbal explanations of the study procedures, potential risks, and the voluntary nature of participation. Written informed consent was obtained prior to study participation; for participants younger than 18 years, consent was also obtained from legal guardians in accordance with institutional review board requirements. Participants were compensated with a gift card valued at approximately US $35 (¥5000).

The target sample size was determined based on the exploratory aims of a qualitative descriptive feasibility design rather than statistical power considerations. A total of 33 individuals expressed interest in participation. Of these, 1 individual did not meet the inclusion criteria, 5 were lost to contact prior to providing informed consent, and 1 was unable to participate due to their condition on the scheduled day. Consequently, 26 participants completed the metaverse-based consultation and postconsultation interview.

Recruitment and data collection were concluded when thematic convergence was observed during qualitative analysis, and no substantially new themes emerged in later interviews, indicating that the sample size was sufficient for the aims of this feasibility study.

Inclusion and Exclusion Criteria

Eligible participants were individuals aged 16 to 25 years at the time of consent, who were aware of their mental health conditions and able to provide written informed consent in Japanese. Individuals currently receiving treatment at a psychiatric medical institution without permission from their primary physician were excluded from participation.

Data Collection

Setting

The study was conducted in a meeting room on the Yokohama City University campus.

Baseline Assessment

As a baseline assessment, the participants completed a self-administered questionnaire consisting of 4 parts: (1) mental and physical conditions, (2) living conditions over the past 3 months, (3) internet usage, and (4) perceptions of mental health and psychiatry.

Metaverse Environment Setup and Technical Support

After the baseline assessment, the participants entered the metaverse environment under the guidance of technical staff, wearing VR goggles and using a controller. Since the participants were unfamiliar with metaverse operations, the technical staff provided detailed instructions for fitting the goggles, launching the application, and logging in.

For this study, experimental worlds and avatars were set up on commercially available metaverse applications, VRChat and Workrooms. VRChat is a social VR platform that allows users to create, share, and explore user-generated virtual worlds through customizable avatars, fostering social interaction and creative expression in immersive settings. In contrast, Workrooms is a virtual collaboration tool designed for professional and educational use, enabling users to hold virtual meetings in shared spaces with avatars and facilitating interactive discussions, presentations, and collaborative tasks. These applications provided structured and flexible virtual spaces suitable for the purposes of this study.

The virtual environments included 3 distinct spaces: an outer space environment, a hospital examination room, and a meeting room with selectable views of a beach or mountains. The avatars included both human and nonhuman fictional characters, providing participants with diverse ways of interacting and expressing themselves. Figure 1 illustrates one of the settings and avatars used in the study.

XSL•FO
RenderX

Figure 1. Overview of the metaverse-based psychiatric consultation process for youth with mental health conditions (N=26), conducted at Yokohama City University Hospital between July 2023 and November 2023. The image illustrates a participant and a clinician wearing head-mounted displays and engaging in interaction via avatars.

Metaverse Psychiatric Consultation

The metaverse consultations were conducted one-on-one in private virtual rooms with a psychiatrist, who also wore VR goggles and appeared as an avatar from a remote location. The psychiatrists included both male and female professionals with an average of 19 years of clinical experience.

Each consultation lasted 30 to 40 minutes and followed a semistructured framework resembling a standard initial psychiatric consultation. The consultation consisted of 5 phases. First, the psychiatrist introduced themselves via their avatar and explained the goals and structure of the session, allowing time to establish rapport and ensure that the participant felt comfortable in the virtual environment. Second, the psychiatrist reviewed the participant's medical and life history, including prior diagnoses, treatments, daily routines, relationships, and environmental stressors. Third, the session focused on symptom observation and assessment, during which participants described their current mental health concerns while the psychiatrist observed verbal and nonverbal cues through avatar-mediated interaction. Fourth, tailored guidance and recommendations were provided, including coping strategies, lifestyle adjustments, and referrals to additional mental health resources when appropriate. Finally, the session concluded with a summary of the discussion and an opportunity for the participant to ask questions or provide feedback.

The consultations were designed to balance structure and flexibility, providing participants with a supportive environment to explore their mental health concerns in a novel virtual setting. During the session, the participants navigated the 3 distinct spaces within the metaverse environment alongside the psychiatrist and interacted via various types of avatars. Technical staff members were available to assist as needed, ensuring a seamless and accessible experience.

Postconsultation Interviews

After the metaverse consultation, participants were individually interviewed to explore their experiences and perceptions of the metaverse consultation. The semistructured interview guide used in this study was developed specifically for this project. An English version of the interview form is provided in Multimedia Appendix 1 for reference. The interviews explored several key areas, including the usability of the metaverse devices, any physical or psychological discomfort experienced, acceptance compared with face-to-face consultations, perceptions of individuals or situations in which metaverse consultations may be beneficial, and implementation-related requirements or concerns.

Following the participant interviews, the psychiatrists who conducted the consultations, along with members of the research team—including psychiatric social workers and a nurse—engaged in structured discussions regarding usability, expectations, and implementation challenges. These discussions addressed the functionality and user experience of the metaverse platform, technical considerations, perceived differences between metaverse and face-to-face consultations, compatibility with participant characteristics and symptom profiles, and additional observations noted during the study.

XSL•FO
RenderX

The reporting of this study conforms to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines (Checklist 1) [21].

Data Analysis

Qualitative Analysis Strategy

The interview data were analyzed via thematic analysis following established qualitative research methodologies. First, all interviews were transcribed verbatim to ensure accuracy. The transcripts were then reviewed multiple times to facilitate familiarity with the data and identify potential patterns.

All interview transcripts were reviewed in full by the first author (MI), who performed line-by-line coding to identify meaningful units of text. Codes were generated inductively from the data and iteratively refined in consultation with 2 coauthors (JF and NT) to enhance credibility. Codes were clustered into categories, and higher-order themes were developed through repeated team discussions until consensus was reached.

Initial coding was conducted by systematically highlighting recurring phrases, concepts, and key statements. These codes were grouped into broader categories that represented emerging themes. To enhance rigor, the themes were independently reviewed and refined by multiple researchers to ensure consistency and reduce bias.

The analysis focused on 3 primary domains: the usability of the metaverse platform, the acceptance of metaverse consultations in comparison with face-to-face interactions, and the identification of participant characteristics or conditions for which metaverse consultations may offer particular advantages, such as social anxiety or sensory sensitivities.

Methodological Integrity

To ensure trustworthiness, we maintained an audit trail of coding decisions, documenting how codes were generated, refined, and grouped into themes. Triangulation was achieved through iterative discussions among the research team, allowing for verification and consensus in data interpretation. Representative quotations were included in the *Results* section and labeled with participant numbers to illustrate key themes and to provide transparency between the participants' voices and the researchers' interpretations.

Data collection was concluded when data saturation was reached, defined as the point at which no new themes or insights emerged from the interviews. No substantially new themes emerged in the latter part of the interviews. Thus, we consider the sample size sufficient for the exploratory aims of this qualitative feasibility study.

This study was conducted and reported in accordance with the APA Journal Article Reporting Standards for qualitative research [22].

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and the Ethical Guidelines for Life Science and Medical Research Involving Human Subjects, established by the Ministry of Education, Culture, Sports, Science and Technology and the Ministry of Health, Labour and Welfare of Japan (partially revised in 2023). The study protocol was approved by the institutional review board of Yokohama City University, Japan (approval number F230400055). Written informed consent was obtained from all participants prior to their inclusion in the study; for participants younger than 18 years, additional consent was obtained from their legal guardians along with participant assent, where applicable. To ensure privacy and confidentiality, all study data were deidentified during handling and analysis, and interview transcripts were meticulously reviewed to remove any personal identifiers. Participants were compensated for their time with a gift card valued at approximately US $35 (¥5000). Regarding the visual materials, this manuscript includes images of staff members demonstrating the metaverse consultation environment. All identifiable individuals in these images provided explicit written informed consent for the publication of their likeness, and the relevant consent forms are available for verification upon request.

Since this study was conducted as an observational study, clinical trial registration was not applicable.

Results

The results presented below were derived through thematic analysis of postconsultation interviews, identifying common patterns across participants' responses.

Baseline Characteristics

A total of 26 young people were assigned from a pool of 33 volunteers. One individual was excluded for not meeting the inclusion criteria, 5 lost contacts before providing informed consent, and 1 could not participate because of his condition on the day. Table 1 summarizes the baseline characteristics, while Table 2 presents the results of the baseline assessments of the participants. The mental and physical conditions of the participants varied, with many expressing difficulties in social interactions, feelings of isolation, and discomfort in seeking support from others. Most participants reported frequent internet use, although their levels of social engagement and self-care practices differed. All participants owned internet-enabled devices, and most used the internet daily. However, some participants acknowledged neglecting essential tasks due to internet use, whereas others reported difficulties in limiting their internet usage. The participants' initial impressions of the metaverse also varied. While only a minority had extensive knowledge or prior experience with the metaverse, many expressed curiosity. Their views on associating with individuals who had a history of psychiatric treatment ranged widely, with some expressing comfort and acceptance, whereas others showed apprehension or stigma toward mental health issues.

XSL•FO
RenderX

Table . Baseline demographic and clinical characteristics of youth participants with mental health conditions (N=26) in an observational feasibility study of metaverse-based psychiatric consultation conducted at Yokohama City University between July 2023 and November 2023.

Characteristic	Description
Age (y), mean (range)	19.9 (16-25)
Gender	
Male	15
Female	7
Nonbinary	3
No response	1
Recruitment sources	Yokohama City University Health Management Center, medical students and junior residents, outpatient of University Hospital and neighborhood clinics, participants from other survey of our research team, neighborhood private high school, Facebook
History of psychiatric treatment	
Yes	22
No	4

XSL•FO
RenderX

Table . Summary of baseline clinical assessments, including psychological symptoms and functional status, of youth participants (N=26) prior to the metaverse-based psychiatric consultation.

Aspect of assessment	Summary of findings
1. Physical and psychological condition	
Perceived health issues	Most participants (21/26, 80%) indicated concerns about their physical or mental health, with common issues including insomnia, social anxiety, general anxiety, depression, learning disabilities, difficulty speaking with others, and lack of motivation.
Comfort being in social situations	Nearly two-thirds (18/26, 68%) reported discomfort being in front of others, and 64% (17/26) felt uncomfortable talking with strangers, indicating significant social anxiety.
Help-seeking behavior	Approximately 65% (17/26) reported that they seek advice when facing problems, primarily consulting family, friends, teachers, or medical staff.
Sense of social connection	Mixed responses: 50% (13/26) reported sometimes feeling a lack of social connections, while others felt relatively neutral or disagreed with feelings of isolation.
Feelings of exclusion and isolation	Most participants (20/26, 77%) felt neither excluded nor isolated, suggesting a generally stable sense of belonging, despite some individual cases of feeling isolated or left out.
2. Living condition of the past 3 months	
Frequency of going out	Most participants went out frequently, with 73% (19/26) going out at least once a week, while 27% (7/26) went out less frequently.
Conversations with family members	The majority (23/26, 88%) reported daily conversations with family, indicating strong familial interaction.
Conversations with nonfamily members	Over half (20/26, 77%) had conversations with nonfamily members at least weekly, suggesting moderate social interaction outside the family.
Ability to maintain a regular routine	Only 27% (7/26) of participants maintained a regular routine daily or several times per week, with the remainder struggling to do so consistently.
Sufficient sleep	Approximately half (13/26, 50%) reported receiving sufficient sleep almost daily or several times per week, while others reported inconsistent sleep.
Frequency of bathing or showering	Nearly all participants (24/26, 92%) bathed or showered daily, indicating consistent personal hygiene practices.
Concentration on work or studies	Only 27% (7/26) could concentrate on work or studies consistently (daily or multiple times per week), with others facing challenges in concentration.
Engagement in nonwork/study activities	Most participants (16/26, 61%) engaged in hobbies, sports, or social activities at least once a week, while others were less active.
3. Internet usage	
Access to internet-enabled devices	All participants (26/26, 100%) owned internet-enabled devices, such as computers, smartphones, or gaming consoles.
Frequency of internet use	Nearly all participants reported daily internet use, indicating high internet engagement.
Neglect of responsibilities due to internet use	A majority (19/26, 73%) sometimes or always neglected responsibilities due to internet use, with only one participant reporting no impact.
Making new connections online	Around half (12/26, 46%) reported occasionally or frequently making new connections online, while 27% (7/26) never made new connections through the internet.
Difficulty reducing internet use	Nearly half (12/26, 46%) indicated difficulty reducing internet use, with the remainder having little to no issues in managing their usage.
Familiarity with the metaverse	Knowledge about the metaverse varied, with 46% (12/26) reporting some familiarity, while the rest indicated limited or no knowledge.
Experience with the metaverse	Only a small portion (6/26, 23%) had direct experience with the metaverse, indicating it was relatively new to most participants.
4. Attitudes toward mental health and psychiatry	

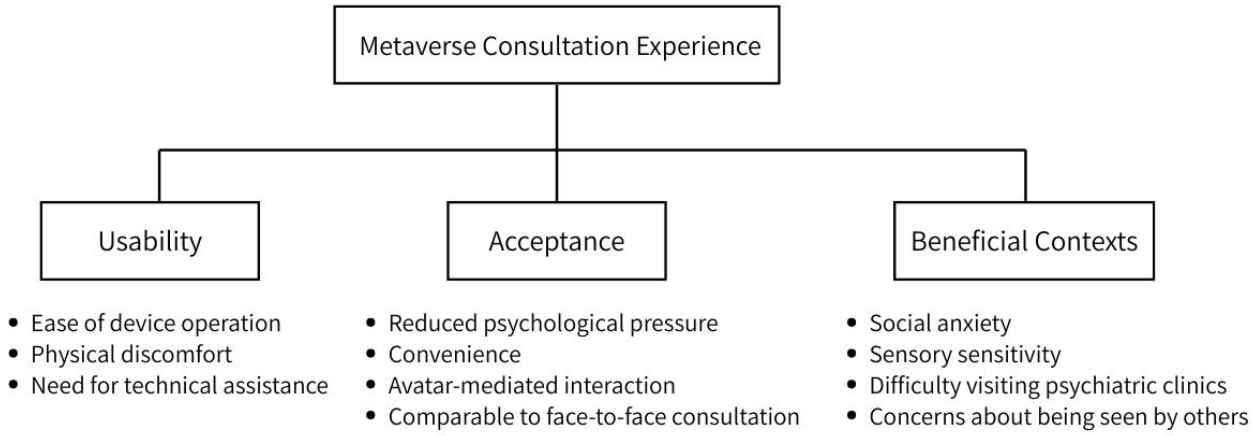
XSL•FO
RenderX

Aspect of assessment	Summary of findings
Friendship with individuals who have seen a psychiatrist	Most participants (23/26, 88%) expressed openness, agreeing that they could become close friends with someone who has sought psychiatric care.
Trust in individuals with a history of psychiatric hospitalization	A majority (19/26, 73%) believed that individuals with past psychiatric hospitalization are just as trustworthy as others.
Perception of psychiatric treatment as a personal failing	Most participants (22/26, 85%) disagreed with the notion that seeking psychiatric help is a sign of personal failure.
Reluctance to hire individuals with psychiatric history	Over half (17/26, 65%) felt that many people would be reluctant to hire someone with a history of psychiatric treatment, reflecting a perceived societal stigma.
Willingness to date someone with a history of psychiatric treatment	Around half (14/26, 54%) were neutral or positive about dating someone with a psychiatric history, while 27% (7/26) expressed reluctance.

Interviews

Thematic analysis was used to identify patterns and themes from participants' postconsultation interviews. A thematic map summarizing the relationships among the identified themes is presented in Figure 2.

Figure 2. Thematic map of participants' experiences with metaverse-based psychiatric consultations (N=26) conducted at Yokohama City University Hospital between July 2023 and November 2023. The map illustrates the relationships among the key themes identified through thematic analysis of postconsultation interviews: usability of the metaverse device, acceptance of metaverse consultations, and contexts in which metaverse consultations may be particularly beneficial for young people with mental health conditions.



Usability of the Metaverse Device

Participants expressed a range of views regarding usability. Approximately 46% reported positive experiences, describing the device as intuitive or comparable to gaming interfaces. About 23% expressed neutral views, noting that while manageable for short sessions, prolonged use might be physically demanding. In contrast, 31% reported negative experiences, primarily related to physical discomfort, device fit, or technical complexity.

Participants who reported positive experiences noted that they "quickly got used to the controls" (Participant 10, 19, 20), and found the device "easy to use due to regular use of gaming devices" (Participant 9, 13, 16, 22). Optimism for future accessibility was evident in comments such as "I would like to use it when it becomes easier" (Participant 4) and "more affordable" (Participant 23, 24).

Participants with neutral views indicated that although they adapted to the operations, more complex tasks might require assistance. One participant commented,

> *I feel comfortable for short sessions, but longer sessions might be tough with these devices.* [Participant 4]

Those expressing negative views raised concerns about the device's physical fit and comfort. Nine participants experienced issues such as the goggles feeling "too tight" (Participant 18), "too loose and slipping" (Participant 13, 14), or challenging to fit over glasses. Five participants reported mild physical discomfort, including VR-related discomfort or head heaviness, although no psychological discomfort was noted. A participant remarked, "It might be difficult to connect and operate remotely without staff assistance" (Participant 18), highlighting barriers to independent use.

Acceptance of the Metaverse Consultation

Responses regarding the acceptance of metaverse consultations revealed a range of perspectives.

Thirty-eight percent of participants expressed a preference for metaverse consultations over face-to-face sessions, citing benefits such as reduced psychological pressure and convenience. The comments included, "It's a big help not having to leave home when I'm feeling too unwell to go out"

XSL•FO
RenderX

(Participant 12) and "Meeting a new doctor is less nerve-wracking through an avatar" (Participant 3). Another participant noted,

> *I don't have to endure the intimidating atmosphere of psychiatric institutions or the doctor's imposing presence.* [Participant 19]

Another 38% of participants expressed neutral responses, describing the content and quality of metaverse consultations as comparable to face-to-face consultations. Many noted that their preference would depend on their condition at the time. For example, one participant remarked,

> *I prefer face-to-face if I can go out, but there are times I'd rather use the metaverse.* [Participant 18]

Others suggested that

> *The metaverse would be preferable if issues such as cost, ease of access, and privacy were addressed.* [Participant 24]

The remaining 24% preferred face-to-face consultations, emphasizing the reassurance of observing facial expressions and body language. One participant shared,

> *If I can't see their face, I worry whether they're truly understanding or accurately interpreting what I want to convey.* [Participant 4]

Another added,

> *For personal and serious discussions, I want to see the person face-to-face.* [Participant 8]

Participants also shared insights into specific scenarios where metaverse consultations might be beneficial. One participant stated,

> *It's ideal for people like me who are self-conscious about being seen. I can focus more on the consultation and express what I truly want to say.* [Participant 9]

Another remarked,

> *It's helpful for those who find it challenging to visit psychiatric clinics.* [Participant 1]

Keywords and Concerns for Metaverse Consultations

Participants highlighted several key requirements and concerns regarding implementation. Security and privacy were frequently emphasized, particularly the need for a secure and private consultation environment. The participants requested assurances that psychiatrists understood them accurately. One participant commented,

> *I don't want to show my face, but if the avatar's facial expressions and body movements are synchronized with reality, it feels more reassuring than a phone call.* [Participant 1]

Concerns about information security for private consultations were also raised.

Customizability of avatars and virtual spaces was another recurring theme. Participants frequently emphasized the importance of customizable settings, using terms such as "relaxing," "free from distractions," and "self-selectable." Twelve participants expressed a preference for avatars that felt authentic to them, with one stating,

> *choosing my avatar is a way of expressing myself, and it makes me feel more comfortable in the consultation.* [Participant 10]

Preferences for human avatars (6 participants) and nonhuman characters (6 participants) were evenly split.

Psychiatric Team's Perspective

The psychiatric team provided valuable insights into the potential benefits and challenges of metaverse consultations.

Team members noted that individuals with sensory hypersensitivity, such as those with traits associated with autism spectrum disorder, appeared to focus more effectively in the immersive VR environment. They also reported that individuals with significant social anxiety or tension seemed to relax more quickly, enabling broader and deeper conversations.

At the same time, concerns were raised regarding the suitability of metaverse consultations for participants with pronounced symptoms, as certain clinical cues may be less observable in virtual settings. Suggestions included simplifying metaverse operations for easier access and expanding the range of avatars and room settings to allow for more personalized consultations. They also recommended developing diagnostic support tools to enhance the consultation experience.

Discussion

Principal Findings

This study evaluated the feasibility and user experience of metaverse-based psychiatric consultations for young people with mental health conditions. The findings demonstrated that, in this study, such consultations are technically feasible and safe, with no major adverse events. Participants reported high levels of engagement, and their postexperience feedback highlighted the unique advantages of avatar-mediated interaction. While some physical discomfort was noted, the overall acceptance of metaverse consultations suggests their potential as a complementary modality in youth mental health care.

Interpretations of Findings

Our results identified specific groups of young people who showed a strong affinity for metaverse consultations, including those with social anxiety, sensory sensitivities (eg, autistic traits), difficulty leaving home due to mental illness, concerns about stigma, and discomfort with their physical appearance. For these individuals, avatar-based interaction lowered psychological barriers and provided a sense of safety and autonomy. These findings are consistent with existing literature showing that digital interventions can enhance help-seeking among youth when designed to offer privacy, flexibility, and user control [9,15,20].

The immersive VR environment was particularly appreciated by participants with sensory sensitivities, as one participant remarked,

XSL•FO
RenderX

> *After experiencing the metaverse consultation, I felt I truly understood what a consultation was meant to be for the first time.* [Participant 2]

This finding may be interpreted through the lens of presence theory, which suggests that immersive environments can enhance users' sense of "being there," potentially facilitating emotional engagement and attentional focus [23].

The ability to interact through avatars in a spatialized environment may also support a form of digital embodiment, enabling participants to regulate interpersonal distance and self-presentation more flexibly than in face-to-face or video-based consultations. This capacity to modulate proximity and representation aligns with findings from embodiment research, which suggest that virtual body ownership and perspective-taking can influence affective and social responses [24]. Such mechanisms may partially explain why some participants reported feeling more relaxed and able to engage in deeper conversations within the metaverse setting.

On the other hand, previous studies caution that VR may also induce discomfort in some individuals with autism spectrum disorder [24]. This duality highlights the need for individualized approaches and flexible options.

In terms of accessibility, our findings revealed operational challenges such as the need for technical support, device heaviness, and VR-related discomfort. These barriers align with broader concerns regarding the digital divide and unequal access to advanced immersive technologies, particularly among young people with mental health conditions. Previous research has demonstrated that individuals with mental ill health, including socioeconomically and digitally marginalized youth, may be limited or nonusers of digital interventions despite device access, suggesting that skills, motivation, and confidence can constitute significant barriers to digital engagement [25,26]. Access to high-end VR hardware may require financial, technological, and spatial resources that are not evenly distributed, potentially exacerbating existing disparities in MHS utilization.

From a technology acceptance perspective, perceived ease of use and perceived usefulness are central determinants of user adoption in health technologies [26]. When devices are physically uncomfortable or operationally complex, perceived ease of use may decline, thereby reducing users' willingness to engage with the platform. Therefore, simplifying hardware requirements and offering lighter, cross-platform alternatives may be essential to ensure equitable and scalable implementation of metaverse-based MHSs. This underscores the importance of exploring less resource-intensive alternatives, such as "light metaverse" platforms that are accessible via widely available devices, including smartphones or personal computers. Expanding metaverse options beyond VR goggles could help reach a wider population and improve inclusivity [18].

In interpreting these findings, it is also important to situate metaverse-based consultations in relation to existing digital mental health modalities, particularly video-based teleconferencing, which is currently the most widely used form of telepsychiatry. Teleconferencing has demonstrated effectiveness in increasing access to care and reducing geographical barriers [27,28]. However, because it typically relies on real-time videoconferencing, it retains key interpersonal elements of face-to-face encounters (eg, visual and vocal cues, facial interaction, and eye contact), while also introducing heightened self-presentation and appearance-related attentional demands [29]. In contrast, metaverse-based consultations have been proposed to offer distinct experiential affordances through avatar-mediated interaction and immersive environments [18]. For some participants, these features appeared to reduce psychological barriers associated with being seen, judged, or scrutinized, thereby facilitating disclosure and engagement. At the same time, the use of VR technology introduced additional usability burdens, including physical discomfort, device complexity, and the need for technical support, which are less prominent in conventional teleconferencing. These findings suggest that metaverse-based consultations should not be viewed as a replacement for teleconferencing, but rather as a complementary modality that may be particularly valuable for specific subgroups of young people for whom video-based interaction remains challenging. Future research directly comparing immersive metaverse approaches with standard teleconferencing is needed to clarify relative advantages, limitations, and appropriate indications within stepped or hybrid models of youth mental health care.

Participants also emphasized the value of choice in how they receive mental health support. The concept of choice emerged consistently in participants' responses, highlighting their desire for flexibility and personalization. The ability to choose avatars and participate in consultations from private spaces fostered a greater sense of agency. This aligns with prior research emphasizing the importance of autonomy and co-designed services in engaging youth with mental health needs [6,17,30,31].

Ethical challenges remain. While our study obtained consent in person, widespread adoption of metaverse consultations will require mechanisms to ensure informed consent, protect privacy, and uphold safety in fully virtual settings [32]. Addressing these ethical and operational hurdles is essential for translating the potential of virtual environments into sustainable clinical practice.

Strengths and Limitations

Strengths

This is the first observational trial to explore the feasibility of psychiatric consultations in a metaverse setting, offering novel insights into how digital environments can be leveraged to support youth mental health. The study benefited from a youth-centered design, the involvement of experienced psychiatrists, and the incorporation of diverse participant feedback. This design enabled a nuanced understanding of how metaverse-based consultations may meet the psychological and logistical needs of underserved youth populations, including those with interpersonal anxiety, sensory sensitivities, or barriers to in-person care.

Limitations

Several limitations should be noted. First, the sample size was modest, and participants were primarily recruited from urban

XSL•FO
RenderX

or institutional settings, which may limit the generalizability of the findings to broader or more diverse populations. Second, the study focused solely on avatar-based communication within VR platforms, without evaluating broader features of the metaverse, such as social engagement spaces, gamified environments, or decentralized user governance structures such as DAOs. Third, populations such as nonnative speakers or individuals with speech impairments were not included, which may limit the inclusivity of the findings. Future studies should explore comparative modalities (eg, light vs heavy metaverse), include more diverse participants, and assess additional functional elements of digital environments to further understand the potential of metaverse-based mental health support.

Implications and Conclusions

This study provides early evidence that metaverse-based psychiatric consultations are a feasible and acceptable approach for supporting young people with mental health conditions. The innovation of this study lies in its focus on avatar-mediated psychiatric consultation within a metaverse environment, moving beyond conventional telepsychiatry models by examining how immersive, customizable virtual spaces may influence user experience and engagement. Unlike prior digital mental health studies that primarily evaluate video- or text-based teleconferencing, this study highlights the distinct role of immersive and self-representational features—such as avatars and virtual environments—in potentially reducing psychological barriers, enhancing autonomy, and supporting help-seeking among youth who experience anxiety, stigma, or discomfort with face-to-face care. By identifying specific subgroups of youth for whom metaverse-based consultations may be particularly beneficial, this study contributes to the emerging field of youth digital mental health by offering practical insights into when and for whom immersive technologies may add value beyond existing modalities.

From a real-world perspective, metaverse-based psychiatric consultations may serve as a complementary access point within stepped or hybrid care models, particularly for young people who face psychological or logistical barriers to in-person or video-based services. However, careful consideration of usability, digital equity, and ethical safeguards will be essential to ensure responsible and scalable implementation. Future research should compare immersive metaverse approaches with conventional teleconferencing, examine long-term clinical outcomes, and explore scalable implementation strategies to support equitable and youth-centered mental health care across diverse settings.

Acknowledgments

We thank Jun Tamura for statistical advice, Minori Ito for assisting with recruitment and trial implementation, and Arimatsu and Miyakawa from Tomoiku Co., Ltd., for providing technical support with the metaverse platform. Generative artificial intelligence tools ChatGPT (OpenAI; GPT-4o) were used to assist with language editing and grammar refinement during manuscript preparation. The authors reviewed, edited, and approved all artificial intelligence–assisted outputs and take full responsibility for the accuracy and integrity of the final manuscript.

Funding

This work was supported by the Japan Science and Technology Agency (grant JPMJPF2203). The funder had no role in the study design, data collection, analysis, interpretation, or the decision to submit the manuscript for publication.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to ethical restrictions and the sensitive nature of the participants' mental health information. Requests for access to the deidentified data can be directed to the corresponding author and may be considered on a case-by-case basis, subject to institutional review board approval. The datasets generated and/or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: MI, JF, NT, KI, T Moroga, T Miyazaki
Data curation: NT, T Moroga, MT
Formal analysis: MI, JF, NT
Funding acquisition: MI
Investigation: MI, JF, NT, KI, T Moroga, MT, TA
Methodology: MI, JF, KI
Project administration: MI, T Miyazaki
Resources: T Moroga, MT, T Miyazaki
Supervision: TA, T Miyazaki
Writing – original draft: MI, JF
Writing – review & editing: MI, JF, NT, KI, T Moroga, MT, TA, T Miyazaki

XSL•FO
RenderX

Conflicts of Interest

MI, NT, KI, TM, MT, TA, and TM declare no competing interests. JF serves as the Executive Director of the Seisa Yokohama Education Counselling Center.

Multimedia Appendix 1
Interview guide that is used for semistructured interview for participants.
[PDF File, 27 KB - xr_v3i1e83688_app1.pdf]

Checklist 1
COREQ checklist.
[PDF File, 61 KB - xr_v3i1e83688_app2.pdf]

References

1. Lu W. Adolescent depression: national trends, risk factors, and healthcare disparities. Am J Health Behav 2019 Jan 1;43(1):181-194. [doi: 10.5993/AJHB.43.1.15] [Medline: 30522576]
2. Kowalchuk A, Gonzalez SJ, Zoorob RJ. Anxiety disorders in children and adolescents. Am Fam Physician 2022 Dec;106(6):657-664. [Medline: 36521463]
3. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S. Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. JAMA Pediatr 2021 Nov 1;175(11):1142-1150. [doi: 10.1001/jamapediatrics.2021.2482] [Medline: 34369987]
4. Kessler RC, Angermeyer M, Anthony JC, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007 Oct;6(3):168-176. [Medline: 18188442]
5. Mental health of adolescents. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health [accessed 2024-11-18]
6. Aguirre Velasco A, Cruz ISS, Billings J, Jimenez M, Rowe S. What are the barriers, facilitators and interventions targeting help-seeking behaviours for common mental health problems in adolescents? A systematic review. BMC Psychiatry 2020 Jun 11;20(1):293. [doi: 10.1186/s12888-020-02659-0] [Medline: 32527236]
7. Radez J, Reardon T, Creswell C, Lawrence PJ, Evdoka-Burton G, Waite P. Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies. Eur Child Adolesc Psychiatry 2021 Feb;30(2):183-211. [doi: 10.1007/s00787-019-01469-4] [Medline: 31965309]
8. Bui AL, Ball AM, Ko LK, Ng M, Rivara FP, Coker TR. Foregone preventive care and unmet mental healthcare needs among U.S. youth. Am J Prev Med 2025 Aug;69(2):107644. [doi: 10.1016/j.amepre.2025.107644] [Medline: 40339830]
9. McGorry PD, Mei C, Chanen A, Hodges C, Alvarez-Jimenez M, Killackey E. Designing and scaling up integrated youth mental health care. World Psychiatry 2022 Feb;21(1):61-76. [doi: 10.1002/wps.20938] [Medline: 35015367]
10. Roulston CA, Ahuvia I, Chen S, Fassler J, Fox K, Schleider JL. "My family won't let me." Adolescent-reported barriers to accessing mental health care. J Res Adolesc 2025 Mar;35(1):e70013. [doi: 10.1111/jora.70013] [Medline: 40017261]
11. Wahdi AE, James C, Nadhira DA, Fine SL. Barriers and facilitators of seeking help for mental health challenges among adolescents across 13 countries: a qualitative investigation. J Adolesc Health 2025 Sep;77(3):421-427. [doi: 10.1016/j.jadohealth.2024.02.035] [Medline: 40908059]
12. Ma KKY, Anderson JK, Burn AM. Review: school-based interventions to improve mental health literacy and reduce mental health stigma - a systematic review. Child Adolesc Ment Health 2023 May;28(2):230-240. [doi: 10.1111/camh.12543] [Medline: 35005830]
13. Seedaket S, Turnbull N, Phajan T, Wanchai A. Improving mental health literacy in adolescents: systematic review of supporting intervention studies. Trop Med Int Health 2020 Sep;25(9):1055-1064. [doi: 10.1111/tmi.13449] [Medline: 32478983]
14. McGorry P, Gunasiri H, Mei C, Rice S, Gao CX. The youth mental health crisis: analysis and solutions. Front Psychiatry 2024;15:1517533. [doi: 10.3389/fpsyt.2024.1517533] [Medline: 39906686]
15. Hollis C. Youth mental health: risks and opportunities in the digital world. World Psychiatry 2022 Feb;21(1):81-82. [doi: 10.1002/wps.20929] [Medline: 35015352]
16. Opie JE, Vuong A, Welsh ET, et al. Outcomes of best-practice guided digital mental health interventions for youth and young adults with emerging symptoms: part I. A systematic review of socioemotional outcomes and recommendations. Clin Child Fam Psychol Rev 2024 Jun;27(2):424-475. [doi: 10.1007/s10567-024-00469-4] [Medline: 38489101]
17. Navas-Medrano S, Soler-Dominguez JL, Pons P. Mixed reality for a collective and adaptive mental health metaverse. Front Psychiatry 2023;14:1272783. [doi: 10.3389/fpsyt.2023.1272783] [Medline: 38250268]
18. Buragohain D, Khichar S, Deng C, Meng Y, Chaudhary S. Analyzing metaverse-based digital therapies, their effectiveness, and potential risks in mental healthcare. Sci Rep 2025 May 16;15(1):17066. [doi: 10.1038/s41598-025-00916-4] [Medline: 40379748]

XSL•FO
RenderX

19. Oh HJ, Kim J, Chang JJC, Park N, Lee S. Social benefits of living in the metaverse: the relationships among social presence, supportive interaction, social self-efficacy, and feelings of loneliness. Comput Human Behav 2023 Feb;139:107498. [doi: 10.1016/j.chb.2022.107498]
20. The world's first hikikomori research lab @ Kyushu University. Kyushu University. URL: https://www.hikikomori-lab.com/en [accessed 2024-11-18]
21. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
22. 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] [Medline: 29345485]
23. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. Front Robot AI 2016;3. [doi: 10.3389/frobt.2016.00074]
24. Satu P, Minna L, Satu S. Immersive VR assessment and intervention research of individuals with neurodevelopmental disorders is dominated by ASD and ADHD: a scoping review. Rev J Autism Dev Disord 2025 Mar;12(1):50-68. [doi: 10.1007/s40489-023-00377-3]
25. Piers R, Williams JM, Sharpe H. Review: can digital mental health interventions bridge the "digital divide" for socioeconomically and digitally marginalised youth? A systematic review. Child Adolesc Ment Health 2023 Feb;28(1):90-104. [doi: 10.1111/camh.12620] [Medline: 36510375]
26. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
27. Mucic D, Shore J, Hilty DM. Overview of the World Psychiatric Association telepsychiatry global guidelines. J technol behav sci 2023 Sep 8;9(3):572-579. [doi: 10.1007/s41347-023-00339-w]
28. Duffy CMC, Benotsch EG. Nonverbal behavior in telehealth visits: a narrative review. Patient Educ Couns 2025 Mar;132:108600. [doi: 10.1016/j.pec.2024.108600] [Medline: 39662111]
29. Gibson K. What Young People Want from Mental Health Services: A Youth Informed Approach for the Digital Age: Routledge; 2021. URL: https://www.taylorfrancis.com/books/mono/10.4324/9780429322457/young-people-want-mental-health-services-kerry-gibson [accessed 2026-03-23]
30. Hetrick SE, Bailey AP, Smith KE, et al. Integrated (one-stop shop) youth health care: best available evidence and future directions. Med J Aust 2017 Nov 20;207(10):S5-S18. [doi: 10.5694/mja17.00694] [Medline: 29129182]
31. Raja US, Al-Baghli R. Ethical concerns in contemporary virtual reality and frameworks for pursuing responsible use. Front Virtual Real 2025;6:1451273. [doi: 10.3389/frvir.2025.1451273]
32. Bouchard S, et al. Virtual reality exposure therapy for social anxiety disorder: a randomized controlled trial. J Consult Clin Psychol 2017;85(3):196-207. [doi: 10.1037/ccp0000181]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
ICT: information and communication technology
MHS: mental health service
VR: virtual reality

Edited by S Brini; submitted 07.Sep.2025; peer-reviewed by L Zhang, M Rubin; revised version received 09.Mar.2026; accepted 09.Mar.2026; published 05.May.2026.

Please cite as:
Ishii M, Fujita J, Toyohara N, Ide K, Moroga T, Takayama M, Asami T, Miyazaki T
Metaverse-Based Psychiatric Consultation for Youths With Mental Health Conditions: Qualitative Descriptive Feasibility Study
JMIR XR Spatial Comput 2026;3:e83688
URL: https://xr.jmir.org/2026/1/e83688
doi:10.2196/83688

© Mio Ishii, Junichi Fujita, Nao Toyohara, Keiko Ide, Tomoko Moroga, Mizuho Takayama, Takeshi Asami, Tomoyuki Miyazaki. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 5.May.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 XR and Spatial

XSL•FO
RenderX

Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Virtual Reality Intervention to Promote Uptake of Medications for Opioid Use Disorder in the Emergency Department Following Opioid-Involved Overdose Based on Input From People With Lived Experience: Development Study

Zina Trost[1*], PhD; Andrea Stevenson Won[2*], PhD; Hari Prasad Naidu Boyapati[3], MBBS; Keri Costa[3], MSN; Josh Davenport[4], BA; Gail D'Onofrio[5,6], MD; Leslie Drager[3], BSN; Phillip Corey Shum[4]; Michael Wolfe Pierce[3], MD; Liana Fraenkel[3,5*], MPH, MD; William Soares III[7*], MS, MD

[1]Department of Psychological and Brain Sciences, Texas A&M University, College Station, TX, United States
[2]Department of Communication, Cornell University, 237 Mann Drive, Suite 492, Ithaca, NY, United States
[3]Berkshire Medical Center, Pittsfield, MA, United States
[4]Immersive Experience Laboratories LLC, Birmingham, AL, United States
[5]Yale School of Medicine, New Haven, CT, United States
[6]Yale School of Public Health, New Haven, CT, United States
[7]Department of Emergency Medicine, Baystate Medical Center, Springfield, MA, United States
[*]these authors contributed equally

Corresponding Author:
Andrea Stevenson Won, PhD
Department of Communication, Cornell University, 237 Mann Drive, Suite 492, Ithaca, NY, United States

Abstract

Background: Opioid use disorder (OUD), which leads to thousands of deaths per year, continues to pose a tremendous challenge to health care systems. While medications for opioid use disorder (MOUDs), such as buprenorphine and methadone, are the most effective treatments for OUD, the initial patient conversation and decision to engage in MOUDs remain major barriers. Decision-making is strongly influenced by the patients' current emotional state. While the emergency department (ED) is a promising access point for initiating MOUDs, patients' readiness to engage in treatment can be limited, in part, by the negative emotions and experiences associated with the ED visit.

Objective: We co-designed and developed a novel emotion-based virtual reality (VR) intervention to increase patient readiness to engage in shared decision-making for treatment with MOUDs following an ED admission for an opioid-related overdose.

Methods: This paper describes a 3-stage, iterative process involving 59 stakeholders, including people with lived experience (LE), to design and develop an intervention that resulted in a VR app and preliminary assessments of user experience.

Results: Our process incorporated the perspectives of LE stakeholders across all aspects of the design and refinement of the VR platform VR-Choice. VR-Choice consists of a VR journey up a mountain ("Mountain Journey"), culminating in the presentation of nine 180-degree videos showing different positive aspects of life during recovery ("Possible Worlds"). Of the 20 participants who experienced the final platform version, none reported more than slight motion sickness symptoms. Nineteen of the 20 participants stated they liked the videos shown in the final version, considered it to be potentially helpful for patients in the emergency room after an opioid-involved overdose, and thought the experience should be offered to those patients. All participants agreed that the program was easy to use, and only 1 participant described it as confusing.

Conclusions: To our knowledge, this is the first effort to use feedback from participants with LE to design an engaging VR experience for people in the ED who may be candidates for MOUDs. Participants' feedback was positive, and we will move forward with development.

(JMIR XR Spatial Comput 2026;3:e84397) doi:10.2196/84397

KEYWORDS

virtual reality; VR; opioid use disorder; opioid agonist medication treatments; patient and public involvement; development

XSL•FO
RenderX

Introduction

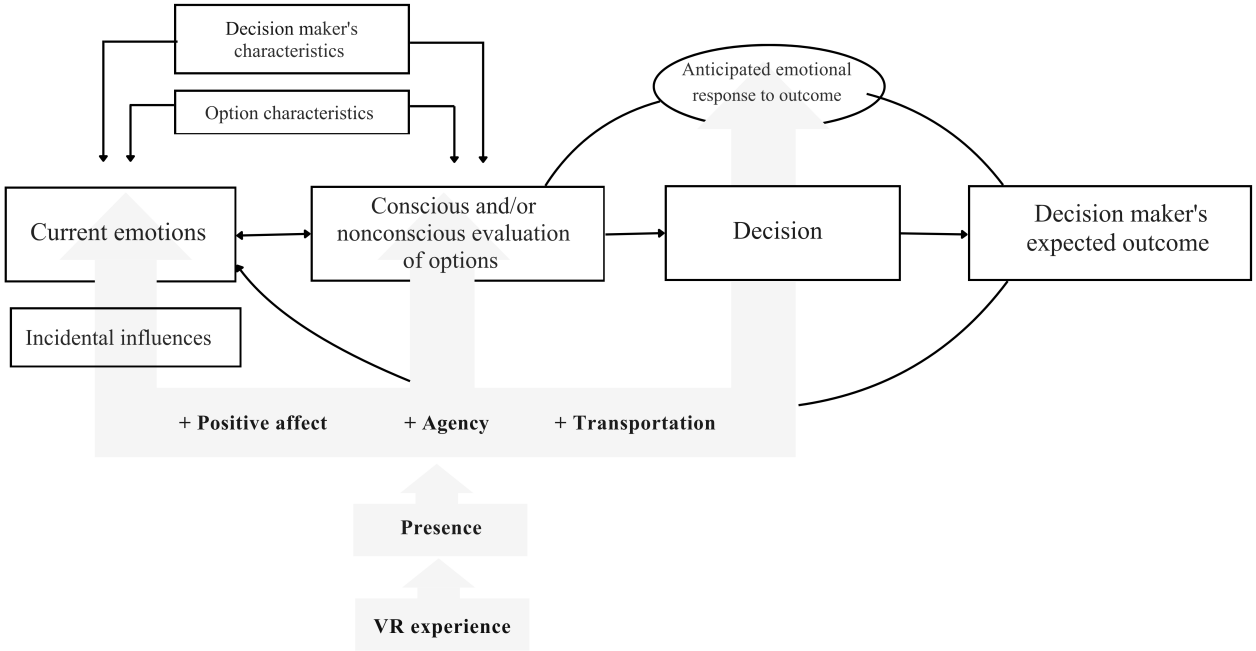
This paper describes the development of an immersive virtual reality (VR) experience, "VR-Choice," designed to increase patient readiness to engage in shared decision-making (SDM) for medications for opioid use disorder (MOUDs) treatment in the emergency department (ED). The VR intervention was developed for use by trained hospital staff in collaboration with ED clinicians.

While deaths related to drug overdoses have recently decreased, opioid use disorder (OUD) remains a leading cause of death in the United States, with 79,384 drug overdose deaths occurring in 2024 [1]. The continuing national opioid crisis necessitates novel initiatives aimed at increasing the adoption of effective treatment. In particular, methadone and buprenorphine (MOUDs) are Food and Drug Administration–approved for OUD and are associated with an approximately 50% reduction in mortality [2-4]. MOUDs also decrease the risk of overdose and acute care use compared with nonpharmacologic treatment [2-5]. Additional benefits include increased retention in treatment, reduced criminal behavior, improved employment, better maternal or infant outcomes, and reduced transmission of HIV and hepatitis C [4,6]. There is strong support that the ED represents a critical access point to engage individuals in care and OUD medication treatment [7]. A landmark randomized controlled trial found that 78% of patients randomized to ED-initiated buprenorphine were engaged in treatment at 30 days, compared to 37% referred to primary care and 45% given a brief motivational interview [7]. Bogan et al [8] extended these findings, showing feasibility across 3 ED sites, with 77% attending follow-up and 60% remaining in treatment at 30 days. Despite their effectiveness, the uptake of MOUDs remains extremely low, with only 19.4% of individuals with OUD receiving MOUDs [9]. ED physicians cite low interest in MOUDs as a top barrier to buprenorphine uptake [10,11]. Data from the US National Survey on Drug Use and Health found that among adults with a substance use disorder who did not receive substance use treatment, 95.6% did not *perceive* that they needed treatment [12]. While education regarding the benefits associated with MOUDs may increase perceived need, existing strategies (eg, motivational interviewing, peer coaches) have not been sufficient [13,14]. Decision-making is strongly influenced by patients' current emotional state. While the ED is a promising access point to initiate MOUDs, patients' readiness to engage in treatment can be limited, in part, due to the negative emotions and experiences associated with the ED visit.

VR-Choice design was informed by the emotion-imbued choice (EIC) model of Lerner et al [15] (Figure 1), which describes the centrality of emotions and affects decision-making. The model posits that an individual's current emotions affect how options are appraised—for example, a positive mood state increases attention to potential benefits. In contrast, negative cognitive or mood states (anxiety, anger) that characterize many individuals with OUD [16-18] are likely to hinder effective decision-making [19-21]. For instance, fear and anxiety may increase (1) preference for the existing state, even when harmful, due to an exaggerated aversion to change (status quo bias); (2) hasty rejection of treatment options (ie, MOUDs) due to a failure to consider all relevant information (premature closure bias); and (3) the tendency to focus solely on information that confirms pre-existing beliefs (confirmation bias) [19]. Each of these biases hinders patients' ability to become fully informed. Conversely, positive emotions encourage attention to benefits and may increase the perceived value of MOUDs; thus, an improvement in a patient's mood may result in greater valuation of MOUDs' benefits. The model also highlights future-oriented processes: anticipated emotions (eg, imagining feeling better on MOUDs) can enhance evaluations of treatment options. Finally, current emotions and expected outcomes interact bidirectionally, reinforcing each other over time [15,22-25].

XSL•FO
RenderX

Figure 1. The emotion-imbued choice model [15] with potential mechanisms of VR intervention. VR: virtual reality.



Drawing on the EIC framework, VR was conceptualized as a promising modality to enhance receptivity toward SDM for MOUD treatment [26]. Specifically, VR has emerged as an effective technology capable of eliciting intense positive affective states in both healthy and clinical populations [27,28], including among acutely distressed individuals and those experiencing acute physical pain [29-31]. VR has also been used to improve affect and reduce pain and anxiety specifically in the ED setting [32,33]. Relatedly, VR can transport people away from their physical environment by replacing sensory information from the physical world with virtual sensory information [34]. Transportation is particularly important because the characteristics of the external physical (ie, the ED) as well as internal (emotional, physical) environment may be acutely distressing to individuals with OUD. Facilitating virtual transport away from these negative influences is expected to improve affect and enable a more open mindset from which positive decisions can be made [35]. Similarly, drawing on interactivity, VR can support feelings of agency, perceived control, and forward momentum, which are consistently lower in individuals with substance use disorder [36-39]. In this sense, VR-facilitated agency can provide individuals in the ED (ostensibly an environment in which they have little control) with evidence of perceivable sensory effects associated with their own actions and choices. Finally, as a simulative technology, VR can support imagination by enabling the simulation of future experiences (“Possible Worlds,” see below). In this way, VR can help counteract the tendency of patients with OUD to focus on immediate, short-term rewards and to discount future outcomes [40,41].

Accordingly, VR-Choice was developed to leverage the ED as a low-threshold, effective approach to engage individuals with OUD. We describe the development of a curated VR experience to increase patient readiness to engage in SDM, a patient-centered, collaborative process in which providers and patients jointly evaluate options, considering individual needs and preferences to determine a course of treatment.

Methods

Overview

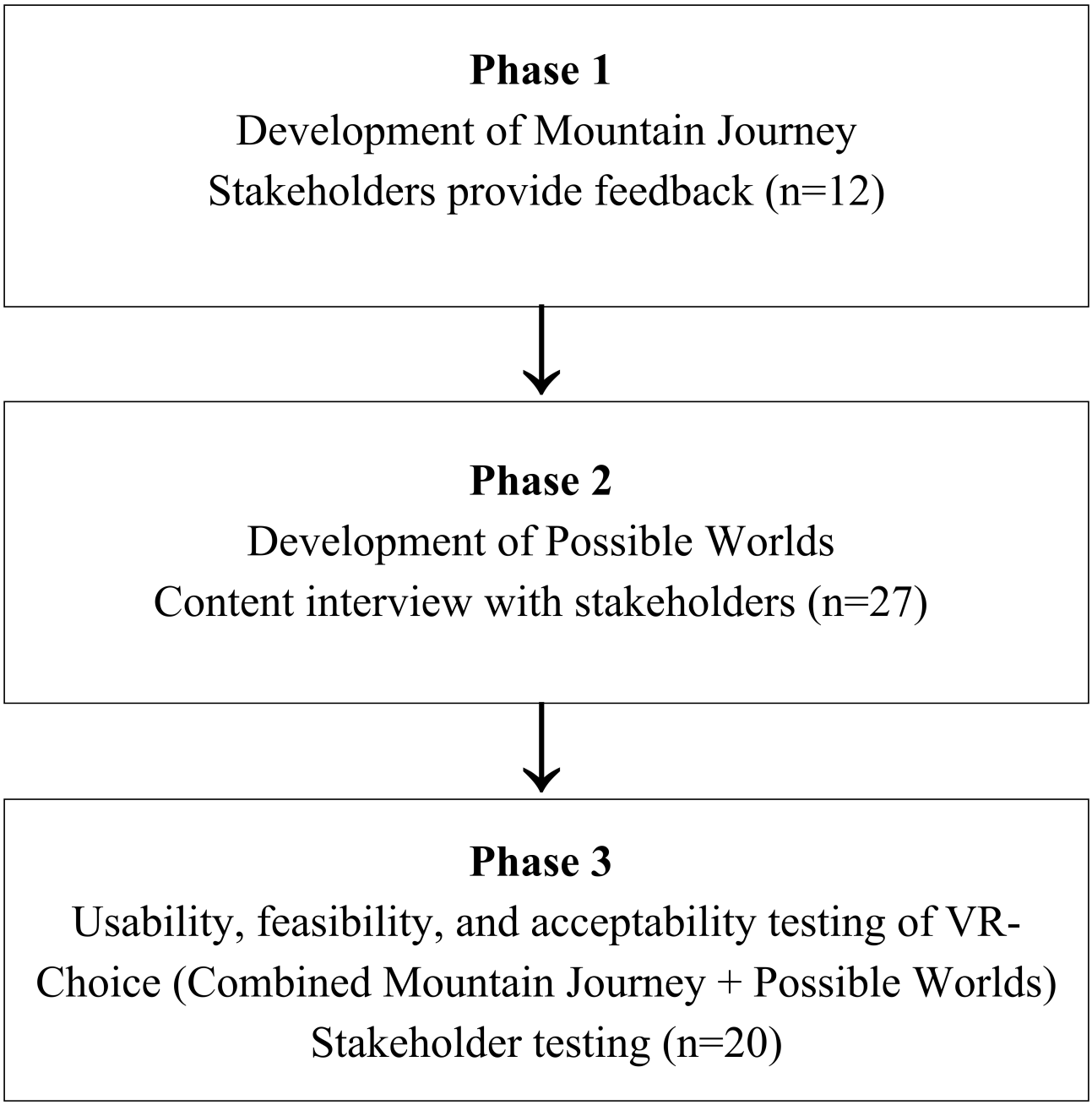
VR-Choice operates using the HTC Vive XR Elite headset and Vive Ultimate Trackers for the left and right wrists. The VR-Choice experience has 2 components. First, participants are presented with a calming virtual environment, developed using the Unity 3D gaming platform (version 2022.3.57f1) [42] intended to induce positive affect and an open mindset; we focused on facilitating the emotion of awe, as it involves cognitive processes—appraisals of vastness, feelings of smallness and connectedness, and accommodation of new ideas—that overlap with the mechanisms within empirically supported interventions for substance use disorder (eg, mindfulness [43]). In this environment, participants proceeded through 4 scenes of increasing elevation to the top of a mountain (“Mountain Journey”). Subsequently, participants were presented with an array of virtual bubbles, each of which, when touched, allowed an immersive glimpse of a 180-degree video showing a positive representation of life after recovery (“Possible Worlds”).

Our process for developing the 2 components of the virtual experience was iterative. In phase 1, we developed the virtual environment for Mountain Journey, which was characterized by relaxation, awe, and forward momentum, modifying it iteratively based on feedback from the research team, academic labs, and interviews with participants with lived experience (LE), who were community members with past or current nonprescribed opioid use. In phase 2, we ideated, selected, storyboarded, and filmed nine 180-degree videos representing positive aspects of life during recovery for Possible Worlds, beginning with interviews with LE stakeholders. In phase 3, we collected feedback on the combined VR-Choice experience

XSL•FO
RenderX

(Mountain Journey + Possible Worlds) from LE stakeholders (see *Phase 3 Feedback Process* for a detailed description of measures). Verbal consent was obtained from all potential stakeholders prior to obtaining feedback or audio recording VR sessions. We describe the process for each phase in Figure 2.

Figure 2. Development of VR-Choice modules for use in the emergency department (ED). VR: virtual reality.



Research Team

The research team consisted of clinicians, researchers, and developers from 4 different institutions, all based in the continental United States. Two members of the team were clinical investigators with expertise in substance use and MOUDs. They received funding for the project and directed the overall research team. Two team members headed VR laboratories specializing in clinically oriented VR interventions. They both engaged their laboratory trainees in the iterative evaluation of the VR environments. The development team consisted of 2 developers with experience in creating VR environments for clinical applications, focusing on pain and rehabilitation. The team met via Zoom weekly or semiweekly during the course of the project and shared stimuli by email and through Google Drive.

Ethical Considerations

All procedures were reviewed and approved as exempt by the Baystate Medical Center Institutional Review Board, and all participants signed informed consent forms or consented verbally. Participants were compensated US $50 per hour, prorated by time (ie, if the interview lasted only half an hour, the participant received US $25). All data presented have been anonymized.

XSL•FO
RenderX

Results

Phase 1: Developing Mountain Journey

Phase 1 Participants

Participants in this stage of development included 6 LE stakeholders who took part in Zoom meetings, 6 LE stakeholders who viewed and commented on videos in person, and members of the participating laboratories and development team (Table 1).

Table . Demographic information of stakeholders with lived experience who provided feedback on Mountain Journey on Zoom or in-person.

Stakeholders	Zoom, n (N=6)	In-person, n (N = 6)
Gender		
Male	2	4
Female	4	2
Self-identified race		
Black	1	4
Hispanic	0	0
White	5	2
American Indian or Alaska Native	0	0
Time in recovery from OUD[a]		
Currently using drugs	0	0
No drug use for 1-3 months	0	3
No drug use for 4-12 months	0	1
No drug use for 1-2 years	0	1
No drug use for 2-5 years	1	1
No drug use for over 5 years	2	0
Unknown	3	0

[a]OUD: opioid use disorder.

Phase 1 Feedback Process

Overview

Once the initial virtual imagery was programmed, participant feedback was sought in 2 ways. First, the research teams met with core LE stakeholders (n=6) over a series of 4 Zoom meetings. During these meetings, a video of the virtual scenery was shared and discussed with all the participants, and recommendations were made for further iteration in terms of journey stages, content, and pacing. LE stakeholders did not have VR equipment at this early point in the development process; however, equipment was available to members of the laboratories of authors ASW and ZT to explore navigation options. Following agreement on content and navigation, the Trost laboratory team collected in-person feedback from additional on-site LE stakeholders (n=6, data shown in Table 1), and feedback on content and navigation was iteratively incorporated as described below. It was not feasible to engage the same stakeholders repeatedly across all phases of the study.

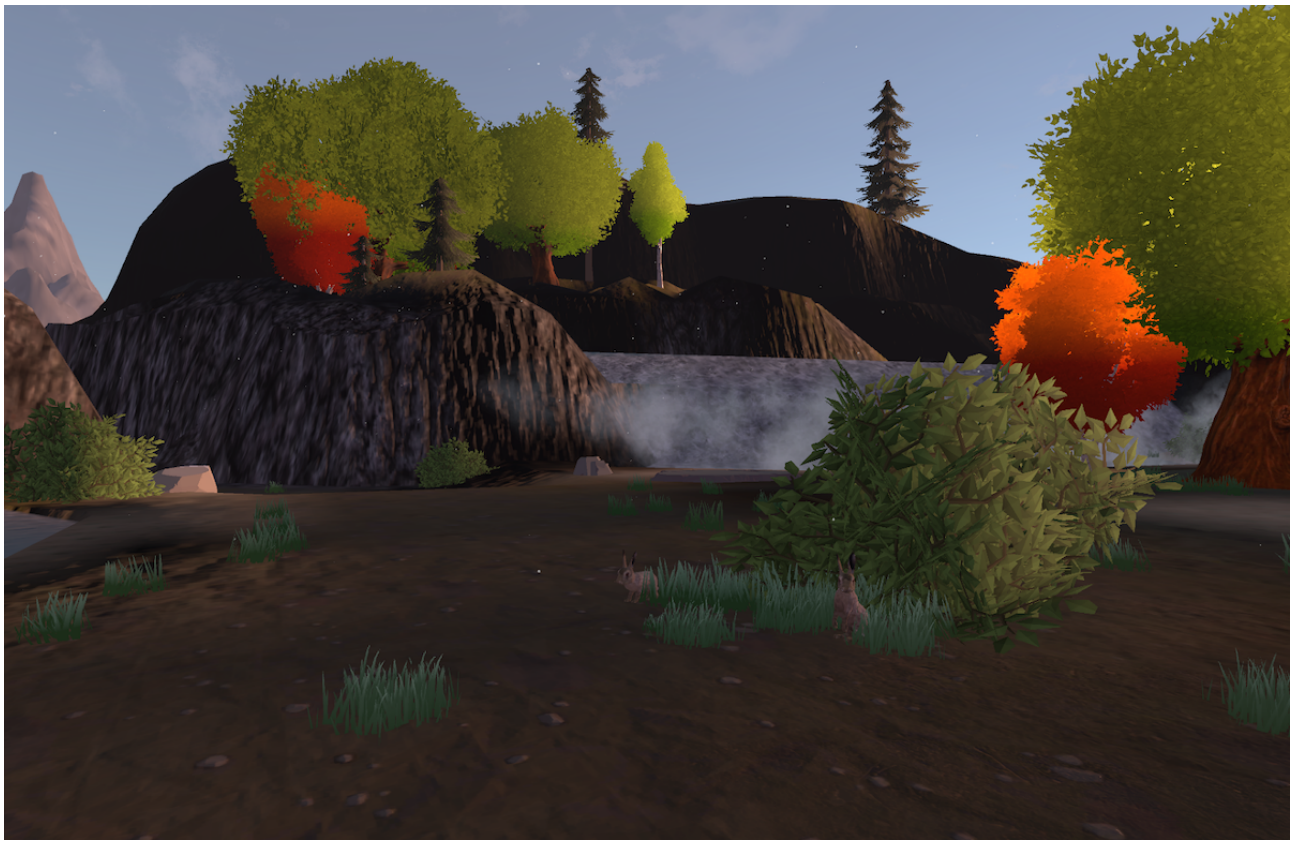
Phase 1 Navigation

Given the environmental constraints (eg, attached medical equipment) and the sensitive physical condition of the target population, considerable discussion focused on the pacing of these scenes and the method by which participants would progress up the mountain. Several methods were tested, including joystick navigation and transportation or teleportation to successive points along the ascent. The various ascent and pacing options were tested successively by students in the research labs, as well as by an LE stakeholder. To alleviate concerns about motion sickness and prevent participant confusion, we settled on a timed scene-fade approach that ensured all participants advanced through the mountain stages at the same pace. The opening scene faded to a scene one-third of the way up the mountain, a scene two-thirds of the way up the mountain, and the final scene at the mountain top (Figures 3-6). As participants progressed up the mountain, snow fell more heavily (to indicate increasing elevation), and the music became louder.

XSL•FO
RenderX

Figure 3. This shows the opening scene of the environment. Birds fly overhead, directing users toward the mountain.

Figure 4. The second scene introduces rabbits browsing in the foreground of the scene.



XSL•FO
RenderX

Figure 5. This shows the third scene, in which the snow is more prominent, indicating a higher altitude. Deer are browsing near the waterfall.

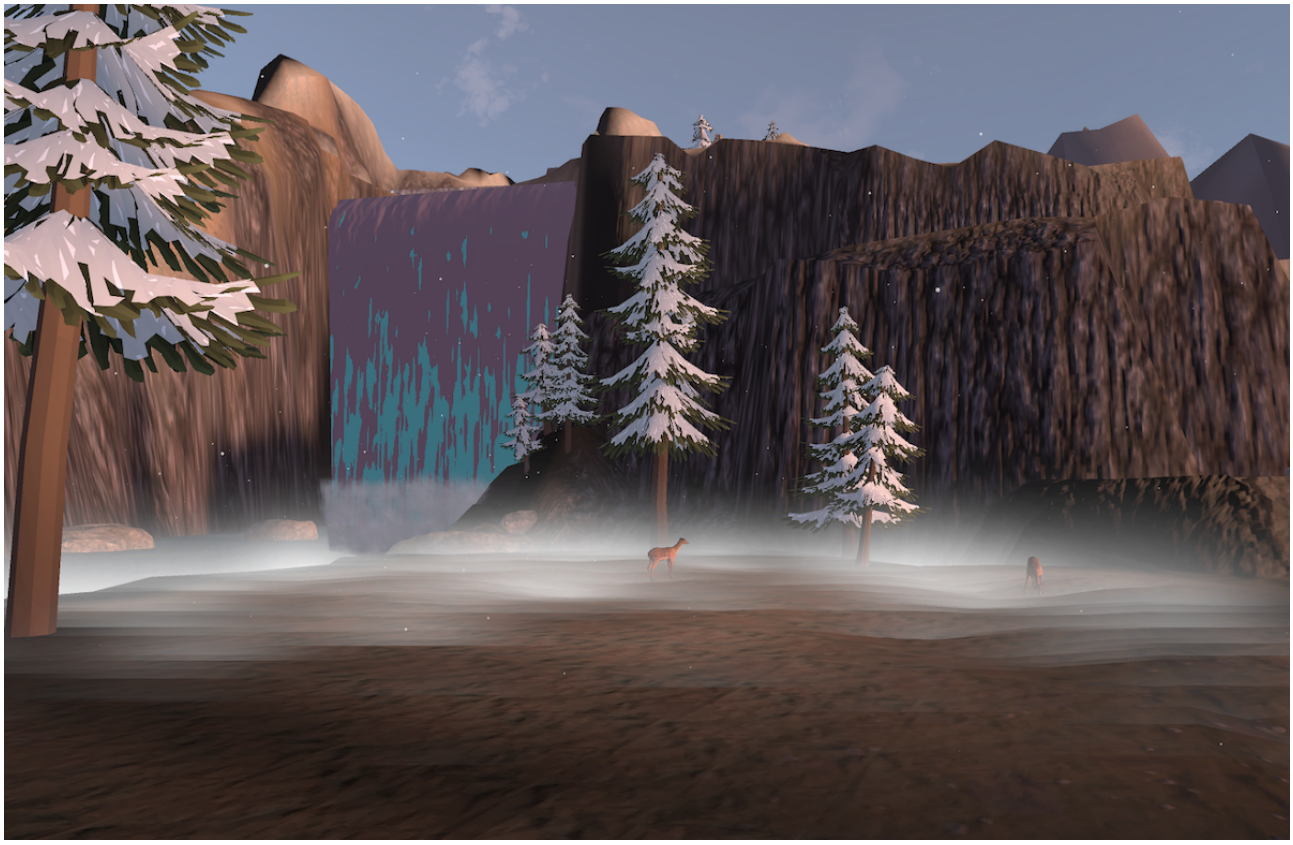
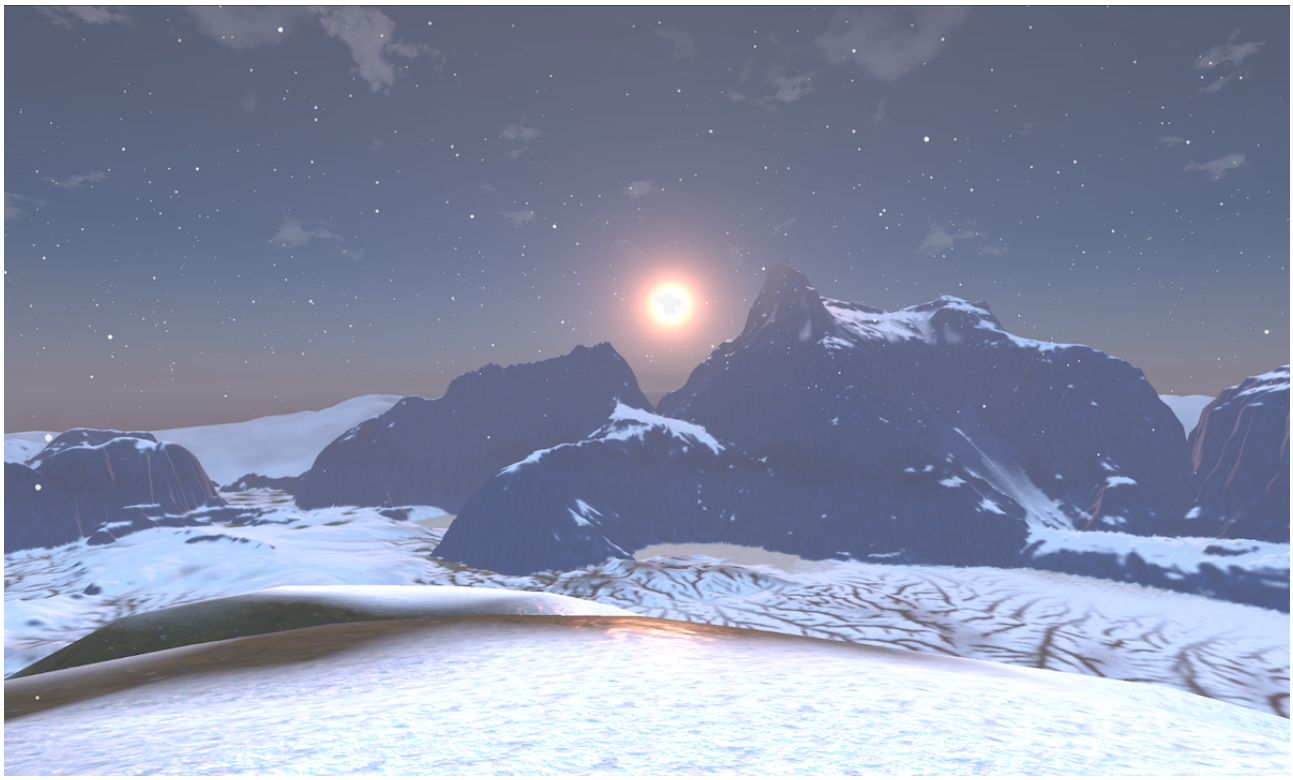


Figure 6. This shows the scene at the top of the mountain.



Phase 1 Content

The initial content for Mountain Journey was collaboratively developed by members of the research team. To generate content for Mountain Journey, the research team searched the literature to identify previous representations of awe in VR, which were limited. We were then provided with existing 360-degree VR imagery, which had been validated to elicit feelings of awe in nonclinical populations [44,45]. Second, we searched for and identified awe-inspiring music, which was likewise specifically

XSL•FO
RenderX

validated in the literature [46] to be used as the background for the intervention. Due to concerns about space on the headset, the designers selected a more abstract, lower-resolution style for the mountain scenes. The app was developed using Unity 2019 and later upgraded to Unity 2022 for performance and analytics enhancements. Gaia (Unity asset) was used for terrain generation within the engine to speed up the process of assigning materials to denote terrain height and slope, as well as to generate a procedural sky for "time of day" manipulation. GeNa (Unity asset) was used to lay out rivers and paths with splines. Vulkan rendering was implemented in place of OpenGL to improve performance and reduce draw calls through object instancing. Other optimizations included reducing texture sizes and video bitrate to achieve a balance between performance, image integrity, and file size. Ideally, we wanted to achieve a framerate that would reduce the risk of nausea for the user while also presenting an open and expansive outdoor environment. In discussions, LE stakeholders emphasized the need for a period of initial relaxation and orientation to the VR space in order to counteract the effects of a busy ER or existing anxiety. Early discussion also emphasized avoiding direct exhortations toward recovery during the VR experience, instead focusing on participants' emotional experience. LE stakeholders also requested the addition of movement and animals in the virtual scenes to keep the participants' attention. These elements were iteratively added. Methods of providing instructions (visible and spoken words) were discussed. The final script can be found in Multimedia Appendix 1. Mountain Journey development resulted in a 2-minute, 35-second-long experience, with approximately 25 seconds at each scene following the initial orientation, as participants proceeded up the mountain.

Phase 2: Developing Possible Worlds

Phase 2 Participants

Participants in this part of development consisted of LE stakeholders who were interviewed to generate bubble content and participated in Zoom meetings to provide feedback on the storyboards, members of participating laboratories, and the development team.

To optimize geographical diversity, the Baystate and ZT laboratories first conducted interviews with LE stakeholders (n=27) in-person, via phone, and on Zoom (relative to stakeholder availability) in order to generate relevant content for the 180-degree videos that would appear in the bubble scenes described above. One member of the research team was bilingual and served as a translator for 3 individuals who were interviewed in Spanish. In total, stakeholders included 10 women and 17 men, ranging in age from 21 to 71 years old (mean 50.30, SD=11.84). Thirteen stakeholders identified as Black or African American, 8 identified as Hispanic, and 6 as White. Recovery ranged from less than 1 month (n=4) to over 11 years (n=5). Demographic information, including recovery characteristics, is outlined in Table 2.

Table . Demographic information of stakeholders with lived experience who provided suggestions for the 180-degree video content for Possible Worlds.

Stakeholders	Participants, n
Gender	
Male	17
Female	10
Self-identified race	
Black	13
Hispanic	8
White	6
American Indian or Alaska Native	0
Time in recovery from OUD[a]	
Currently using drugs	4
No drug use for 1-3 months	4
No drug use for 4-12 months	4
No drug use for 1-2 years	1
No drug use for 2-5 years	6
No drug use for over 5 years	7
Unknown	1

[a]OUD: opioid use disorder.

Phase 2 Feedback Process

Stakeholders were introduced to the concept and purpose of Possible Worlds and were invited to suggest future situations that might be hopeful or motivating for recovery (eg, "What would you look forward to when you are in recovery from your substance use disorder and not using drugs?"). A minority of participants suggested depicting the negative consequences of opioid use, with the idea that these would motivate people toward recovery. Such negative suggestions included depictions

XSL•FO
RenderX

of jail and overdose. Because our aim was to calm and empower participants, we decided to depict only positive situations.

Following these interviews, the collected feedback was reviewed over Zoom with the research team and LE stakeholders. Three principal themes emerged: belongingness, physical and financial security, and a physical transformation to a healthy body. The subsequent Zoom discussions with the videographer focused on how to best represent the selected themes visually in the VR environment. Storyboards were iteratively presented and discussed, and each recorded scene was reviewed by the entire research team.

Content: Possible Worlds Scenes

All videos were shot in 4K and were 180 degrees. In other words, participants could turn their heads 90 degrees to the left or the right. If they tried to turn completely around (unlikely in the VR environment), they would see only black space. Although the user's body was not represented in the VR (when a user looked down, they would only see blank space), the camera was positioned in each scene to create the illusion of embodied presence in the environment. For example, the camera was positioned at an average seated or standing height, in a position where a person could plausibly be located (eg, in a chair, rather than embedded in an object or the ground). Actors in scenes were directed to maintain plausible social distance and make "eye contact" with the camera, so that viewers would feel they were being personally addressed. Although users' individual bodies were not represented, we still needed to consider how to represent diversity through scenes and actors, bearing in mind that individuals' experiences would vary. Because there was no avatar or other way to provide users with a visible body, we selected scenes where the body of the participant could be "invisible" without evoking uncanniness. Videos ranged from 32 to 40 seconds each (Figure 7).

XSL•FO
RenderX

Figure 7. A screenshot from each video offered to participants is shown next to a description of the video itself. Barbecue:

XSL•FO
RenderX

Navigation

In addition to the content and composition of the videos, we also discussed how participants would navigate this part of the experience. The bubbles were laid out in a grid pattern on a slightly concave curvature to allow for easier viewing and reachability. The bubble videos are rendered at 720p and compressed to 360p on a slightly transparent material to allow for a less claustrophobic presentation while keeping the file sizes and processing overhead low enough to run 9 videos simultaneously without catastrophically plunging the framerate. Vignettes of each scene were visible within each "bubble" so that participants could preview the material and avoid selecting content that might be distressing for personal reasons. For example, some participants might choose to avoid viewing materials with children, animals, or religious imagery. Participants could select bubbles more than once and could view as many of the videos as they liked. A voiceover at the beginning of the scene suggests that participants view at least 3 bubbles and that they could leave any scene at any time.

Phase 3: Testing the Usability, Feasibility, and Acceptability of the Combined VR-Choice App

Phase 3 Participants

Participants for this part of the development were LE stakeholders who were recruited from a local transitional housing site serving people with OUD and from a nonprofit community health prevention and education drop-in center that provides health screenings as well as opioid-involved overdose prevention training, education, and naloxone distribution. All LE participants in this portion of the study had a past or current history of OUD and overdose.

This VR-Choice app was tested with an additional group of 20 LE stakeholders. Stakeholders included 13 men and 7 women, ranging in age from 33 to 72 years old (mean 51.80, SD 11.96). Ten stakeholders identified as Black or African American, 2 identified as Hispanic, 7 as White, and 1 as American Indian or Alaska Native. Recovery status ranged from currently using drugs to over a 5-year duration. Recovery characteristics are outlined in Table 3. Of the participants, 1 endorsed regular VR use, 10 reported having tried VR a few times, and 9 indicated no VR experience.

Table . Demographic information of 20 stakeholders with lived experience who provided feedback on the combined VR-Choice experience that included both Mountain Journey and Possible Worlds.

Stakeholders	Participants, n
Gender	
Male	13
Female	7
Self-identified race	
Black	10
Hispanic	2
White	7
American Indian or Alaska Native	1
Time in recovery from OUD[a]	
Currently using drugs	2
No drug use for 1-3 months	4
No drug use for 4-12 months	3
No drug use for 1-2 years	3
No drug use for 2-5 years	4
No drug use for over 5 years	4
Unknown	0

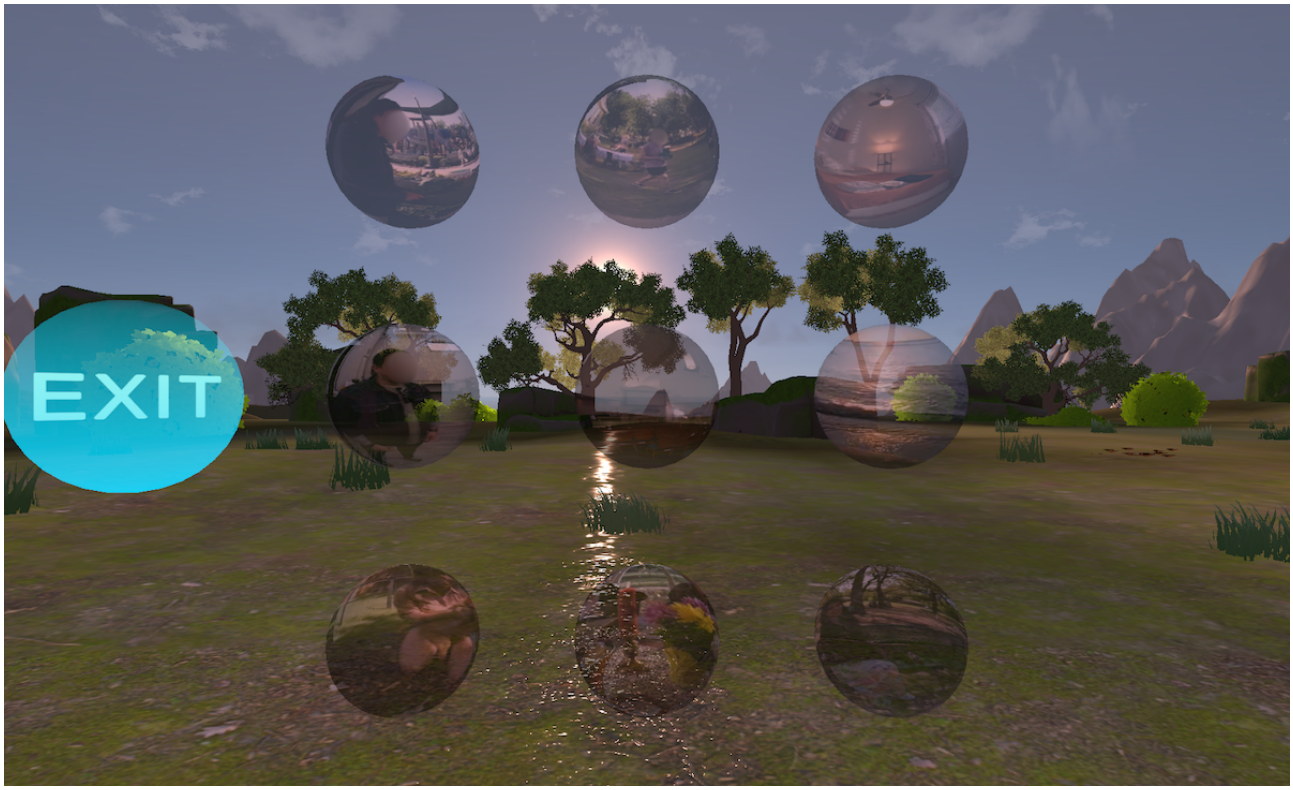
[a]OUD: opioid use disorder.

Phase 3 Feedback Process

The final VR-Choice app combined the Mountain Journey and Possible Worlds sections. Following the ascent up the mountain, participants are automatically transported back to the orientation location at the beginning of Mountain Journey, where they are gradually presented with a grid of 9 bubbles (Figure 8). Three bubbles appear in front of the participant, and after they choose and view one, another set of 3 bubbles enters the scene. In both written and spoken instructions, participants were asked to select at least 3 bubbles, with no upper limit on the number they could choose.

XSL•FO
RenderX

Figure 8. The bubbles presenting the video content appeared in groups of 3, with a new set of videos being added after participants had watched at least 1 from each group. This figure shows the final array of videos.



After experiencing the VR-Choice app, participants completed several measures assessing simulation sickness, usability, and acceptability. The following measures were employed, and the results are summarized in Table 4. With the exception of usability items, which were specific to the content of VR-Choice, all measures had previously been validated in VR-related intervention research.

1. Simulation sickness: Participants also responded to 9 items drawn from the Simulation Sickness Questionnaire [47], inquiring about VR-related symptoms, such as vertigo, blurred vision, and eye strain. Items were rated on a 4-point scale with response options ranging from not at all (0), slightly (1), moderately (2), to very (3).
2. Usability: Usability included specific questions about navigation (“I chose which items I wanted to experience”, “I controlled my actions and experiences in the bubble space”, and “I chose how many different bubbles I wanted to experience”) and general ease of use (“I found this VR program easy to use” and “I think this VR program was confusing”). Items were rated on a scale of 1 (strongly disagree) to 5 (strongly agree).
3. Acceptability: Acceptability items were drawn and adapted from the Treatment Evaluation Inventory [48] and included “I think this VR program is likely to help patients in the Emergency Room after an overdose”, “I like the scenes in the bubbles in the VR program”, and “I think this VR program should be offered to patients in the Emergency Room after an overdose.” Items were rated on a scale of 1 (strongly disagree) to 5 (strongly agree).

A few participants endorsed items from the Simulation Sickness Questionnaire (Table 4), but all were characterized as being “Slight.” In terms of the usability of the experience of interacting with bubbles, 9 agreed and 11 strongly agreed with “I chose which bubbles I wanted to experience.” Similarly, 10 agreed, and 10 strongly agreed with “I chose how many different bubbles I wanted to experience.” Responses to “I controlled my actions and experiences in the bubble space” were more varied, with 9 agreeing, 8 strongly agreeing, 2 being neutral, and 1 disagreeing.

Five participants agreed, and 15 strongly agreed with the statement “I found this program easy to use”; similarly, 3 disagreed, and 16 strongly disagreed with the statement “I think this program was confusing”—1 person agreed with this statement. In terms of acceptability, 12 individuals strongly agreed with the statement “I like the scenes in the bubbles in the VR program”; 6 agreed, 1 was neutral, and 1 disagreed. Thirteen people strongly agreed with the statement “I think this VR program is likely to help patients in the Emergency Room after an overdose”; 6 agreed, and 1 was neutral. Similarly, 15 people strongly agreed with the statement “I think this VR program should be offered to patients in the Emergency Room after an overdose”; 4 agreed, and 1 was neutral.

XSL•FO
RenderX

Table . Participant responses to survey measures.

Domain	Participants, n
Simulation sickness[a]	
General discomfort	2
Fatigue	0
Eye strain	3
Difficulty focusing	2
Headache	1
Fullness of the head	3
Blurred vision	3
Dizziness	1
Vertigo	1
Usability	
“I chose which bubbles I wanted to experience”	
Strongly agree	11
Agree	9
“I chose how many different bubbles I wanted to experience”	
Strongly agree	10
Agree	10
“I controlled my actions and experiences in the bubble space”	
Strongly agree	8
Agree	9
Neutral	2
Disagree	1
“I found this program easy to use”	
Strongly agree	15
Agree	5
“I think this program was confusing”	
Agree	1
Disagree	3
Strongly disagree	16
Acceptability	
“I like the scenes in the bubbles in the VR[b] program”	
Strongly agree	12
Agree	6
Neutral	1
Disagree	1
“I think this VR program is likely to help patients in the Emergency Room after an overdose”	
Strongly agree	13
Agree	6
Neutral	1
“I think this VR program should be offered to patients in the Emergency Room after an overdose”	
Strongly agree	15

XSL•FO
RenderX

Domain	Participants, n
Agree	4
Neutral	1

[a]All symptoms of simulator sickness were labeled as "Slight."
[b]VR: virtual reality.

Discussion

Principal Findings

This paper describes a 3-stage, iterative process that drew heavily on incorporating the experiences and perspectives of people with LE in all aspects of the design and refinement of the VR-Choice platform. VR-Choice was intended to address affect and motivation among patients admitted to the ED following overdose, with the goal of increasing their readiness to engage in SDM for treatment with MOUDs. VR-Choice consists of a VR journey up a mountain ("Mountain Journey") culminating in the presentation of nine 180-degree videos showing different positive aspects of life during recovery ("Possible Worlds"). The bubble content reflects themes derived from LE interviews conducted across the country. In total, we were able to engage with 59 participants with past or current OUD during the design process. Of the 20 participants who experienced the final version of the platform, none reported more than slight motion sickness symptoms. Nineteen of 20 participants stated they liked the videos shown in the final version, considered it to be potentially helpful for patients in the emergency room after an opioid-involved overdose, and thought the experience should be offered to those patients. All participants agreed that the program was easy to use, and only 1 participant described it as confusing.

Both the context and patient population targeted by VR-Choice necessitated careful consideration of multiple ethical and design factors. Specific efforts were made to mitigate simulation sickness, given the likely physical state of patients following overdose or overdose treatment, which informed the design of transitions "up the mountain" and between scenes. Once the intervention is deployed clinically, participants will experience the virtual environment seated and under the observation of health care workers. The positive feedback from non–ED-based OUD stakeholders is encouraging but will need to be further examined within the specific context of postoverdose ED care.

In order to mitigate the risk of psychological distress in a vulnerable population, substantial attention was given to the content of Possible Worlds bubbles, balancing the inclusion of themes endorsed as meaningful by stakeholders (eg, caring for children) with sensitivity to content that could be personally triggering. To that end, the content of each bubble is made readily apparent to viewers (visible inside the bubble) to inform participants' selection, and mechanisms were incorporated to allow for a quick and easy exit if desired. All content was carefully curated to be positive. While the development of the administration protocol draws heavily on input from LE experts, including individuals with OUD and ED personnel, issues around informed consent remain relevant given the potential differences in states of readiness among ED patients during an opioid-related visit. For future evaluations of VR-Choice involving patients in the ED, the informed consent process will include a detailed screening to ensure patients are not altered in mental status, feeling sick, or in significant distress prior to the experience. Consent will also include an overview of VR-Choice—including the content of the Possible Worlds bubbles—and a clear emphasis that patients may remove the headset and stop the experience at any time, for any reason. Following preliminary evidence supporting the feasibility and efficacy of VR-Choice within the ED, future large-scale studies should incorporate implementation metrics, including the integration of VR-Choice within ED workflows and evaluation of the intervention's cost-effectiveness.

Limitations

This experience was limited by both technical factors and constraints in the co-design process. In terms of technical limitations, there was no interactivity, personalization, or embodiment in any of the VR or 180-degree video scenes. Participants could not reach out and touch or manipulate objects. In the video scenes, while we did our best to simulate a sense of social presence, the actors were only pretending to make eye contact and engage with an imaginary speaker and as such could not respond to any behavior on the part of the participants. Because the camera was set at a height of 5′5″, it did not necessarily match participants' actual height. Similarly, neither hands nor body were portrayed in any of the scenes to avoid a mismatch with participants' actual bodies. Therefore, when participants looked down, they would see only empty space. As technology continues to advance, we expect to eventually gain the ability to rapidly create avatars of a user and integrate them into digital scenes; current platforms capable of this, such as the Apple Vision Pro, remain prohibitively expensive and technically demanding for the purposes of this project.

In addition, there were challenges in working with our target population. First, obtaining a representative sample of participants with LE was especially challenging. Participants were often difficult to contact, so follow-up interviews were challenging to conduct. Second, many people do not have personal experience with VR, meaning that for many stakeholders, including many clinical members of the research team, it was challenging to imagine possible outcomes. Future work developing in this area could consider how to provide a sample VR experience before asking people to comment on screenshots or other nonimmersive content. However, this brings other issues, including the fact that the novelty of participants' "first VR experience" might create an anchoring effect.

Future Directions

Next steps will include feasibility, acceptability, and preliminary efficacy testing in the actual ED setting with the population of interest (individuals following opioid-involved overdose). The

XSL•FO
RenderX

data collected from this work will subsequently inform the planning of a larger randomized controlled trial testing the VR-Choice intervention against a standard intervention type.

Future studies may further explore mechanisms of change and participant responses, drawing on both EIC theory and established models of health behavior adoption. For instance, the ability to choose elements within the VR intervention aligns with principles of self-determination theory by potentially enhancing autonomy and competence through interactive engagement with the Possible Worlds bubbles [49]. Similarly, the intervention incorporates aspects of the transtheoretical model of change [50] by targeting the precontemplation and contemplation stages through visualization of recovery-related benefits. Given the acute nature of the ED setting, such questions will likely be best addressed through collaboration with laboratory-based human testing. Pragmatic considerations will ultimately include conducting a cost-benefit analysis comparing the VR-Choice platform with more established (but, as noted, potentially insufficient) approaches, such as motivational interviewing, contingency management, and peer recovery coaching, to highlight the unique contribution of the VR approach. Future studies may also wish to examine combining VR-Choice with more traditional approaches for a potential synergistic impact.

The potential applications of this intervention are wide, including acute treatment programs (detoxification centers), outpatient intake assessments, or harm reduction sites, with appropriate modifications to address the specific barriers relevant to each context. In addition to applications for substance use, it may be valuable to explore the use of such motivationally oriented VR experiences in other health behavior change contexts, such as smoking cessation, eating disorders, and physical activity promotion.

Conclusions

In this paper, we present VR-Choice, a virtual environment designed to improve affect and motivation among patients with OUD in the ED following an overdose and to increase their readiness to engage in shared decision-making for treatment with MOUDs. While the efficacy of this app has not yet been tested, initial feedback from participants with LE points to the feasibility of VR as a way to engage patients undergoing an extremely difficult experience in a challenging environment. This work contributes to the growing body of innovative nonpharmacological approaches for addressing the opioid crisis. Our VR intervention should be positioned within the spectrum of novel digital health interventions being developed for substance use disorders, highlighting how it addresses specific gaps in existing approaches for engaging patients in the critical ED setting.

Acknowledgments

The authors would like to thank Dr Jennifer Michaels, Caitlin McKinnon, and Paula Ciepiela for their help with the recruitment of people with lived experience. Special thanks go to Anthony Najera of Colectivo Creative for video production services. They would also like to thank Isabelle McLeod Daphnis for her help in copy-editing the paper.

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following task was delegated to GAI tools under full human supervision: recommendations. The GAI tool used was Copilot. 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.

Copilot was used to recommend the appropriate reporting guideline for this paper.

Funding

The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: this work was supported by the funding mechanism “Using Innovative Digital Healthcare Solutions to Improve Quality at the Point of Care (R21/R33 - Clinical Trial Optional)” under the grant “Virtual Reality at the Point of Care to Increase Uptake of Medications for Opioid Use Disorder (MOUD) in the Emergency Department” (grant number R21 HS029536).

Data Availability

Because the dataset contains sensitive and potentially identifiable information from hospitalized patients, it cannot be shared. This restriction follows our institutional review board-approved protocol and the ethical requirements for research with vulnerable populations.

Authors' Contributions

Conceptualization: LF, WS, ASW, ZT, PCS, JD, GDO, MWP
Data curation: LD, ZT, HPNB
Formal analysis: ZT
Funding acquisition: LF, WS
Investigation: LF, WS
Methodology: LF, WS, ASW, ZT, PCS, JD, GDO, MWP

XSL•FO
RenderX

Project administration: LD, KC
Resources: LF, WS, ZT
Software: PCS, JD
Supervision: LF, WS
Visualization: PCS, JD
Writing – original draft: ASW, ZT
Writing – review & editing: ZT, ASW, HPNB, KC, JD, GDO, LD, PCS, MWP, LF, WS

Conflicts of Interest

ASW has previously received funding from Meta. The other authors declare no conflicts of interest.

Multimedia Appendix 1
Script heard by participants during the virtual reality simulation.
[DOCX File, 7 KB - xr_v3i1e84397_app1.docx]

References

1. Garnett MF, Miniño AM. Drug overdose deaths in the United States, 2023-2024. NCHS Data Brief 2026 Jan(549). [doi: 10.15620/cdc/174639] [Medline: 41770984]
2. Larochelle MR, Bernson D, Land T, et al. Medication for opioid use disorder after nonfatal opioid overdose and association with mortality: a cohort study. Ann Intern Med 2018 Aug 7;169(3):137-145. [doi: 10.7326/M17-3107] [Medline: 29913516]
3. Santo T Jr, Clark B, Hickman M, et al. Association of opioid agonist treatment with all-cause mortality and specific causes of death among people with opioid dependence: a systematic review and meta-analysis. JAMA Psychiatry 2021 Sep 1;78(9):979-993. [doi: 10.1001/jamapsychiatry.2021.0976] [Medline: 34076676]
4. Sordo L, Barrio G, Bravo MJ, et al. Mortality risk during and after opioid substitution treatment: systematic review and meta-analysis of cohort studies. BMJ 2017 Apr 26;357:j1550. [doi: 10.1136/bmj.j1550] [Medline: 28446428]
5. Pierce M, Bird SM, Hickman M, et al. Impact of treatment for opioid dependence on fatal drug-related poisoning: a national cohort study in England. Addiction 2016 Feb;111(2):298-308. [doi: 10.1111/add.13193] [Medline: 26452239]
6. Nielsen S, Tse WC, Larance B. Opioid agonist treatment for people who are dependent on pharmaceutical opioids. Cochrane Database Syst Rev 2022 Sep 5;9(9):CD011117. [doi: 10.1002/14651858.CD011117.pub3] [Medline: 36063082]
7. D'Onofrio G, O'Connor PG, Pantalon MV, et al. Emergency department-initiated buprenorphine/naloxone treatment for opioid dependence: a randomized clinical trial. JAMA 2015 Apr 28;313(16):1636-1644. [doi: 10.1001/jama.2015.3474] [Medline: 25919527]
8. Bogan C, Jennings L, Haynes L, et al. Implementation of emergency department-initiated buprenorphine for opioid use disorder in a rural southern state. J Subst Abuse Treat 2020 Mar;112S:73-78. [doi: 10.1016/j.jsat.2020.02.007] [Medline: 32220414]
9. Medications for opioid use disorder among adults who had an opioid use disorder. : Substance Abuse and Mental Health Services Administration (SAMHSA); 2025 URL: https://www.samhsa.gov/data/sites/default/files/reports/rpt56616/2024-nsduh-spotlight-moud.pdf
10. Lowenstein M, Kilaru A, Perrone J, et al. Barriers and facilitators for emergency department initiation of buprenorphine: a physician survey. Am J Emerg Med 2019 Sep;37(9):1787-1790. [doi: 10.1016/j.ajem.2019.02.025] [Medline: 30803850]
11. Pourmand A, Beisenova K, Shukur N, Tebo C, Mortimer N, Mazer-Amirshahi M. A practical review of buprenorphine utilization for the emergency physician in the era of decreased prescribing restrictions. Am J Emerg Med 2021 Oct;48:316-322. [doi: 10.1016/j.ajem.2021.06.065] [Medline: 34274576]
12. Key substance use and mental health indicators in the united states: results from the 2024 national survey on drug use and health. : Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2025 HHS Publication No. PEP25-07-007, NSDUH Series H-60 URL: https://www.samhsa.gov/data/sites/default/files/reports/rpt56287/2024-nsduh-annual-national-report.pdf [accessed 2026-06-11]
13. Beaudoin FL, Jacka BP, Li Y, et al. Effect of a peer-led behavioral intervention for emergency department patients at high risk of fatal opioid overdose: a randomized clinical trial. JAMA Netw Open 2022 Aug 1;5(8):e2225582. [doi: 10.1001/jamanetworkopen.2022.25582] [Medline: 35943744]
14. Santos C, Ramdin C, Becceril J, Al Kik B, Jefri M, Nelson L. The impact of an Emergency Department Peer Navigator (EDPN) program in improving clinical outcomes and healthcare utilization in an urban setting. Am J Emerg Med 2023 Jun;68:10-16. [doi: 10.1016/j.ajem.2023.02.031] [Medline: 36905880]
15. Lerner JS, Li Y, Valdesolo P, Kassam KS. Emotion and decision making. Annu Rev Psychol 2015 Jan 3;66:799-823. [doi: 10.1146/annurev-psych-010213-115043] [Medline: 25251484]
16. Blum K, Braverman ER, Holder JM, et al. Reward deficiency syndrome: a biogenetic model for the diagnosis and treatment of impulsive, addictive, and compulsive behaviors. J Psychoactive Drugs 2000 Nov;32 Suppl:1-112. [doi: 10.1080/02791072.2000.10736099] [Medline: 11280926]

XSL•FO
RenderX

17. Kiluk BD, Yip SW, DeVito EE, Carroll KM, Sofuoglu M. Anhedonia as a key clinical feature in the maintenance and treatment of opioid use disorder. Clin Psychol Sci 2019 Nov;7(6):1190-1206. [doi: 10.1177/2167702619855659] [Medline: 32042509]
18. Lister JJ, Lee G, Ellis JD, Pasman E, Agius E, Resko SM. Comorbid affective and substance use disorders of Medicaid/Medicare beneficiaries at an opioid treatment program serving small urban and rural communities. Front Psychiatry 2022;13:881821. [doi: 10.3389/fpsyt.2022.881821] [Medline: 35586409]
19. Bartholomeyczik K, Gusenbauer M, Treffers T. The influence of incidental emotions on decision-making under risk and uncertainty: a systematic review and meta-analysis of experimental evidence. Cogn Emot 2022 Sep;36(6):1054-1073. [doi: 10.1080/02699931.2022.2099349] [Medline: 35838421]
20. Isen AM. An influence of positive affect on decision making in complex situations: theoretical issues with practical implications. J Consum Psychol 2001 Jan;11(2):75-85. [doi: 10.1207/S15327663JCP1102_01]
21. Harlé KM, Guo D, Zhang S, Paulus MP, Yu AJ. Anhedonia and anxiety underlying depressive symptomatology have distinct effects on reward-based decision-making. PLoS One 2017;12(10):e0186473. [doi: 10.1371/journal.pone.0186473] [Medline: 29059254]
22. Baumeister RF, Vohs KD, DeWall CN, Zhang L. How emotion shapes behavior: feedback, anticipation, and reflection, rather than direct causation. Pers Soc Psychol Rev 2007 May;11(2):167-203. [doi: 10.1177/1088868307301033] [Medline: 18453461]
23. Loewenstein G, O'Donoghue T, Rabin M. Projection bias in predicting future utility. Q J Econ 2003 Nov 1;118(4):1209-1248. [doi: 10.1162/003355303322552784]
24. Rottenstreich Y, Hsee CK. Money, kisses, and electric shocks: on the affective psychology of risk. Psychol Sci 2001 May;12(3):185-190. [doi: 10.1111/1467-9280.00334] [Medline: 11437299]
25. Schwarz N. Feelings as information: informational and motivational functions of affective states. In: Sorrentino RM, Higgins ET, editors. Handbook of Motivation and Cognition: Foundations of Social Behavior (Volume 2): Guilford Press; 1990:527-561.
26. Rosén J, Kastrati G, Reppling A, Bergkvist K, Åhs F. The effect of immersive virtual reality on proximal and conditioned threat. Sci Rep 2019 Nov 22;9(1):17407. [doi: 10.1038/s41598-019-53971-z] [Medline: 31758051]
27. 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] [Medline: 32868750]
28. Botella C, Riva G, Gaggioli A, Wiederhold BK, Alcaniz M, Baños RM. The present and future of positive technologies. Cyberpsychol Behav Soc Netw 2012 Feb;15(2):78-84. [doi: 10.1089/cyber.2011.0140] [Medline: 22149078]
29. Hennessy RW, Rumble D, Christian M, Brown DA, Trost Z. A graded exposure, locomotion-enabled virtual reality app during walking and reaching for individuals with chronic low back pain: cohort gaming design. JMIR Serious Games 2020 Aug 10;8(3):e17799. [doi: 10.2196/17799] [Medline: 32773381]
30. Trost Z, Zielke M, Guck A, et al. The promise and challenge of virtual gaming technologies for chronic pain: the case of graded exposure for low back pain. Pain Manag 2015;5(3):197-206. [doi: 10.2217/pmt.15.6] [Medline: 25971643]
31. Trost Z, Anam M, Seward J, et al. Immersive interactive virtual walking reduces neuropathic pain in spinal cord injury: findings from a preliminary investigation of feasibility and clinical efficacy. Pain 2022 Feb 1;163(2):350-361. [doi: 10.1097/j.pain.0000000000002348] [Medline: 34407034]
32. Mosadeghi S, Reid MW, Martinez B, Rosen BT, Spiegel BMR. Feasibility of an immersive virtual reality intervention for hospitalized patients: an observational cohort study. JMIR Ment Health 2016 Jun 27;3(2):e28. [doi: 10.2196/mental.5801] [Medline: 27349654]
33. Andreatta PB, Maslowski E, Petty S, et al. Virtual reality triage training provides a viable solution for disaster-preparedness. Acad Emerg Med 2010 Aug;17(8):870-876. [doi: 10.1111/j.1553-2712.2010.00728.x] [Medline: 20670325]
34. Norton WJ, Sauer J, Gerhard D. A quantifiable framework for describing immersion. Presence (Camb) 2020 Dec 1;29:191-200. [doi: 10.1162/pres_a_00361]
35. Keltner D, Lerner JS. Emotion. In: Fiske ST, Gilbert DT, Lindzey G, editors. Handbook of Social Psychology: John Wiley & Sons; 2010:317-352. [doi: 10.1002/9780470561119.socpsy001009]
36. Bandura A. Self-efficacy: the foundation of agency. In: Perrig WJ, Grob A, editors. Control of Human Behavior, Mental Processes, and Consciousness Essays in Honor of the 60th Birthday of August Flammer: Psychology Press; 2000:17-33.
37. Baler RD, Volkow ND. Drug addiction: the neurobiology of disrupted self-control. Trends Mol Med 2006 Dec;12(12):559-566. [doi: 10.1016/j.molmed.2006.10.005] [Medline: 17070107]
38. Kadden RM, Litt MD. The role of self-efficacy in the treatment of substance use disorders. Addict Behav 2011 Dec;36(12):1120-1126. [doi: 10.1016/j.addbeh.2011.07.032] [Medline: 21849232]
39. Weinberg D. Post-humanism, addiction and the loss of self-control: reflections on the missing core in addiction science. Int J Drug Policy 2013 May;24(3):173-181. [doi: 10.1016/j.drugpo.2013.01.009] [Medline: 23485123]
40. Ahn WY, Vassileva J. Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence. Drug Alcohol Depend 2016 Apr 1;161:247-257. [doi: 10.1016/j.drugalcdep.2016.02.008] [Medline: 26905209]

XSL•FO
RenderX

41. Cheng J, González-Vallejo C. Hyperbolic discounting: value and time processes of substance abusers and non-clinical individuals in intertemporal choice. PLoS One 2014;9(11):e111378. [doi: 10.1371/journal.pone.0111378] [Medline: 25390941]
42. Unity 2022.3.57f1. Unity. 2025. URL: https://unity.com/releases/editor/whats-new/2022.3.57f1 [accessed 2025-11-13]
43. Schuman-Olivier Z, Goodman H, Rosansky J, et al. Mindfulness training vs recovery support for opioid use, craving, and anxiety during buprenorphine treatment: a randomized clinical trial. JAMA Netw Open 2025 Jan 2;8(1):e2454950. [doi: 10.1001/jamanetworkopen.2024.54950] [Medline: 39836426]
44. Mancuso V, Borghesi F, Chirico A, et al. IAVRS-international affective virtual reality system: psychometric assessment of 360° images by using psychophysiological data. Sensors (Basel) 2024 Jun 28;24(13):4204. [doi: 10.3390/s24134204] [Medline: 39000988]
45. Chirico A, Ferrise F, Cordella L, Gaggioli A. Designing awe in virtual reality: an experimental study. Front Psychol 2017;8(2351):2351. [doi: 10.3389/fpsyg.2017.02351] [Medline: 29403409]
46. Ji Q, Janicke-Bowles SH, De Leeuw RNH, Oliver MB. The melody to inspiration: the effects of awe-eliciting music on approach motivation and positive well-being. Media Psychol 2021 May 4;24(3):305-331. [doi: 10.1080/15213269.2019.1693402]
47. Kim HK, Park J, Choi Y, Choe M. Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. Appl Ergon 2018 May;69:66-73. [doi: 10.1016/j.apergo.2017.12.016] [Medline: 29477332]
48. Kelley ML, Heffer RW, Gresham FM, Elliott SN. Development of a modified treatment evaluation inventory. J Psychopathol Behav Assess 1989 Sep;11(3):235-247. [doi: 10.1007/BF00960495]
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. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
50. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]

Abbreviations

ED: emergency department
EIC: emotion-imbued choice
LE: lived experience
MOUD: medications for opioid use disorder
OUD: opioid use disorder
SDM: shared decision-making
VR: virtual reality

Edited by I Steenstra, T Leung; submitted 12.Dec.2025; peer-reviewed by K Johnson, R Bhandari, X Jiang; revised version received 09.May.2026; accepted 12.May.2026; published 06.Jul.2026.

Please cite as:
Trost Z, Won AS, Boyapati HPN, Costa K, Davenport J, D'Onofrio G, Drager L, Shum PC, Pierce MW, Fraenkel L, Soares III W
A Virtual Reality Intervention to Promote Uptake of Medications for Opioid Use Disorder in the Emergency Department Following Opioid-Involved Overdose Based on Input From People With Lived Experience: Development Study
JMIR XR Spatial Comput 2026;3:e84397
URL: https://xr.jmir.org/2026/1/e84397
doi:10.2196/84397

© Zina Trost, Andrea Stevenson Won, Hari Prasad Naidu Boyapati, Keri Costa, Josh Davenport, Gail D'Onofrio, Leslie Drager, Phillip Corey Shum, Michael Wolfe Pierce, Liana Fraenkel, William Soares III. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 6.Jul.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Augmented Reality–Assisted Training Tool for Mental Health Task-Sharers: Pilot Mixed Methods Usability Study

Ling Li Vivian Ngiam[1], MPS; Lucas Piotr Wozniak[1], MPS; Catherine Dinh-Le[1,2], MPH; Ayanna Elon Seals[1], PhD; Victoria K Ngo[2], PhD; Jose Fernando Florez-Arango[2,3,4], MD, MS, PhD

[1]Neurohue, LLC, 600 N Broad Street, Suite 5 #3720, Middletown, DE, United States

[2]Graduate School of Public Health and Health Policy, CUNY, New York, NY, United States

[3]Department Population Health Sciences, Weill Cornell Medicine, New York, NY, United States

[4]Center of Biomedical Informatics and Biostatistics (CB2), University of Arizona, 1230 North Cherry Avenue, PO Box 210242, Tuscon, AZ, United States

Corresponding Author:
Jose Fernando Florez-Arango, MD, MS, PhD
Graduate School of Public Health and Health Policy, CUNY, New York, NY, United States

Abstract

Background: The growing global mental health (MH) burden, especially in underresourced communities, calls for innovative, scalable, and culturally responsive training approaches to expand care access and improve outcomes. Task-sharing has shown promise in addressing workforce shortages but is limited by training and supervision challenges. Traditional methods, such as role-playing and standardized patients, are resource-intensive and less scalable. Virtual simulations, including augmented reality (AR), present novel opportunities for immersive and interactive training. An AR-assisted training tool can enable culturally sensitive training while fostering empathy, communication skills, and confidence in handling nuanced MH scenarios. However, AR's usability and effectiveness for MH task-sharing training remain underexplored.

Objective: This study aimed to assess the usability of an AR-assisted MH task-sharing training tool that uses virtual patient (VP) simulation and evaluate its potential to enhance training. Additionally, we developed design recommendations for future related XR-assisted clinical training tools.

Methods: We conducted a formative, explorative sequential mixed-methods usability study. A convenience sample of 5 MH trainees or workers (ages 18‑60 years; female: n=3, male: n=2; identifying as African American, Asian, and Hispanic) participated. Participants were recruited through a university-affiliated MH training program. The usability testing protocol included a semistructured prestudy interview, orientation to the AR headset (Magic Leap 2), a think-aloud user testing session, and a poststudy quantitative questionnaire and qualitative interview. Usability was assessed using a modified Post-Study System Usability Questionnaire (PSSUQ), which measures system usefulness, information quality, and interface quality. Data were analyzed using descriptive statistics and thematic analysis.

Results: The AR simulation was positively received by participants, demonstrating above-average usability. The overall mean PSSUQ score was 3.46 (SD 1.71), with subscale scores for system usefulness (mean 3.46, SD 1.77), information quality (mean 3.76, SD 1.73), and interface quality (mean 2.83, SD 1.59). Thematic analysis highlighted high realism fostered trainees' empathy toward the VP, while increasing immersion and interaction quality. Despite some hardware limitations and user discomfort that broke immersion, participants recognized the tool's potential and usefulness for training in various MH scenarios. Based on these findings, we proposed design recommendations across environmental context, training structure, VP behavioral realism (body language, voice, eye movement, and technical hardware considerations).

Conclusions: This pilot study is among the first to evaluate AR-based VP simulation for training lay MH task-sharers, filling the technology and population gaps of prior VR- and desktop-focused simulation research. Preliminary empirical usability evidence from a validated instrument (PSSUQ) demonstrates above-average usability, and findings informed design recommendations for future research. These findings suggest AR-based training could provide a less resource-intensive solution for realistic MH training practice in underresourced settings. Further research includes larger sample sizes for inferential analysis, comparison studies with traditional methods, and generalizing to other MH conditions.

(JMIR XR Spatial Comput 2026;3:e80711) doi:10.2196/80711

KEYWORDS

augmented reality; augmented reality glasses; usability; task-sharing; participatory design; mental health care training; role-playing practice; soft skills training; simulation-based training; VP simulation; virtual standardized patient

XSL•FO
RenderX

Introduction

Background

Current Global Mental Health Needs

The worldwide shortage of providers to meet mental health (MH) needs calls for innovative solutions to expand access to care and improve patient outcomes. Depression is a leading cause of global disability; evidence shows that investment in screening and treatment yields substantial health and economic returns [1,2].

In underresourced communities, there are greater barriers to MH care, which calls for context-specific and creative solutions. One factor that may prevent people from seeking appropriate care is MH stigma, which has been found to be higher in racial and ethnic minority populations [3]. Minority groups, including Hispanic and Black communities, are disproportionately affected and less likely to receive treatment [4]. Those who do receive treatment often receive lower-quality services [5].

To address these specific barriers to MH care, there is a need for increasing the supply of providers that exercise cultural humility, as such approaches are more effective at addressing the MH needs of culturally heterogeneous populations [6]. Populations impacted by discrimination and racism often express mistrust of nonminority providers [7]. Addressing these MH care gaps includes expanding appropriate training in cultural competence and instilling humble approaches of care in existing providers [8]. Thus, innovative approaches to training additional providers are needed to meet the growing MH needs of underresourced communities.

Current MH Training Landscape

MH task-sharing is a promising approach for increasing culturally sensitive MH services, particularly in underresourced communities [9]. Task-sharing shifts certain MH services from specialists to less-trained providers [10]. Studies show that task-sharing simple treatments, such as problem-solving therapy and medication management, have been implemented effectively [11,12].

Training and supervision are essential for successful task-sharing, as highlighted by the World Health Organization [13]. Systematic and locally specific efforts are needed to enable intervention sites to design and monitor their own training and supervision [14]. Additionally, simpler therapies, such as behavioral activation delivered by more junior MH workers, have been found to be as effective as more complex therapies delivered by specialists in improving outcomes for patients with depression [15]. Task-sharing interventions are also strengthened by incorporating the services of providers with strong community ties, shared lived experiences, credibility, and nonstigmatizing attitudes [10]. By training these community members and providers who are usually already providing informal support, communities can build their capacity to provide trustworthy MH services.

Resolving challenges associated with supervision and training is essential to improving MH treatment. Although clinical supervision and role-playing are useful training techniques, the conflicting priorities of task-sharing trainees often make it difficult to maintain regular training schedules, hindering the effective implementation of these methods [16]. Additionally, less experienced personnel may struggle to identify and treat subtle symptoms, such as reading facial expressions in patients with severe schizophrenia [17]. This underscores the importance of realistic simulation and accessible, flexible training tools. Traditionally, realistic clinical scenarios have been provided by standardized patients (SPs), trained persons who simulate actual patients and their symptoms [18]. However, this approach is resource-intensive, requiring significant time, travel, financial cost, and supervision [19,20].

Virtual standardized patients are computer-generated, interactive agents designed to simulate clinical presentations, allowing for repeated practice of MH skills such as interviewing and counseling [21]. Because virtual standardized patients do not require the same level of real-time oversight from supervisors, they can alleviate the burden on trainers by enabling self-paced, independent learning.

AR, VR, and Virtual Simulation

Virtual reality (VR) is described as a fully immersive virtual environment, while augmented reality (AR) overlays virtual elements onto the real world [22,23]. AR can foster a strong sense of presence and embodiment, making it a powerful tool in educational and training settings [24,25]. According to research, AR-driven identification and transportation can meaningfully affect users' behaviors, physiological reactions, and psychological perception [26-31].

Social cognitive and situated learning theories support VR's potential for observation-based learning in realistic contexts [32]. VR's immersive and interactive nature also enhances experiential learning, making it effective in training by fostering empathy and user-centered thinking [33]. VR enables high-presence interactions with virtual patients (VPs) and creates realistic training environments [34].

Experiential learning through AR and VR in MH improves users' comprehension, memory retention, and the development of practical skills by allowing them to participate in interactive simulations that closely mimic real-life scenarios [35]. These simulations elicit emotional responses similar to those experienced in real-life situations, nurturing empathy and profound emotional understanding, and providing valuable educational opportunities [36]. Additionally, AR and VR allow learners to practice in realistic environments, developing key skills like procedural recall, communication, problem-solving, and coping strategies [37]. Furthermore, AR and VR have been shown to be interactive, engaging, and convenient, hence making them valuable tools for virtual simulations in MH training and task-sharing settings [38].

Landscape of VR Training

VR training is increasingly used in health care education for both technical and nontechnical skill development [39,40]. In nursing education, a review of virtual simulations found them to be as effective as, or more effective than, conventional methods, offering self-paced, economical, and space-efficient learning opportunities [41].

XSL•FO
RenderX

In MH nursing, VR simulations facilitate emotional connections between learners and VPs by presenting realistic, interactive scenarios [42]. While earlier studies used nonimmersive, desktop-based VR, immersive VR with head-mounted displays (HMDs) has shown greater efficacy due to its sensory immersion, which intensifies focus, presence, emotional engagement, and learning outcomes [43-48]. In a usability test, a 360-degree VR simulation portraying schizophrenia scenarios was found to be user-friendly, engaging, educationally relevant, and highly immersive [49]. Another study on VR simulation training for behavioral health anticipatory guidance and motivational interviewing found that VR significantly increased pediatric residents' engagement and critical skill development by providing an immersive, accessible, and realistic but patient-free learning and practice environment, and reduced their anxiety and stress [50].

Landscape of AR Training

AR is easier to use and has fewer negative physical side effects than VR, including the "cybersickness" that is sometimes felt after long VR usage [51]. AR also increases knowledge transfer by including real-world context that VR does not have [52]. In the medical field, AR is gaining popularity as a tool for clinical training and patient education by enabling medical professionals to visually communicate information about new treatments, medication mechanisms, and surgical procedures [53-56]. AR's applications in anatomical and physiological training are noted for their maturity, showing substantial promise for transforming medical education through innovative prototypes and teaching approaches [57]. An integrative review by Zhu et al [58] found that AR has the potential to improve health care education by lowering failure rates and enhancing precision, while other studies also highlight AR's role in enhancing knowledge retention, concept integration, and confidence among medical students [58-62]. For example, a pilot study contrasting traditional mannequin-based training with AR-enhanced approaches for patient decompensation recognition demonstrated that AR significantly improved participants' assessment skills and clinical confidence [63]. Despite promising early findings, there are still relatively few published experimental and observational studies assessing AR's application and efficacy in medical education [23].

AR in medical education has shown promise in enhancing learning and practical skills, and it has further potential in providing immersive environments for the development of important social skills [64]. By superimposing virtual components within the real environment, AR is very suitable for effective simulation training [58]. Ward et al [65] expressed that exposing users to challenging or rare scenarios in AR simulations creates a safe environment for practicing adaptive skills needed in the clinical space. Despite AR's potential to deliver intricate and precise training across diverse social skills applications, the literature remains lacking, especially within MH training. A recent systematic review pointed out that while simulation-based training in health care education is advancing, there is limited systematic research on its impact on learning outcomes, with most studies to date focusing on feasibility and face validity [64].

Using an AR tool for MH simulation training with VPs is novel; hence, usability testing, defined as the evaluation of elements that impact users' experience with a product for its intended purpose, is an essential first step to determine the strengths, weaknesses, and effectiveness of using AR for task-sharing [66]. Thus, we conducted a mixed-methods pilot study assessing the usability of our AR tool.

Research Objective

Building on the above existing research, we first engaged in a participatory design process to investigate the various needs and challenges pertaining to AR-assisted training of lay MH workers in conducting MH screening and assessments of community members. We then built a prototype which was pilot-tested with stakeholders from the Harlem Strong Initiative (HSI). The HSI, led by the CUNY Center for Innovation in Mental Health (CIMH) team, uses community-engaged planning across a multisectoral coalition of organizations to support MH task-sharing and evaluates the impact of this model in low-income housing and primary care sites in the Harlem community [67,68].

The goals of this study are to assess the usability of the AR-assisted MH task-sharing tool in the current training landscape and to propose design recommendations for future iterations of AR-assisted training tools for MH task-sharing skills and related use cases.

Methods

Research Design Overview

The study design used a formative mixed-methods approach incorporating both qualitative and quantitative data. This design triangulates objective usability scores with subjective user experiences. We chose a modified "think-aloud" protocol for user testing to gather insights into study participants' cognitive processes and any usability difficulties they had during the experience [69]. As a group, we outlined questions for pre- and posttesting qualitative interviews grounded in literature, and we used the Post-Study System Usability Questionnaire (PSSUQ), which has prior validity evidence to assess usability quantitatively [70].

Prototype

The prototype was co-designed with community leaders and members of the Harlem community. Methods used during the participatory design included interviews, field ethnography and observations, early user testing workshops, and multistakeholder co-design sessions.

To better understand and design for the community, we explored three research questions: first, what are the various needs and challenges pertaining to training community-based organization workers in conducting MH screenings of community members; second, how can an AR-assisted MH task-sharing training tool potentially address those needs and challenges; and third, what interaction design strategies can be used to improve the usability of such a tool.

Interviews and brainstorming workshops with community stakeholders highlighted key needs for the training tool- being

XSL•FO
RenderX

time-efficient, offering engaging and realistic content mirroring the nuanced scenarios social workers often face, catering to diverse backgrounds, especially the African American and Hispanic communities, while accommodating varying levels of MH literacy, embedding emotional support for social workers during intense scenarios, and fostering a sense of self-efficacy.

During a field observation session of CIMH's existing online training, screening skills were highlighted as a curriculum section that many trainees wanted to explore further. The Patient Health Questionnaire-4 (PHQ-4) form specifically offers a self-contained, well-structured, brief training scenario that requires soft skill expertise to administer effectively. It consists of four questions that assess symptoms of both anxiety and depression, frequently used for clinical assessments and research purposes [71]. Thus, administering the PHQ-4 was integrated into the design of the prototype.

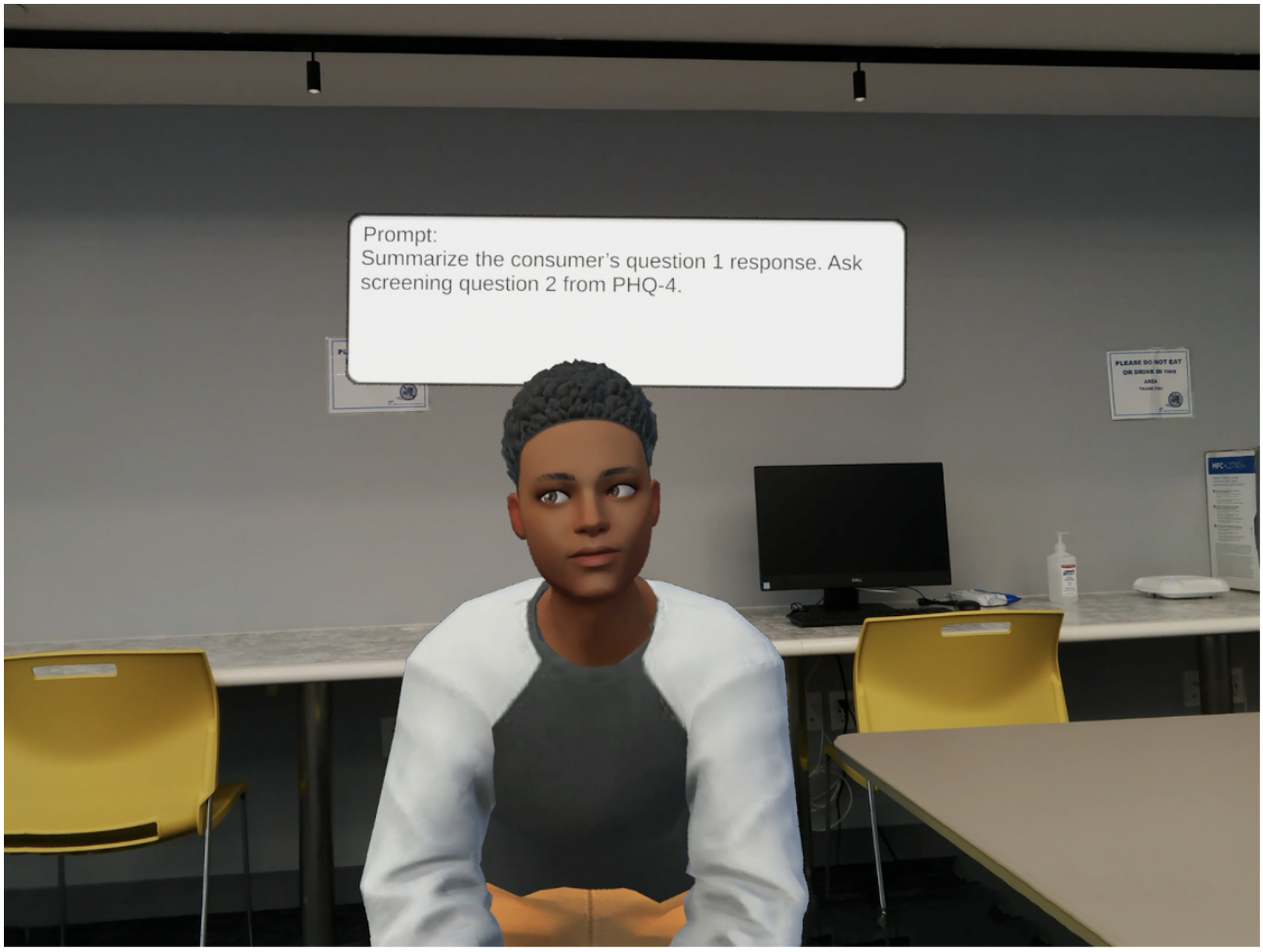
After several co-design sessions, the prototype was finalized as a single MH screening skills module guiding trainees through administering the PHQ-4 appropriately (Multimedia Appendix 1). As seen in Figure 1, the user puts on the AR glasses, Magic Leap 2, and role-plays as an MH provider and interacts with a young adult, Hispanic, male VP named Alvaro, who exhibits symptoms of depression and hesitates about opening up about them [72].

Trainees are guided by 24 different dialogue prompts (Multimedia Appendix 2) that appear above the VP in a roughly 10-minute simulated interaction (Figure 2).

Figure 1. Profile view of a participant wearing the Magic Leap 2 augmented reality headset while interacting with the virtual patient training prototype. The virtual patient appears through the headset and is positioned across from the participant to simulate a face-to-face mental health screening and communication scenario.

XSL•FO
RenderX

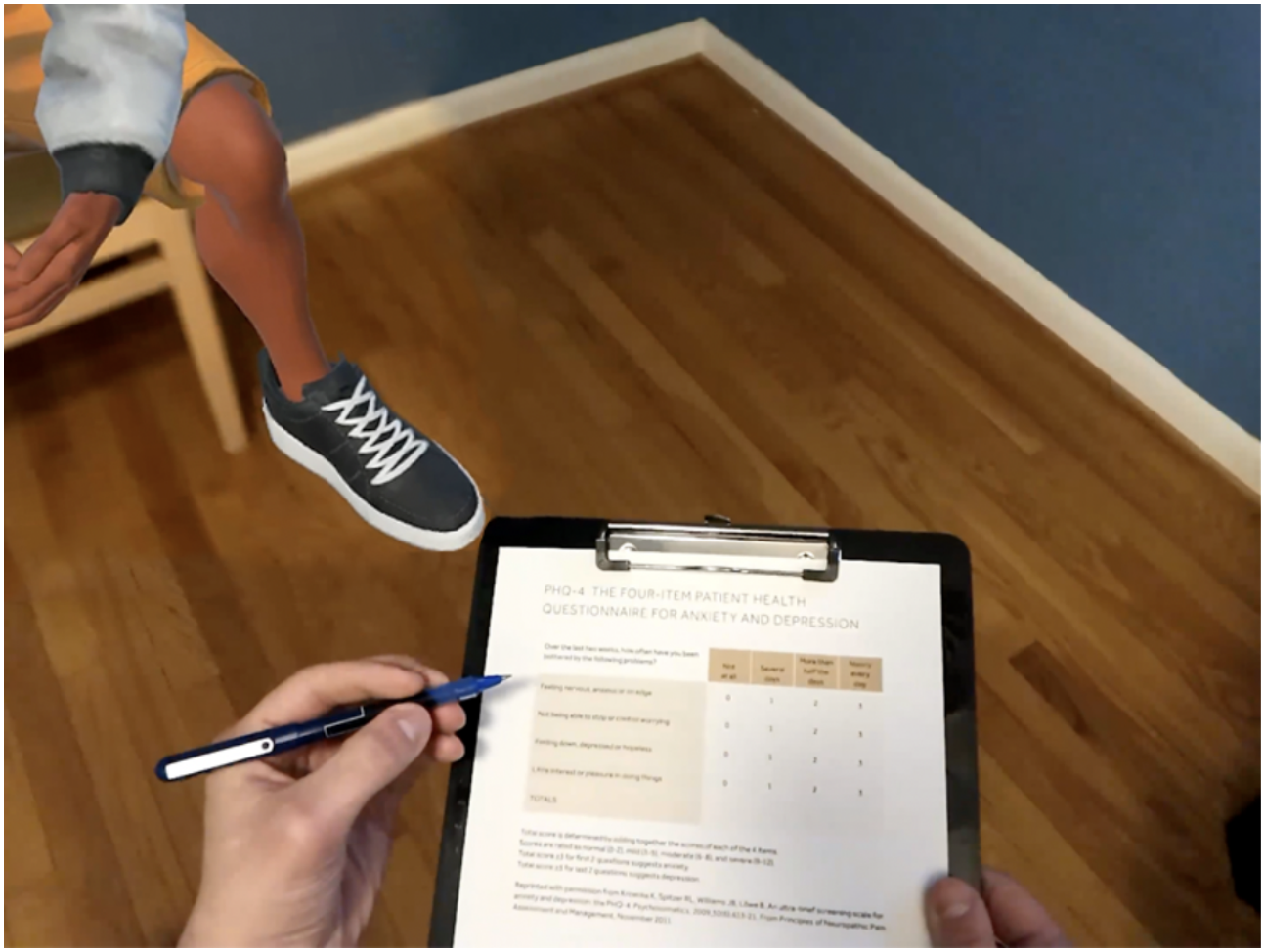
Figure 2. First-person headset view of the virtual patient embedded within the real-world lab environment. The virtual patient is designed to simulate conversational screening interactions and respond to user input during the augmented reality-based training experience. PHQ-4: Patient Health Questionnaire-4.



The VP responds with human-recorded spatial audio voiceovers, lip-synced to animated facial expressions, as well as various body language animations, which enhance the realism and emotional aspect of the experience. Trainees can ask follow-up questions at times or ask the VP to repeat what they said. The experience allows them to practice new communication skills verbally and in an embodied way, engaging with the VP while filling out four responses on a physical PHQ-4 form (Figure 3).

XSL•FO
RenderX

Figure 3. First-person view of the Patient Health Questionnaire-4 screening form presented during the simulation. Participants completed the standardized questionnaire as part of the virtual patient interaction workflow. PHQ-4: Patient Health Questionnaire-4.



Data Collection

Participants

A convenience sample of participants was recruited from CIMH's community network through listservs and word of mouth. Inclusion criteria were broad to capture a wide range of task-sharing perspectives, requiring that participants were members of or adjacent to the community served by the HSI and could speak to the needs of local MH training. Participants recruited for this study group did not overlap with study participants in the main HSI study. Education and MH training levels were varied to generalize the usability of the training tool; participants ranged from a college student with no MH training to an experienced licensed clinical social worker. While five participants were initially recruited and consented, one participant withdrew before beginning the user testing stage. Consequently, the final analytical sample consisted of four participants (n=4), representing a 20% proportion of missing data for the simulation and poststudy measures.

Usability Tests

There were 5 back-to-back usability tests scheduled in a day, with each session lasting between 45 minutes and 60 minutes. Participants were compensated with a US $50 Amazon gift code for their contributions. Their informed consent was first obtained through a physical form, after which we carried out a five-stage data collection procedure:

Participants first engaged in a semistructured prestudy interview, detailing their training background, including their experience with CUNY's MH skills modules and their role in the community. They also discussed the challenges and strategies associated with interactions with distressed community members who could be experiencing MH issues. (Multimedia Appendix 3).

Participants were then briefed about the study goals and AR simulation objectives. They were provided with clear instructions about the AR glasses usability test, emphasizing the experience over their knowledge.

After an orientation to the AR glasses, participants navigated an AR simulation, practicing the administration of the PHQ-4 form with a VP. To facilitate the "think-aloud" protocol, researchers provided a structured briefing and used standardized verbal probes during the simulation. Participants were prompted to describe visual elements (eg, "Describe what you see"), instructed on how to handle interaction cards (eg, "Show us how you would proceed"), and asked targeted questions during planned pauses regarding the character's body language and the use of the PHQ-4 task. If participants deviated from expected interactions, researchers asked clarifying questions such as

XSL•FO
RenderX

"What did you expect to happen?" to capture their cognitive processes.

A semistructured poststudy interview guide was used to assess participants' perceptions of the usability of the AR simulation (Multimedia Appendix 3). Interview guide questions were adapted from usability testing literature [73-75]. The interview aimed to query feedback on the prototype's use, effectiveness, and potential improvement areas. At the end of each interview, the researcher confirmed all relevant information had been included to ensure the quality and accuracy of the data.

Finally, a 17-question poststudy questionnaire (Multimedia Appendix 4), adapted from the PSSUQ with a 7-point Likert scale ranging from strongly agree (1) to strongly disagree (7) with N/A as an additional option [70], was administered. The PSSUQ is a validated instrument with high internal consistency (typically Cronbach $\alpha>0.89$), designed to measure user satisfaction across system usefulness, information quality, and interface quality. Additionally, demographic questions with multiple choice options were included that determined Table 1.

Table . Overview of interview participant characteristics, including age range, gender, ethnicity, prior augmented reality experience, and highest level of education. One enrolled participant withdrew prior to completing the study.

Participant ID	Age range (years)	Sex	Ethnicity	No. of times used an AR[a] or similar device	Highest level of education
P1	35 - 40	Male	Latino or Hispanic	0	Master's degree
P2	26 - 30	Female	Asian	1	Bachelor's degree
P3	56 - 60	Female	African American	0	Master's degree
P4	16 - 20	Female	Asian	0	High school
P5 (dropped out)[b]	—[c]	Male	African American	—	—

[a]AR: augmented reality.

[b]Participant wore glasses, went to get their contact lenses but never came back.

[c]Not available.

The testing took place at one of CIMH's facilities, in a large private room with several tables and chairs. The first and second authors conducted the usability tests and were the only people present in the space besides the study participant. The hardware used includes the Magic Leap 2 and two laptops for filling out the study questionnaires. The prototype was developed in Unity, using Ready Player Me for the custom avatar, Salsa LipSync Suite for facial animations, and Mixamo for body animations [76-79]. The usability lab featured two facing chairs (Figure 4), one for the participant and one for the VP. Researchers were positioned outside the participant's field of view (Figure 5) and intervened only for technical issues or at specific checkpoints (Multimedia Appendix 3).

XSL•FO
RenderX

Figure 4. Individual participant setup during usability testing, showing the seated configuration used for augmented reality-based virtual patient interaction and observation.

XSL•FO
RenderX

Figure 5. Usability lab setup including facilitator workstations and participant seating arrangement used during augmented reality training sessions.

Data Analysis

Quantitative

We used descriptive and summary statistics to analyze participant demographics and scores on the adapted PSSUQ. Due to the pilot nature of the study and the small sample size (n=4), inferential statistical testing was not performed; data are presented as mean (SDs) to indicate trends.

The PSSUQ consists of a set of 17 questions on a 7-point Likert-scale (1=strongly agree to 7=strongly disagree). The overall PSSUQ score is the average of the scores of questions 1 to 15 and 17. The subscales include: system usefulness (SYSUSE), which is the average score of questions 1 to 6, information quality (INFOQUAL), which is the average score of questions 7 to 12, and interface quality (INTERQUAL), which is the average score of questions 13 to 15.

Qualitative

All interviews were recorded using Zoom, with identifying details removed to maintain confidentiality [80]. Initial transcriptions, facilitated by Zoom's automatic service, were refined by the first author, taking out any names, occupations, and other identifying details and using a prescribed participant number instead. After transcriptions were verified for accuracy, they were entered into the qualitative analysis software ATLAS.ti (ATLAS.ti Scientific Software Development GmbH), where a 2-step coding process was adopted [81].

Two independent coders (VN and SF) were used to minimize bias and positionality and ensure accurate depiction of participants' perceptions [82]. The first author (VN) and a research assistant (SF) used an inductive approach based on conventional qualitative content analysis to analyze the interview data [83,84]. In contrast to summative content analysis, which focuses mostly on counting and measuring data, this method emphasizes the identification of recurrent patterns throughout the data collection, similar to traditional thematic analysis [85].

The researchers independently reviewed the interviews and subsequently coded key concepts in line with the interview questions while also drawing from codes from the participatory design phase, such as "Challenges," "Workflow," "Ease of use," "Immersion," "Usefulness," and "Future recommendations." Then, the researchers discussed preliminary findings since multiple analysts for code development enhance the findings' credibility [82]. Using an open coding strategy, codes were refined and added, reflecting significant patterns in the data. Discrepancies in coding led to discussions until consensus was reached. The frameworks described in a related study by Hess et al [86] served as further guidance for the coding process. Domains from the study like "Experiential satisfaction," "Learning engagement," "Technology learning curve," and

XSL•FO
RenderX

"Opportunities for improvement" formed the basis for theme identification.

Transcripts were then restructured according to common codes for thorough examination of themes and patterns. Participants' frequently overlapping and closely related codes were grouped together to create categories, which in turn led to the formation of primary themes. The categories were formed and summarized individually. Comparing these categories and summaries between the 2 researchers constantly and discussing to reach consensus allowed for principal themes (Multimedia Appendix 5) to be formed with a consistency of more than 90%, suggesting additional reviews were unnecessary [82].

Ethical Considerations

This study was approved by CUNY's Institutional Review Board (IRB2021-2031) as part of the Harlem Strong Mental Health Coalition study (U01OD033245). All procedures performed were in accordance with the ethical standards of the institutional and/or national research committee, and written informed consent was obtained from all participants. The consent forms included information on the study purpose, expected procedures, time commitment, potential risks, discomforts, and/or benefits, such as participation payment, and participant confidentiality and rights. Participants were given descriptions of how and with whom their data would be safeguarded, used, and shared; details in the consent form included the potential for deidentified data submission to the National Institute of Mental Health Data Archive (NDA) for secondary analysis without additional consent.

Authorized members of the research team managed the password-protected data for any identifiable information that might be linked to participants from this study. Codes were assigned to replace names of participants, and identifying information was removed when no longer needed. Participants were compensated with $50 Amazon gift cards after the completion of their session. In terms of included images of participants completing the study, only the study team members are identifiable in photos. Participants all wore the AR headset, and written permission was obtained to include their deidentified images.

Validity, Reliability, and Methodological Integrity

In accordance with JARS-Qual standards, we acknowledge the influence of the research team's background on data interpretation. The primary investigators include digital health informaticists, HCI researchers, and clinical psychologists with extensive experience in MH task-sharing. This diverse expertise allowed for a rigorous evaluation of the tool's technical performance while remaining grounded in clinical reality.

Quantitative reliability was supported using the validated PSSUQ instrument. Qualitative integrity was maintained through investigator triangulation and the use of the think-aloud protocol, which captured real-time cognitive data, reducing retrospective recall bias.

Results

Overview

We recruited a total of 5 lay MH trainees and workers with various ranges of training and experience in MH care (for participant details, see Table 1).

Prestudy Interview

To set some context for this inquiry, we present a brief background of relevant current workflows described by the pilot study's participants during the prestudy interviews.

Building Trust Is Challenging but Important

Building trust with the community was highlighted as crucial by two participants. P1 told us that they

> *definitely use a lot of open-ended questions,*

While P3 highlighted that

> *Trust is a huge factor in getting community members to know that you're there for them, that you're concerned about their well-being. You'd like to maybe introduce them to some practices that they may not be aware of or might not understand fully and let them know that it's okay.*

P3 further described:

> *I might tend to go a little bit further than intended for whatever the session is supposed to be, as more of a listening ear type thing, and just try to engage in trying to get more information from them that I might not necessarily use.*

Difficulty in Adapting Theoretical Knowledge to Practice in Current Training

Participants experienced difficulty adapting theoretical knowledge to practice. P2 noted that

> *because every encounter is different, it's kind of very foundational. You have to develop your own way of counseling,*

While P1 emphasized that

> *when you come into actually working in the field, it's a little different. You might be in a setting where you're dealing with things that are the least favorable settings.*

Current Training Included Role-Play and Usually Did Not Include Difficult Scenarios

P1 told us that in the current workflow they: "*covered* The *Diagnostic and Statistical Manual of Mental Disorders*," which is a comprehensive classification system published by the American Psychiatric Association (APA), and "*did a lot of role-play.*" They added that "*normally, when we do role-plays in class, nobody goes down that road of re-enacting more challenging scenarios like schizophrenia, people tend to present the lighter stuff that might not be entirely accurate.*" This highlighted that current training included role-play and usually did not include difficult scenarios, which corroborated with our prototype design to focus on role-play.

XSL•FO
RenderX

Content of Sessions

P1 also noted that typical sessions were brief:

> *Most of the sessions that I work with were 20 min sessions. So if I only had 20 min to gauge how a person's week went, you know, to go through anything larger than 5 questions to gauge their temperament will be challenging.*

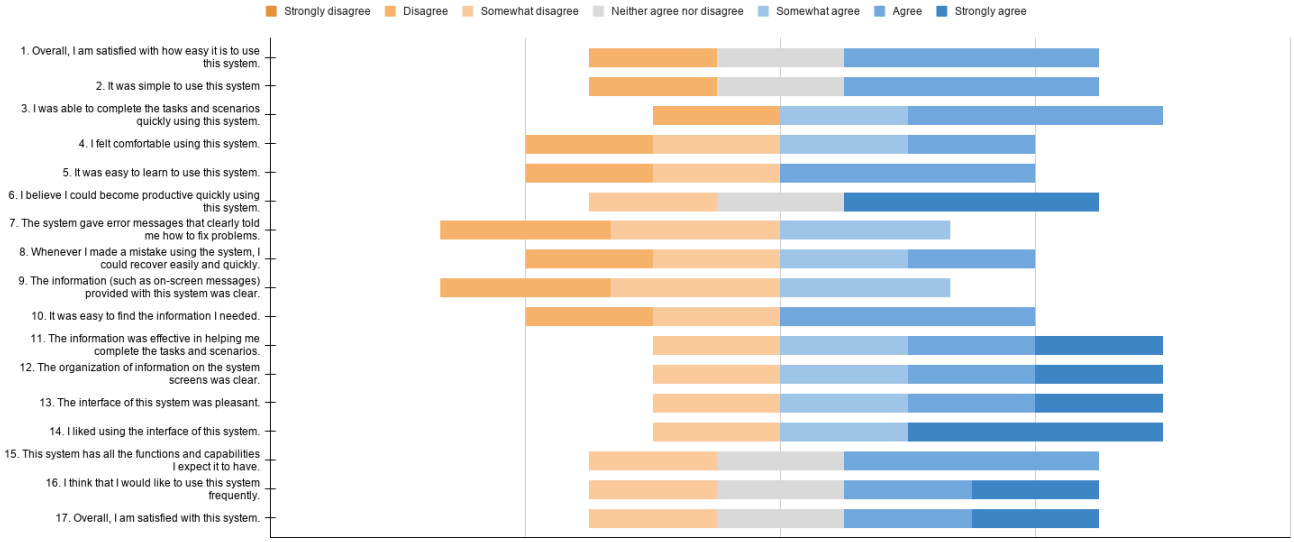
This informed our prototype design choice of administering the brief PHQ-4 form.

Usability Findings

Overall, the four participants rated the usability of the prototype as above-average usability (Multimedia Appendix 6) in the modified PSSUQ survey which includes an additional question on the desired frequency to use the system (question 16).

The prototype received an overall PSSUQ score of mean 3.46 (SD 1.71), SYSUSE score of mean 3.46 (SD 1.77), INFOQUAL score of mean 3.76 (SD 1.73), and INTERQUAL score of mean 2.83 (SD 1.59). Desire to use the system frequently (question 16) received a mean score of 3 (SD 1.83). Figure 6 displays the distribution of responses across all items. One participant chose N/A for questions 7 and 9, which skewed those results. Participants who identified as Asian experienced difficulty with the headset fitting comfortably on their faces (due to a missing nose piece kit) and gave higher scores on their PSSUQ questionnaire.

Figure 6. Distribution of participant responses to the Post-Study System Usability Questionnaire. Bars leaning toward agreement indicate positive perceptions of system usability. N/A responses were excluded from the visualization.



Think-Aloud Insights

In this section, we outline the observations made during the think-aloud section in which participants interacted with the intervention.

Use of Physical Headset

Two out of the four participants experienced the headset falling off their face, requiring constant adjustment and even manual holding to continue. P2 mentioned that

> *it broke the immersion to me. I think it would feel more like I'm talking to someone if the headset didn't keep falling, I was distracted by it.*

P4 felt similarly:

> *it's just slipping down everywhere. My arms were like sore from holding it up because they're heavy, too. But yeah, I think I was a little bit distracted by how like it kept falling down. I think also like the whole thing with the headset not fitting just like took away from it.*

By P4's session, the headset was also heating up, compounding discomfort.

Delivery of Content

All participants reported difficulty hearing the VP's audio at maximum volume, requiring researcher clarification at multiple points.

P3 and P4 also reported difficulty understanding the VP's accented speech at times, with P4 unable to comprehend certain responses even after replays.

P2 and P4 experienced technical glitches where the VP failed to respond to prompts, requiring researcher intervention to continue.

Key Themes From Postintervention Interviews

Following the think-aloud section, the study's major themes identified from the postintervention interviews are discussed here, mainly experiential satisfaction, navigation of the training experience, and interacting with the VP.

Experiential Satisfaction

The overall impression of the simulation, particularly the emotional valence and level of satisfaction for the participants, was generally positive, consistent with the above-average overall PSSUQ usability score (mean 3.46, SD 1.71). The simulation's realism allowed for participants to feel empathy toward the VP. Participants acknowledged the potential usefulness of the

XSL•FO
RenderX

simulation in other contexts, and first-time AR and VR headset users also voiced some struggles.

General Positive Sentiments

Participants generally had positive sentiments toward the training module and found the simulation to be very useful. P2 expressed that "it's a really good training module and the prompts were very helpful in guiding." Echoing similar sentiments, P3 pointed out that this was their first AR and VR experience and that "it was cool and successful." Drawing connections to current training, P1 mentioned that adding this kind of training and technology would be beneficial for practicing difficult cases during role-playing training as they could better depict challenging scenarios. They emphasized:

> *I think maybe having specific augmented reality trainings for specific DSM categories would be very helpful … like depression and anxiety … if it's tailored towards, you know, something that may be heavier, that would be interesting.*

Furthermore, they highlighted that the usage of the PHQ-4 form itself would be useful for training as it would also be helpful to use it in the field.

Realism Resulted in Empathy for the VP

The realism of the simulation and the quality of the animations led participants to feel empathy toward and connect with the VP. Participants all expressed that they felt like they were talking to a real person, and some also mentioned that they forgot that the researchers were in the same room while having the conversation, as the simulation felt so realistic. Three out of the four participants also noted that the voice of the VP felt very realistic as it had an accent, making the scenario very similar to what it would be like in the field. P1 emphasized that

> *it kind of forces you to empathize. It forced me to look at the AI almost like a real person, like, this person has an accent.*

Potential Usefulness of Simulation in Other Contexts

Furthermore, participants highlighted the potential usefulness and applicability of the training approach in other contexts, especially in different MH and specific *DSM* (*Diagnostic and Statistical Manual of Mental Disorders*) categories. Other contexts mentioned include general one-on-one interactions, for example, in customer sales training. With regards to other MH contexts, P1 pointed out that it would be a lot more beneficial to carry out challenging role-play scenarios through AR than to act them out, as it may misrepresent symptoms of, for example, schizophrenia. P3 also highlighted that it would be extremely useful to have simulations like this used in the context of different MH conditions, as having and practicing the lived application of this knowledge is important for people who interact with people who might display these conditions.

Struggles for First-Time AR and VR Headset Users

Two out of the three first-time AR and VR headset users told us that they were slightly confused at the start when trying to navigate the experience. They felt a little overwhelmed between having to pay attention to the prompts, ask questions from the PHQ-4 form, and pay attention to the VP's speech while trying to understand the accent, be attentive to the VP's body language, and maintain eye contact. The uncertainty of what would come next initially also made the participants feel uneasy. On the other hand, P1 told us that

> *if this was my second or third practice, I would know I can read the prompt and look at the patient.*

Two participants mentioned that with more practice, they would be able to get the hang of the training and reap the benefits from it.

Navigating the Training

This section details the participants' experience with the system's guidance mechanisms, providing context for the information quality subscale (mean 3.76, SD 1.73), where the instructional design was found to be functional but occasionally confusing.

Effectiveness of Prompts

Overall, participants found the prompts (Figure 7) useful in guiding them through the training. P3 appreciated that the prompts helped them focus on the specific goal at that certain time and that the prompts worked well in helping them follow the flow of the training. P4 mentioned how the prompts helped them to learn appropriate responses when they were unsure of how to respond. They told us:

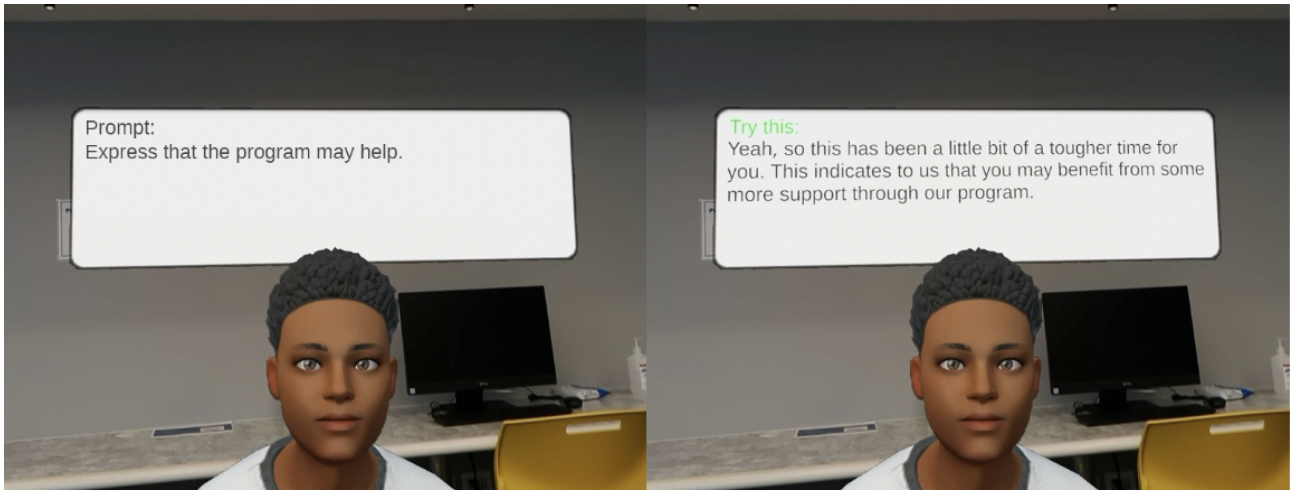
> *I'm not sure if Im like you know, helping or hurting. So that's why having the prompts helped. It's like semantics also have impact on how you affect the client ...*

They further mentioned that

> *… at times I was really lost as to what to say, but also because I didn't hear him so well. That kind of made it more helpful for the prompt, because I was like, okay, this is what I say now.*

XSL•FO
RenderX

Figure 7. Comparison of two in-system prompt styles within the virtual patient simulation: (left) a general guidance prompt encouraging users to respond in their own words, and (right) a structured "Try this" prompt offering a suggested phrasing. The figure illustrates alternative scaffolding strategies tested for communication support.



However, navigating between 2 types of prompts (Figure 7) confused users and reduced immersion. Participants mentioned that having to pay more careful attention to the 2 different kinds of prompts broke the flow of immersion and took away from the experience. P2 elaborated that

> *when it's the 'Try this' prompt, I'm like kind of fed what to say, but then there's the other prompt where I kind of have to think on my feet. Switching back and forth was kind of confusing for me.*

P3 and P1 similarly felt that the simultaneous demands of attending to VP dialogue, body language, and PHQ-4 questions alongside varying prompt types created cognitive overload. P1 wanted more attention to the VP's "body languages, whatever it may be, tone of voice, things of that nature," hence the navigating between the 2 prompts did sometimes confuse them.

Administering the PHQ-4 Form

Participants pointed out that it was difficult to interpret qualitative responses and translate them into a quantitative scale (PHQ-4 form). P1 told us that

> *At times, I was unable to gauge like an appropriate answer from his response so I wanted to interact with him more to see what that meant. I felt like I needed to pull more.*

They further elaborated that they wanted to ask the VP more clarifying questions and have a more dynamic interaction to get the correct scale reading. P4 brought up that people usually don't know how they feel, and it is MH workers' jobs to be their advocate and interpret their response into a number and understand how to help them. Yet P4 expressed that

> *it really depends on the mental health professional right? Who is speaking to them like, what do they think that his, or her, or their response means, and translate all to a quantitative scale, which is difficult.*

Interacting With the VP

In this section, participants' interaction with the VP will be reported, mainly their listening to the VP, observing the VP, and the flow of conversation with the VP. These qualitative insights help explain the strong interface quality score (mean 2.83, SD 1.59) attributed to the VP's realistic body language, as well as the auditory challenges that impacted the Information Quality subscale (mean 3.76, SD 1.73).

Listening to the VP

Participants found that using a real person's voice with an accent increased realism, immersion, and hence the usefulness of training for specific populations or communities. P1 remarked that the VP's accent is very realistic, which reflects people in the community having different accents that might be more difficult to understand. P4 echoed this and felt like the voice was real because it had an accent. P4 elaborated that

> *If it was an automated voice I would have felt weird, that would have added to like the whole unrealness. Using a real person's voice definitely helps… yeah, at one point… like it hit me that he wasn't actually there. It was like just the two of us having a conversation.*

Although P2 felt that the voice was impactful, they mentioned, "I wasn't sure how to respond, because I couldn't understand. I think his accent, and also like the volume was low."

All participants experienced difficulty with hearing due to low volume and accent. However, this motivated them to listen more intently during the training, and they expressed that it was beneficial to train for real-life situations in which some people are more soft-spoken. P1 raised that "it trains me to hear, so it does serve a benefit" echoing P2, who noted "in a real-life situation, when someone is speaking softly or in a different accent, I would need to be listening more intently." P4 reiterated "I really had to process, really listen, which is good, I was really listening."

Observing the VP

Overall, participants highlighted that the body language of the VP was useful in training and increased realism, immersion, connection, and empathy. P1 mentioned that nonverbal cues were extremely helpful, for example, "the moving forward, the eyebrow expression." P3 echoed the movement of the VP:

XSL•FO
RenderX

> *He actually felt kind of connected … there was a point where he kind of leaned back. I felt like it was important to just kind of lean in just a little bit to let him know that I heard and to let him know that he was okay. And I think that that's so valuable and so important. It's kind of given that empathetic piece.*

Similarly, P4 also expressed that they felt empathy and connection toward the VP, highlighting that

> *when he leaned forward, I think maybe it was sort of his way of saying like, 'Oh, this is between me and you, like when he was sharing whether he was feeling depressed.*

P4 further inferred from the movement that "He was afraid that I would tell others. So I think that's why he leaned forward."

On the other hand, there were mixed views on the realism of the VP. P4 mentioned that the distance of the VP in the headset facilitated easier conversation with the VP. They highlighted,

> *I also like that he was actually sitting in front of me and he wasn't too far or too close. I think the distance from where he was to me on the screen was helpful, because it was like it was just us two.*

They also felt that the VP, being not too realistic and an animated avatar, made it easier to engage in conversation in a training setting. They added,

> *I'm like, would that actually be scarier for me, because it feels too real? I'll probably be second guessing my responses, whereas here it felt like I was talking to someone my age, and I feel like, maybe that's good.*

On the other hand, P2 mentioned

> *I was more focused on how his eyes were not moving. It kind of gave me the uncanny valley feeling. I think I was so conscious of trying to maintain eye contact, I didn't notice (the body language of the VP).*

Hence, they requested more realism, specifically in the eye movements.

Flow of Conversation With the VP

Although participants found that the pauses were good in giving space for validation and mimicked what a conversation would be like in the field, more interactivity in responses was desired from the VP. P2 felt that the pauses gave them a chance to acknowledge how difficult the conversation was and recognized that people would appreciate the space and pauses for their feelings to be validated. However, P2 highlighted that

> *I couldn't understand some of his responses. I don't know if they were realistic… After I validated him, all he did was grunt. He never actually said it was hard, so I guess I had to be picking up on that, right? But yeah, I think he just didn't interact that much.*

P3 also echoed, "I was really interested in him elaborating more of what he was experiencing, and then reassuring him that it was okay."

Discussion

Overview

Our findings indicate that the AR simulation demonstrated above-average usability, as reflected in PSSUQ scores and corroborated by qualitative accounts of realism, immersion, and empathetic engagement with the VP. Participants perceived the tool as a promising complement to existing training workflows, and the usability data informed a set of design recommendations organized across four domains: environmental context, training structure, VP behavioral realism, and technical hardware considerations (Multimedia Appendix 7).

To our knowledge, this is among the first studies to investigate the potential of using AR, rather than VR or desktop-based simulation, for training lay MH task-sharing workers with VP simulation. The majority of existing simulation research targets credentialed clinicians or nursing students in controlled settings; our study extends this work to community-based task-sharers who provide frontline MH support in underresourced communities [9,10,39-42]. Furthermore, the tool was co-designed through participatory methods with stakeholders from the HSI, ensuring that the training content and interaction design reflect the cultural and contextual realities of the population served [67,68].

Principal Results

Current Challenges in MH Training

Traditional MH training often grapples with challenges in content simplification due to lack of resources [19,20]. Role-playing has often been used as a method to help trainees develop critical skills such as empathy, communication, and problem-solving [87]. While it is an effective way to simulate real-life scenarios and practice responses, it also presents challenges such as ensuring that role-plays are realistic and emotionally safe for participants, thus requiring skilled facilitators and significant preparation [87]. Echoing the literature, participants mentioned role-play in current training mostly being done between peers. Acting out more challenging scenarios, such as schizophrenia symptoms, was difficult to accurately role-play and might even be left out as peers were not trained actors. Current training often did not use trained actors as providers would require resources like time and money which they mostly do not have [88]. Our findings suggest that adding AR to the current workflow could increase more representative and effective role-play training.

Additionally, trainees may have theoretical knowledge but may require more intensive training in more nuanced practical skills such as reflecting on and interpreting people's feelings and responses [89]. Participants emphasized the struggle of adapting theoretical knowledge to practice with current training methods. They expressed that theoretical knowledge is limited in its effectiveness, and they ultimately had to learn the ropes on the job and develop their own style of counseling. In addition, participants expressed difficulty when interpreting qualitative responses, translating them into a quantitative scale, and identifying how to help people due to individual subjectivity, reflecting their adaptability challenges with current traditional

XSL•FO
RenderX

training. Notably, participants found that the AR training simulation allowed them to practice in a safe and comfortable environment where mistakes were acceptable and did not have detrimental consequences.

A systematic review found that racial and ethnic minority populations faced cultural MH stigma, which led to greater barriers to MH services and lack of trust [90]. Building trust between providers and patients in MH services is important to promote and maintain treatment and engagement [91]. Similarly, participants emphasized that building trust with patients was challenging but important. This type of engagement is difficult to train using traditional methods yet crucial for authentic patient relationships.

Enhancing Training Quality

An integrative review found that realism and debrief were important in MH simulation training and improved trainees' engagement and in bridging the gap between theory and practice [92]. Our study's results highlight that the realism of the AR simulation and the quality of the animations led participants to feel empathy and connection toward the VP. Participants emphasized that they felt like they (the VP and participant) were the only ones in the room, indicating that the AR simulation was successful in immersing participants into a real-life scenario of conversing with a community member. The VP's voice had an accent and was realistic, reflecting the population they would be interacting with, further allowing participants to feel like they were talking to a real person and practice communication. The VP's animation and body language allowed trainees to further empathize with the VP and practice building trust. Through MH training simulation with SPs, trainees could gain empathy, compassion, and confidence, benefiting the people they serve [93]. The AR simulation was realistic, immersive, and sparked empathetic connections, thereby allowing trainees an environment to practice engaging and building trust in a realistic and effective manner.

Additionally, participants highlighted that AR simulation can better depict challenging scenarios, especially specific DSM categories like schizophrenia, depression, and anxiety. They emphasized that a standardized role-play training that best accurately depicts symptoms, rather than peers acting out the symptoms, would be extremely useful for practice. As such, having realistic practice in AR before real-life interactions can build provider efficacy and better care for patients. This kind of applied training is crucial for effective care by providers and MH task-sharers, with studies showing that using AR or VR to practice in a safe environment enhances knowledge and transferable skills, leading to the transfer of knowledge and skills in clinical practice [94]. Thus, this AR-assisted training tool could bridge the gap in training accuracy and enrich current traditional role-playing training.

The Role of Realism and Immersion in AR Training

Overview

Studies have shown that higher realism in immersive virtual experiences leads to a more positive user experience, highlighting that degrees of realism affect the extent of immersion and thus the training's effectiveness [95]. When such training elicits emotional responses similar to those experienced in real-life situations, empathy is found to be nurtured in trainees [36]. Thus, it is important to create a realistic training experience for trainees to develop empathy and immerse fully. In this section, the factors that contribute to realism include the VP's body language, voice, distance from the trainee, appearance, eye movements, and dialogue.

Body Language

Nonverbal cues such as facial and body movements have been shown to facilitate connection with VPs [96]. Participants in our study empathized with the VP and wanted to reassure, validate, and comfort the VP when they perceived distress through the VP's body language. We recommend designing VPs with realistic nonverbal cues to enhance immersion and empathy (Multimedia Appendix 7).

Voice

Virtual assets' realism and appropriate contextualization enhance one's perceptual skills, which are essential for developing sensemaking abilities when using AR [97]. Our study's VP had a real person's voice with an accent, and participants pointed out that this was very useful for training their listening skills, especially for specific populations or communities they may be training for. They added that the headset's technical limitation, that led to the soft volume, was actually useful in training them to listen more attentively to soft-spoken people. These findings emphasize that voice realism in the VP is useful and crucial for participants to hone their sensemaking skills in MH training. Using a representative real person's voice and varying volume levels can train attentive listening skills (Multimedia Appendix 7).

Distance From Trainee

The results highlight that the distance of the VP from the participants was realistic enough, not too close or too far, and that this facilitated easier engagement and immersion in the training. This corroborates with literature findings that training systems which are context-specific enhance learning [97]. Placing the VP at a therapeutically informed comfortable distance enhances immersion (Multimedia Appendix 7) [98].

Appearance

Our findings, also supported with literature, suggest that representing the VP as an animated 3D avatar instead of making the VP too human-like helped in avoiding the uncanny valley effect; this is the discomfort or eeriness experienced when digital avatars closely resemble human beings but fall short in certain aspects of realism [99,100]. Balancing avatar realism to avoid the uncanny valley effect is recommended (Multimedia Appendix 7).

Eye Movements

One participant perceived the VP's eyes as staying still, and it reduced their ability to notice body language cues that were programmed into the training module and broke the immersion for them. The perception of the eyes not moving, even though they were programmed to move, contributed to them experiencing the uncanny valley effect. This highlights the importance of making the eye movements more realistic and

XSL•FO
RenderX

obvious to ensure that the uncanny valley effect is minimized and to increase immersion, especially when gaze-tracking is correlated to trainees' assessment skills [97]. Realistic and obvious eye movements are essential for maintaining immersion (Multimedia Appendix 7).

Dialogue

Although pauses in dialogue were good in giving space for validation and mimicked a real conversation in the field, more interactivity in the VP's responses was desired by participants. In addition, one of the participants felt like the responses were lacking in realism; for example, the VP grunting as a response. These findings align with evidence that dynamic, contextually responsive VP dialogue is crucial to perceived realism in simulation-based training [21,97]. More interactive, dynamic VP responses would enhance conversational realism (Multimedia Appendix 7).

Immersion was also affected by the structure of the training. This included the use of prompts, which have been found to be useful for guiding learning [97]. Although the prompts were useful in guiding participants on how to respond when they were unsure, navigating between two types of prompts confused users and reduced immersion. Switching between a general prompt and a specific word-for-word prompt (Figure 7), paying attention to the VP's body language and dialogue, and asking the questions from the PHQ-4 form broke the flow of immersion and caused confusion in participants at times. Clearly distinguishing prompt types and gradually transitioning from guided to unguided interactions would improve training flow (Multimedia Appendix 7).

Overall Usability and Acceptability of AR Tool

Our findings show that the AR-assisted training tool is useful for MH training and has the potential to scale to different MH contexts, especially more complex scenarios. Qualitative results highlight the overwhelming positive sentiments toward the AR tool and its promising potential in successfully filling the gaps of current training workflows. The overall PSSUQ scores indicated above-average usability, consistent with patterns reported in comparable simulation-based training evaluations [92,101]. The interface quality subscale performed particularly well, suggesting that the AR tool's interface met user expectations for functionality and pleasantness, an outcome that aligns with findings that higher-fidelity immersive environments tend to yield stronger user acceptance [36,95]. High reported desire to use the system frequently further supports the tool's acceptability as a complement to existing training workflows.

The main challenges of integrating this tool into the wider MH training community are the technology learning curve and hardware limitations. Two out of the three first-time users of AR and VR headsets experienced difficulty at the start. Although they expressed that they were overwhelmed by all the different things to pay attention to, they also highlighted that with more practice, they would adjust to the training and benefit from it. This is a positive outcome since AR and VR headsets are known to have a steep learning curve [101]. The two participants who raised this concern were in the 35- to 40- and 56- to 60-year-old range, respectively, which may indicate that the older demographic, who tend to have a lower technology learnability rate, could find this AR tool to have a manageable learning curve [102]. We recommend keeping this challenge in mind and working on building a seamless onboarding process when introducing a similar training tool.

Another challenge for integration is the headset's hardware limitations, including the headset falling down due to a missing accessory pack for supporting a range of nose bridges, the headset heating up, and low audibility even on the headset's maximum volume. This was an internal software issue that could be fixed in hindsight. The headset falling down affected the PSSUQ scores, particularly as this discomfort led to breaking participants' immersion. A study using similar AR HMDs also mentioned that issues such as user discomfort and technical limitations of HMDs need to be addressed before widespread adoption in the health care industry [103]. Thus, these challenges must be addressed for a successful implementation of this tool. We recommend addressing hardware fit, comfort, and cost challenges; integrating familiar training workflows; incorporating personalized feedback mechanisms; and enabling VP response replay (Multimedia Appendix 7).

Limitations

This study has several limitations that should be considered when interpreting the findings. First, as an exploratory study, the work reported here involved a limited number of participants, making it difficult to draw generalizable conclusions. Due to the small sample size, a descriptive analysis was used instead of a test of significance. The selection of the "N/A" option affected the quantitative results. To minimize this, an additional instruction could have been given to participants before they filled out the quantitative survey to remind them that if they did not understand the question, they could clarify with researchers. Due to the attrition of one participant (P5) and the small pilot sample size, thematic saturation was not reached; however, the data collected from the remaining four participants provided consistent insights into the tool's usability. This study was intended to understand usability and user perceptions rather than to evaluate training effectiveness or learning outcomes. The promising results indicate the need for future studies, such as collecting quantitative data from a larger sample size and measuring trainees' actual performance after AR-assisted training.

Second, the hardware challenges might have affected the data. Participant satisfaction with AR training has been shown to vary depending on technical factors [104]. Comfort issues such as overheating from low battery life, accommodating glasses wearers (which caused a participant to drop out of the usability test), and the lack of an adaptive fit for certain nose bridges and head sizes could not be overlooked. The hardware costs also presented challenges for the scalability of being widely used in the community.

Third, the study was conducted in a single site; hence, our results may lack transferability to other settings with different parameters, which is an especially important factor to consider for educational learning interventions [105]. The investigation was also context-specific, which may not be generalizable to screening for other MH conditions. Using more in-depth scales

XSL•FO
RenderX

to test realism and immersion and the relationship it has with effectiveness in a different experiment would also further enhance this study.

Conclusions

This pilot study provides promising preliminary evidence that AR-based VP simulation is a usable and engaging training modality for MH task-sharers. The study is among the first to apply AR glasses with VP simulation to this population, filling the gaps of past studies that have predominantly used VR or desktop-based systems with credentialed clinicians and nursing students. This pilot study brings preliminary empirical usability evidence from a validated instrument (PSSUQ) demonstrating above-average usability, and a set of interaction design recommendations spanning environmental context, training structure, VP behavioral realism, and technical hardware considerations. These findings show promising results that AR-based simulation could enable MH task-sharers in underresourced community settings to access realistic, culturally tailored training practice at a lower financial, labor, and time cost, directly addressing the resource constraints of current training approaches. Future research should explore cost-effective hardware alternatives that address comfort and accessibility challenges, increase sample sizes for inferential analysis, compare AR-assisted training simulations with traditional methods, and extend the tool to additional MH conditions for greater generalizability and scalability.

Acknowledgments

The authors thank Rajshree Saraf and Sounak Ghosh for their support in the development of the Augmented Reality-Assisted Training Tool, Victoria Ngo, PhD, and Marina Weiss, PhD, for their clinical expertise and review of mental health content, and all the community stakeholders who took part in our study. The authors declare the use of generative artificial intelligence (AI) in the writing process. According to the GAIDeT taxonomy (2025), the following task was delegated to a generative AI tool under full human supervision: proofreading and editing. The tool used was Claude (Anthropic). No AI-generated content was included in the manuscript, and all scientific content, analysis, and interpretations were produced and verified by the authors. Responsibility for the final manuscript lies entirely with the authors.

Funding

This study was cofunded by the CUNY Graduate School of Public Health and Health Policy and the Harlem Strong Mental Health Coalition: A Multisector Community-Engaged Collaborative for System Transformation (5U01MH136557). The funding agencies were not involved in designing this study, sampling and recruitment, data collection, data interpretation, or writing this paper.

Data Availability

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

Conflicts of Interest

None declared.

Multimedia Appendix 1
ARATT demo.
[MP4 File, 47020 KB - xr_v3i1e80711_app1.mp4]

Multimedia Appendix 2
Screening Scenario script.
[PDF File, 24 KB - xr_v3i1e80711_app2.pdf]

Multimedia Appendix 3
Semistructured interview questions.
[PDF File, 44 KB - xr_v3i1e80711_app3.pdf]

Multimedia Appendix 4
Qualtrics questionnaire.
[PDF File, 197 KB - xr_v3i1e80711_app4.pdf]

Multimedia Appendix 5
Summary of key themes.
[PDF File, 76 KB - xr_v3i1e80711_app5.pdf]

XSL•FO
RenderX

Multimedia Appendix 6
Quantitative responses table.
[PDF File, 39 KB - xr_v3i1e80711_app6.pdf]

Multimedia Appendix 7
Summary of design recommendations.
[PDF File, 80 KB - xr_v3i1e80711_app7.pdf]

References

1. König H, König HH, Konnopka A. The excess costs of depression: a systematic review and meta-analysis. Epidemiol Psychiatr Sci 2020;29. [doi: 10.1017/S2045796019000180]
2. Chisholm D, Sweeny K, Sheehan P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. Lancet Psychiatry 2016 May;3(5):415-424. [doi: 10.1016/S2215-0366(16)30024-4]
3. Eylem O, de Wit L, van Straten A, et al. Stigma for common mental disorders in racial minorities and majorities a systematic review and meta-analysis. BMC Public Health 2020 Dec;20(1):879. [doi: 10.1186/s12889-020-08964-3]
4. Roy-Byrne PP, Joesch JM, Wang PS, Kessler RC. Low socioeconomic status and mental health care use among respondents with anxiety and depression in the NCS-R. Psychiatr Serv 2009 Sep;60(9):1190-1197. [doi: 10.1176/ps.2009.60.9.1190] [Medline: 19723733]
5. McGuire TG, Miranda J. New evidence regarding racial and ethnic disparities in mental health: policy implications. Health Aff (Millwood) 2008;27(2):393-403. [doi: 10.1377/hlthaff.27.2.393] [Medline: 18332495]
6. Ranjbar N, Erb M, Mohammad O, Moreno FA. Trauma-informed care and cultural humility in the mental health care of people from minoritized communities. Focus (Am Psychiatr Publ) 2020 Jan;18(1):8-15. [doi: 10.1176/appi.focus.20190027] [Medline: 32047392]
7. Gopalkrishnan N. Cultural diversity and mental health: considerations for policy and practice. Front Public Health 2018;6(179):179. [doi: 10.3389/fpubh.2018.00179] [Medline: 29971226]
8. Eken HN, Dee EC, Powers AR III, Jordan A. Racial and ethnic differences in perception of provider cultural competence among patients with depression and anxiety symptoms: a retrospective, population-based, cross-sectional analysis. Lancet Psychiatry 2021 Nov;8(11):957-968. [doi: 10.1016/S2215-0366(21)00285-6]
9. Ngo VK, Weiss B, Lam T, Dang T, Nguyen T, Nguyen MH. The Vietnam Multicomponent Collaborative Care for Depression Program: development of depression care for low- and middle-income nations. Journal of Cognitive Psychotherapy 2014;28(3):156-167. [doi: 10.1891/0889-8391.28.3.156] [Medline: 25568593]
10. Le PD, Eschliman EL, Grivel MM, et al. Barriers and facilitators to implementation of evidence-based task-sharing mental health interventions in low- and middle-income countries: a systematic review using implementation science frameworks. Implement Sci 2022 Jan 12;17(1):4. [doi: 10.1186/s13012-021-01179-z] [Medline: 35022081]
11. Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. JAMA 2000 Jan 12;283(2):212-220. [doi: 10.1001/jama.283.2.212] [Medline: 10634337]
12. Ngo VK, Sherbourne C, Chung B, et al. Community engagement compared with technical assistance to disseminate depression care among low-income, minority women: a randomized controlled effectiveness study. Am J Public Health 2016 Oct;106(10):1833-1841. [doi: 10.2105/AJPH.2016.303304] [Medline: 27552274]
13. World Health Organization. Education and Training: World Health Organization; 2016. URL: https://iris.who.int/handle/10665/252271 [accessed 2025-02-05]
14. Caulfield A, Vatansever D, Lambert G, Van Bortel T. WHO guidance on mental health training: a systematic review of the progress for non-specialist health workers. BMJ Open 2019 Feb 1;9(1):e024059. [doi: 10.1136/bmjopen-2018-024059] [Medline: 30782724]
15. Richards DA, Ekers D, McMillan D, et al. Cost and outcome of behavioural activation versus cognitive behavioural therapy for depression (COBRA): a randomised, controlled, non-inferiority trial. Lancet 2016 Aug 27;388(10047):871-880. [doi: 10.1016/S0140-6736(16)31140-0] [Medline: 27461440]
16. Jacobs Y, Myers B, van der Westhuizen C, Brooke-Sumner C, Sorsdahl K. Task sharing or task dumping: counsellors experiences of delivering a psychosocial intervention for mental health problems in South Africa. Community Ment Health J 2021 Aug;57(6):1082-1093. [doi: 10.1007/s10597-020-00734-0] [Medline: 33161458]
17. Harvey PD, Khan A, Keefe RSE. Using the Positive and Negative Syndrome Scale (PANSS) to define different domains of negative symptoms. Innov Clin Neurosci 2017 Dec 1;14(11–12):18-22. [Medline: 29410933]
18. Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. AAMC. Acad Med 1993 Jun;68(6):443-451. [doi: 10.1097/00001888-199306000-00002] [Medline: 8507309]
19. Adib-Hajbaghery M, Sharifi N. Effect of simulation training on the development of nurses and nursing students' critical thinking: a systematic literature review. Nurse Educ Today 2017 Mar;50:17-24. [doi: 10.1016/j.nedt.2016.12.011] [Medline: 28011333]

XSL•FO
RenderX

20. Williams B, Reddy P, Marshall S, Beovich B, McKarney L. Simulation and mental health outcomes: a scoping review. Adv Simul 2017 Dec;2(1):2. [doi: 10.1186/s41077-016-0035-9]
21. Talbot T, Rizzo A. Virtual human standardized patients for clinical training. In: Rizzo A, Bouchard S, editors. Virtual Reality for Psychological and Neurocognitive Interventions: Springer; 2019:387-405. [doi: 10.1007/978-1-4939-9482-3_17]
22. Milgram P, Takemura H, Utsumi A, Kishino F. Augmented reality: a class of displays on the reality-virtuality continuum. 1995 Presented at: Proc SPIE 2351, Telemanipulator and Telepresence Technologies; Oct 31 to Nov 1, 1994; Boston, MA p. 282-292. [doi: 10.1117/12.197321]
23. Barsom EZ, Graafland M, Schijven MP. Systematic review on the effectiveness of augmented reality applications in medical training. Surg Endosc 2016 Oct;30(10):4174-4183. [doi: 10.1007/s00464-016-4800-6] [Medline: 26905573]
24. McCall R, Wetzel R, Löschner J, Braun AK. Using presence to evaluate an augmented reality location aware game. Pers Ubiquit Comput 2011 Jan;15(1):25-35. [doi: 10.1007/s00779-010-0306-8]
25. Otono R, Genay A, Perusquia-Hernandez M, et al. Studying “avatar transitions” in augmented reality: influence on sense of embodiment and physiological activity. 2022 Presented at: 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct); Oct 17-21, 2022; Singapore p. 503-504. [doi: 10.1109/ISMAR-Adjunct57072.2022.00106]
26. Jacobs R, Schnädelbach H, Jäger N, et al. The performative mirror space. 2019 May 2 Presented at: CHI ’19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland p. 1-14 URL: https://dl.acm.org/doi/proceedings/10.1145/3290605 [doi: 10.1145/3290605.3300630]
27. Hoang TN, Ferdous HS, Vetere F, Reinoso M. Body as a canvas: an exploration on the role of the body as display of digital information. 2018 Presented at: DIS ’18: Designing Interactive Systems Conference 2018; Jun 9-13, 2018; Hong Kong, China p. 253-263. [doi: 10.1145/3196709.3196724]
28. Botvinick M, Cohen J. Rubber hands ‘feel’ touch that eyes see. Nature New Biol 1998 Feb 19;391(6669):756-756. [doi: 10.1038/35784]
29. IJsselsteijn WA, de Kort YAW, Haans A. Is this my hand I see before me? the rubber hand illusion in reality, virtual reality, and mixed reality. Presence: Teleoperators and Virtual Environments 2006 Aug 1;15(4):455-464. [doi: 10.1162/pres.15.4.455]
30. Yuan Y, Steed A. Is the rubber hand illusion induced by immersive virtual reality? Presented at: VR ’10: Proceedings of the 2010 IEEE Virtual Reality Conference; Mar 20-24, 2010. [doi: 10.1109/VR.2010.5444807]
31. Javornik A, Rogers Y, Gander D, Moutinho A. MagicFace: stepping into character through an augmented reality mirror. 2017 Presented at: CHI ’17: CHI Conference on Human Factors in Computing Systems; May 6-11, 2017; Denver, CO p. 4838-4849. [doi: 10.1145/3025453.3025722]
32. Scavarelli A, Arya A, Teather RJ. Virtual reality and augmented reality in social learning spaces: a literature review. Virtual Real 2021 Mar;25(1):257-277. [doi: 10.1007/s10055-020-00444-8]
33. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development: Pearson Education; 2014. URL: https://books.google.com/books?id=jpbeBQAAQBAJ
34. Verkuyl M, Romaniuk D, Mastrilli P. Virtual gaming simulation of a mental health assessment: a usability study. Nurse Educ Pract 2018 Jul;31:83-87. [doi: 10.1016/j.nepr.2018.05.007] [Medline: 29800764]
35. Asad MM, Naz A, Churi P, Tahanzadeh MM. Virtual reality as pedagogical tool to enhance experiential learning: a systematic literature review. Education Research International 2021 Nov 16;2021(1):1-17. [doi: 10.1155/2021/7061623]
36. Li Y, Luo H, Zhang C. 2021 Presented at: Hawaii International Conference on System Sciences; Jan 5-8, 2021 URL: http://hdl.handle.net/10125/71104
37. Lohr DJ, Johnson S, Aziz S, Komogortsev O. Demonstrating eye movement biometrics in virtual reality. Presented at: Proceedings of the 2023 Symposium on Eye Tracking Research and Applications (ETRA '23); May 30 to Jun 2, 2023. [doi: 10.1145/3588015.3590109]
38. Concannon BJ, Esmail S, Roduta Roberts M. Head-mounted display virtual reality in post-secondary education and skill training. Front Educ 2019 Aug 14;4. [doi: 10.3389/feduc.2019.00080]
39. Choudhury N, Gélinas-Phaneuf N, Delorme S, Del Maestro R. Fundamentals of neurosurgery: virtual reality tasks for training and evaluation of technical skills. World Neurosurg 2013 Nov;80(5):e9-19. [doi: 10.1016/j.wneu.2012.08.022] [Medline: 23178917]
40. Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. Ann Surg 2005 Feb;241(2):364-372. [doi: 10.1097/01.sla.0000151982.85062.80] [Medline: 15650649]
41. Verkuyl M, Mastrilli P. Virtual simulations in nursing education: a scoping review. Research & Reviews: Journal of Nursing and Health Sciences 2017 Aug 12 [FREE Full text]
42. Johnsen HM, Fossum M, Vivekananda-Schmidt P, Fruhling A, Slettebø Å. Teaching clinical reasoning and decision-making skills to nursing students: design, development, and usability evaluation of a serious game. Int J Med Inform 2016 Oct;94:39-48. [doi: 10.1016/j.ijmedinf.2016.06.014] [Medline: 27573310]
43. Wu B, Yu X, Gu X. Effectiveness of immersive virtual reality using head‐mounted displays on learning performance: a meta‐analysis. Brit J Educational Tech 2020 Nov;51(6):1991-2005. [doi: 10.1111/bjet.13023]
44. Hollis RB, Was CA. Mind wandering, control failures, and social media distractions in online learning. Learn Instr 2016 Apr;42:104-112. [doi: 10.1016/j.learninstruc.2016.01.007]

XSL•FO
RenderX

45. Dede CJ, Jacobson J, Richards J. Introduction: virtual, augmented, and mixed realities in education. In: Richards J, Dede C, Huang R, Richards J, editors. Virtual, Augmented, and Mixed Realities in Education: Springer; 2017:1-16. [doi: 10.1007/978-981-10-5490-7_1]
46. Slater M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos Trans R Soc Lond B Biol Sci 2009 Dec 12;364(1535):3549-3557. [doi: 10.1098/rstb.2009.0138] [Medline: 19884149]
47. Ahn S(, Bailenson JN, Park D. Short- and long-term effects of embodied experiences in immersive virtual environments on environmental locus of control and behavior. Comput Human Behav 2014 Oct;39:235-245. [doi: 10.1016/j.chb.2014.07.025]
48. Sheridan TB. Human–robot interaction. Hum Factors 2016 Jun;58(4):525-532. [doi: 10.1177/0018720816644364]
49. Lee Y, Kim SK, Eom MR. Usability of mental illness simulation involving scenarios with patients with schizophrenia via immersive virtual reality: a mixed methods study. PLoS One 2020;15(9):e0238437. [doi: 10.1371/journal.pone.0238437]
50. Herbst R, Rybak T, Meisman A, et al. A virtual reality resident training curriculum on behavioral health anticipatory guidance: development and usability study. JMIR Pediatr Parent 2021 Jun 29;4(2):e29518. [doi: 10.2196/29518] [Medline: 34081601]
51. Lawson BD, Stanney KM. Editorial: cybersickness in virtual reality and augmented reality. Front Virtual Real 2021 Oct 1;2:759682. [doi: 10.3389/frvir.2021.759682]
52. Wu HK, Lee SWY, Chang HY, Liang JC. Current status, opportunities and challenges of augmented reality in education. Comput Educ 2013 Mar;62:41-49. [doi: 10.1016/j.compedu.2012.10.024]
53. Tang KS, Cheng DL, Mi E, Greenberg PB. Augmented reality in medical education: a systematic review. Can Med Educ J 2020 Mar;11(1):e81-e96. [doi: 10.36834/cmej.61705] [Medline: 32215146]
54. Lee LN, Kim MJ, Hwang WJ. Potential of augmented reality and virtual reality technologies to promote wellbeing in older adults. Appl Sci (Basel) 2019 Jan;9(17):3556. [doi: 10.3390/app9173556]
55. Yeung AWK, Tosevska A, Klager E, et al. Virtual and augmented reality applications in medicine: analysis of the scientific literature. J Med Internet Res 2021 Feb 10;23(2):e25499. [doi: 10.2196/25499] [Medline: 33565986]
56. Juanes-Méndez JA, Marcos-Pablos S, Velasco-Marcos MA, Izard SG. Medical learning through simulations with immersive virtual reality and augmented reality devices. In: Technological Adoption and Trends in Health Sciences Teaching, Learning, and Practice: IGI Global; 2022. URL: https://www.semanticscholar.org/paper/Medical-Learning-Through-Simulations-With-Immersive-Juanes-M%C3%A9ndez-Marcos-Pablos/e1925e4c95506b8537803898934818402a79b4b [accessed 2025-02-06]
57. Kamphuis C, Barsom E, Schijven M, Christoph N. Augmented reality in medical education? Perspect Med Educ 2014 Sep;3(4):300-311. [doi: 10.1007/s40037-013-0107-7] [Medline: 24464832]
58. Zhu E, Hadadgar A, Masiello I, Zary N. Augmented reality in healthcare education: an integrative review. PeerJ 2014;2:e469. [doi: 10.7717/peerj.469] [Medline: 25071992]
59. Albrecht UV, Folta-Schoofs K, Behrends M, von Jan U. Effects of mobile augmented reality learning compared to textbook learning on medical students: randomized controlled pilot study. J Med Internet Res 2013 Aug 20;15(8):e182. [doi: 10.2196/jmir.2497] [Medline: 23963306]
60. Siebert JN, Ehrler F, Gervaix A, et al. Adherence to AHA guidelines when adapted for augmented reality glasses for assisted pediatric cardiopulmonary resuscitation: a randomized controlled trial. J Med Internet Res 2017 May 29;19(5):e183. [doi: 10.2196/jmir.7379] [Medline: 28554878]
61. Kotcherlakota S, Pelish P, Hoffman K, Kupzyk K, Rejda P. Augmented reality technology as a teaching strategy for learning pediatric asthma management: mixed methods study. JMIR Nurs 2020 Dec 2;3(1):e23963. [doi: 10.2196/23963] [Medline: 34406970]
62. Ingrassia PL, Mormando G, Giudici E, et al. Augmented reality learning environment for basic life support and defibrillation training: usability study. J Med Internet Res 2020 May 12;22(5):e14910. [doi: 10.2196/14910] [Medline: 32396128]
63. Zackoff MW, Cruse B, Sahay RD, et al. Development and implementation of augmented reality enhanced high-fidelity simulation for recognition of patient decompensation. Simul Healthc 2021 Jun 1;16(3):221-230. [doi: 10.1097/SIH.0000000000000486] [Medline: 32910102]
64. Bracq MS, Michinov E, Jannin P. Virtual reality simulation in nontechnical skills training for healthcare professionals: a systematic review. Simul Healthc 2019 Jun;14(3):188-194. [doi: 10.1097/SIH.0000000000000347] [Medline: 30601464]
65. Ward P, Gore J, Hutton R, Conway GE, Hoffman RR. Adaptive skill as the conditio sine qua non of expertise. J Appl Res Mem Cogn 2018;7(1):35-50. [doi: 10.1016/j.jarmac.2018.01.009]
66. Nielsen J. Usability Engineering: Morgan Kaufmann Publishers Inc; 1994.
67. Strengthening community mental health resilience in Harlem. Harlem Strong. URL: https://harlem-strong.org/ [accessed 2025-02-06]
68. Center for innovation in mental health at the CUNY graduate school of public health and health policy. Center for Innovation in Mental Health. URL: https://cimh.sph.cuny.edu/ [accessed 2025-02-06]
69. Ericsson KA, Simon HA. Protocol analysis: verbal reports as data (revised edition). MIT Press Direct. URL: https://direct.mit.edu/books/monograph/4763/Protocol-AnalysisVerbal-Reports-as-Data [accessed 2025-02-05]

XSL•FO
RenderX

70. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int J Hum Comput Interact 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]
71. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ–4. Psychosomatics 2009 Nov;50(6):613-621. [doi: 10.1016/S0033-3182(09)70864-3]
72. Care. Magic Leap 2. URL: https://www.magicleap.com/magic-leap-2 [accessed 2025-02-05]
73. WAI site usability testing questions. Web Accessibility Initiative (WAI). URL: https://www.w3.org/WAI/EO/Drafts/UCD/questions.html [accessed 2025-02-05]
74. Mosadeghi S, Reid MW, Martinez B, Rosen BT, Spiegel BMR. Feasibility of an immersive virtual reality intervention for hospitalized patients: an observational cohort study. JMIR Ment Health 2016 Jun 27;3(2):e28. [doi: 10.2196/mental.5801] [Medline: 27349654]
75. Writing your user test plan: open-ended vs. specific tasks and questions. User Testing. URL: https://www.usertesting.com/blog/open-ended-vs-specific-tasks-and-questions [accessed 2025-02-05]
76. Unity: develop, deploy, and grow | the world's leading 3D game engine. Unity. URL: https://unity.com/ [accessed 2024-03-12]
77. Integrate an avatar creator into your game in days [archived version]. Ready Player Me. URL: https://web.archive.org/web/20251222042725/https://readyplayer.me/ [accessed 2024-03-12]
78. SALSA lipsync suite. Unity Asset Store. URL: https://assetstore.unity.com/packages/tools/animation/salsa-lipsync-suite-148442 [accessed 2024-03-12]
79. Animate 3D characters for games, film, and more. Mixamo. URL: https://www.mixamo.com/#/ [accessed 2024-03-12]
80. One platform to connect. Zoom. URL: https://www.zoom.com [accessed 2025-02-05]
81. A brief history of ATLAS.ti. ATLAS.ti. URL: https://atlasti.com/about-atlas-ti [accessed 2025-02-05]
82. Lincoln YS, Guba EG, Pilotta JJ. Naturalistic inquiry. Int J Intercult Relat 1985 Jan;9(4):438-439. [doi: 10.1016/0147-1767(85)90062-8]
83. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
84. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
85. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
86. Hess O, Qian J, Bruce J, et al. Communication skills training using remote augmented reality medical simulation: a feasibility and acceptability qualitative study. Med Sci Educ 2022 Oct;32(5):1005-1014. [doi: 10.1007/s40670-022-01598-7] [Medline: 35966166]
87. Ayer L, Schultz D, Abbott M, et al. Mental health task-shifting in community-based organizations: implementation, impact, and cost — evaluation of the connections to care program. : RAND Corporation; 2020. [doi: 10.7249/RR3083]
88. Dobmeyer AC, Hunter CL, Corso ML, et al. Primary care behavioral health provider training: systematic development and implementation in a large medical system. J Clin Psychol Med Settings 2016 Sep;23(3):207-224. [doi: 10.1007/s10880-016-9464-9] [Medline: 27484777]
89. Hashemiparast M, Negarandeh R, Theofanidis D. Exploring the barriers of utilizing theoretical knowledge in clinical settings: a qualitative study. Int J Nurs Sci 2019 Oct 10;6(4):399-405. [doi: 10.1016/j.ijnss.2019.09.008] [Medline: 31728392]
90. Misra S, Jackson VW, Chong J, et al. Systematic review of cultural aspects of stigma and mental illness among racial and ethnic minority groups in the United States: implications for interventions. Am J Community Psychol 2021 Dec;68(3-4):486-512. [doi: 10.1002/ajcp.12516] [Medline: 33811676]
91. Gaebel W, Muijen M, Baumann AE, et al. EPA guidance on building trust in mental health services. Eur psychiatr 2014 Feb;29(2):83-100. [doi: 10.1016/j.eurpsy.2014.01.001]
92. Alexander L, Coyte B, Evans A, Dickson J, Guinea S, Foster K. The educational value of mental health simulation in undergraduate bachelor of nursing education: an integrative review. Clin Simul Nurs 2023 Nov;84:101459. [doi: 10.1016/j.ecns.2023.101459]
93. Knutson de Presno Å, Øgård-Repål A, Fossum M. Simulations with standardized patients for nursing students in preparation for clinical placements in mental health care. Clin Simul Nurs 2021 May;54:70-76. [doi: 10.1016/j.ecns.2021.01.009]
94. Gasteiger N, van der Veer SN, Wilson P, Dowding D. How, for whom, and in which contexts or conditions augmented and virtual reality training works in upskilling health care workers: realist synthesis. JMIR Serious Games 2022 Feb 14;10(1):e31644. [doi: 10.2196/31644] [Medline: 35156931]
95. Gonçalves G, Coelho H, Monteiro P, Melo M, Bessa M. Systematic review of comparative studies of the impact of realism in immersive virtual experiences. ACM Comput Surv 2023 Jul 31;55(6):1-36. [doi: 10.1145/3533377]
96. Kasapakis V, Dzardanova E, Nikolakopoulou V, Vosinakis S, Xenakis I, Gavalas D. Social virtual reality: implementing non-verbal cues in remote synchronous communication. In: Bourdot P, Alcañiz Raya M, Figueroa P, Interrante V, Kuhlen TW, Reiners D, editors. Virtual Real Mix Real Cham: Springer International Publishing; 2021:152-157. [doi: 10.1007/978-3-030-90739-6_10]
97. Sushereba CE, Militello LG, Wolf S, Patterson ES. Use of augmented reality to train sensemaking in high-stakes medical environments. J Cogn Eng Decis Mak 2021 Sep;15(2-3):55-65. [doi: 10.1177/15553434211019234]

XSL•FO
RenderX

98. Egozi S, Tishby O, Wiseman H. Therapeutic distance in client-therapist narratives: client attachment, therapist attachment, and dyadic effects. Psychother Res 2021 Nov;31(8):963-976. [doi: 10.1080/10503307.2021.1874069] [Medline: 33471634]
99. Mori M, MacDorman KF, Kageki N. The uncanny valley [from the field]. IEEE Robot Automat Mag 2012 Jun;19(2):98-100. [doi: 10.1109/MRA.2012.2192811]
100. MacDorman KF, Ishiguro H. The uncanny advantage of using androids in cognitive and social science research. IS 2006 Nov 13;7(3):297-337. [doi: 10.1075/is.7.3.03mac]
101. Mohamad Zainal NH, Wahid HH, Mahmud M, et al. The applications of augmented reality (AR) and virtual reality (VR) in teaching medical and dentistry students: a review on advantages and disadvantages. MJMHS 2023 Nov 20;19(s12):65-78. [doi: 10.47836/mjmhs.19.s12.9]
102. Czaja SJ, Charness N, Fisk AD, et al. Factors predicting the use of technology: findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). Psychol Aging 2006 Jun;21(2):333-352. [doi: 10.1037/0882-7974.21.2.333] [Medline: 16768579]
103. Rüger C, Feufel MA, Moosburner S, Özbek C, Pratschke J, Sauer IM. Ultrasound in augmented reality: a mixed-methods evaluation of head-mounted displays in image-guided interventions. Int J CARS 2020 Nov;15(11):1895-1905. [doi: 10.1007/s11548-020-02236-6]
104. Baashar Y, Alkawsi G, Ahmad WNW, et al. Effectiveness of using augmented reality for training in the medical professions: meta-analysis. JMIR Serious Games 2022 Jul 5;10(3):e32715. [doi: 10.2196/32715] [Medline: 35787488]
105. Hoadley C, Campos FC. Design-based research: what it is and why it matters to studying online learning. Educ Psychol 2022 Jul 3;57(3):207-220. [doi: 10.1080/00461520.2022.2079128]

Abbreviations

APA: American Psychiatric Association
AR: augmented reality
CIMH: Center for Innovation in Mental Health
DSM: *Diagnostic and Statistical Manual of Mental Disorders*
HMD: head-mounted display
HSI: Harlem Strong Initiative
INFOQUAL: information quality
INTERQUAL: interface quality
MH: mental health
MR: mixed reality
NDA: National Institute of Mental Health Data Archive
PHQ-4: Patient Health Questionnaire-4
PSSUQ: Post-Study System Usability Questionnaire
SP: standardized patient
SYSUSE: system usefulness
VP: virtual patient
VR: virtual reality

Edited by S Brini, T Leung; submitted 28.Jul.2025; peer-reviewed by C Munsch, C Johnson; revised version received 16.Mar.2026; accepted 17.Mar.2026; published 25.Jun.2026.

Please cite as:
Ngiam LLV, Wozniak LP, Dinh-Le C, Seals AE, Ngo VK, Florez-Arango JF
Augmented Reality–Assisted Training Tool for Mental Health Task-Sharers: Pilot Mixed Methods Usability Study
JMIR XR Spatial Comput 2026;3:e80711
URL: https://xr.jmir.org/2026/1/e80711
doi:10.2196/80711

© Ling Li Vivian Ngiam, Lucas Piotr Wozniak, Catherine Dinh-Le, Ayanna Elon Seals, Victoria K Ngo, Jose Fernando Florez-Arango. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 25.Jun.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Experience and Socioemotional Evaluation of Virtual Therapist Character Types in a Simulated Virtual Reality Exposure Therapy: Quantitative Study

Michelle Celina Hallmann[1], BSc; Nicolina Laura Peperkorn[2], MSc; Francesco Vona[1], PhD; Youssef Shiban[2], PhD; Jan-Niklas Voigt-Antons[1], PhD

[1]Immersive Reality Lab, Hamm-Lippstadt University of Applied Sciences, Marker-Allee 76-78, Hamm, Germany
[2]Department of Clinical Psychology, PFH Private University of Applied Sciences, Göttingen, Germany

Corresponding Author:
Jan-Niklas Voigt-Antons, PhD
Immersive Reality Lab, Hamm-Lippstadt University of Applied Sciences, Marker-Allee 76-78, Hamm, Germany

Abstract

Background: Virtual reality exposure therapy (VRET) has emerged as an effective treatment for individuals with specific phobias, particularly for those unable to access traditional exposure therapy. However, several aspects of VRET applications still require further investigation, including the role of embodied conversational agents as virtual therapists.

Objective: This study aims to investigate how the realism level and species of a virtual therapist influence user experience and perceived socioemotional relationship skills in a VRET intervention for dog phobia.

Methods: We conducted a quantitative study that examined user experience and perceived socioemotional relationship skills of virtual therapist character types, where participants were recruited from a convenience sample and were not required to have a dog phobia. In 4 simulated VRET interventions, participants interacted with 4 virtual therapists through a role-play dialogue (3.5 minutes per scenario), followed by an exposure scenario with a dog. The 4 virtual character types included a realistic adult woman, a stylized adult woman, a stylized cat, and a stylized robot. We collected quantitative data through standardized questionnaires assessing user experience and presence, cybersickness, trust, and willingness to share information, and an exploratory, theory-informed questionnaire assessing the virtual therapist's socioemotional relationship skills.

Results: A total of 40 participants completed the experiment. The realistic human scored significantly lower in hedonic quality than the stylized human (z=2.53; P=.01; P_{adj}=.02; rbb=−0.54), the robot (z=2.726; P=.006; P_{adj}=.01; rbb=−0.56), and the cat (z=3.772; P<.001; P_{adj}=.004; rbb=−0.81). Additionally, the cat scored higher in hedonic quality than the stylized human (z=2.528; P=.02; P_{adj}=.02; rbb=−0.52). While no significant differences were found among the virtual character types regarding trust or willingness to share information, the therapeutic relationship skills of the robot were rated significantly lower than those of the stylized human (z=−2.315; P=.02; P_{adj}=.02; rbb=−0.42). However, when asked to indicate a preferred virtual therapist, 30% (12/40) of the participants chose the robot, 27.5% (11/40) the stylized human, 25% (10/40) the cat, and 17.5% (7/40) the realistic human, making the latter the least favored preferred option in our study sample.

Conclusions: The realistic human was the least preferred virtual therapist, whereas the stylized human received more favorable evaluations. This finding suggests that increased visual realism does not necessarily result in more positive perceptions of virtual therapists and may, in some cases, lead to less favorable user responses. The cat's higher hedonic quality and the robot's popularity suggest that user perceptions may vary across different virtual therapist embodiments. These findings underscore the need for further research on virtual therapist design in VRET.

(JMIR XR Spatial Comput 2026;3:e81537) doi:10.2196/81537

KEYWORDS

embodied conversational agent; virtual character; virtual therapist; Uncanny Valley; virtual reality exposure therapy

Introduction

Background

Specific phobias are among the most common mental disorders, with lifetime prevalence estimates ranging from 7.7% to 12.5% among adults [1]. Individuals with specific phobias experience intense, irrational fear responses and exhibit avoidance behavior toward a specific object or situation [2]. Exposure therapy, particularly in vivo treatment, is the preferred method for helping individuals replace maladaptive fear responses with adaptive reactions [3]. Exposure therapy for specific phobias

XSL•FO
RenderX

typically begins with psychoeducation, during which patients learn how phobias develop and persist through avoidance behavior, and which fears are rational or irrational [4,5]. Patients are then systematically confronted with the feared object or situation, enabling cognitive restructuring by reducing maladaptive behaviors and replacing irrational fear responses with more adaptive interpretations [2,4,5]. Despite the proven effectiveness of exposure therapy [6], fewer than 8% of individuals with specific phobias seek treatment [7]. Barriers include the unpredictability of exposure, challenges in implementing therapy, and a shortage of mental health care providers [6,8,9].

An alternative approach, known as in virtuo treatment, uses virtual reality (VR) to address many of these barriers. VR enables users to experience computer-generated environments that can evoke a strong sense of presence or "sense of being there," which is commonly attributed to the replacement of real-world sensory input with computer-generated stimuli [10]. Through this effect, the user cognitively knows that the world around them is unreal, yet experiences a perceptual illusion that makes the brain and the body automatically react to sensory cues [11]. This characteristic of VR allows patients to repeatedly confront realistic phobic stimuli in a manner comparable with traditional in vivo exposure conducted in real-life settings. But in contrast to real-world exposure, virtual reality exposure therapy (VRET) allows therapists to precisely control and gradually adjust exposure intensity according to individual patient needs, offering high repeatability and flexibility while maintaining a sense of safety for the patient [12]. This controlled environment may reduce barriers to treatment and encourage engagement with exposure exercises that might otherwise be avoided in real-life situations. Several meta-analyses have reported greater patient willingness to engage in VRET compared with in vivo exposure [13-15], which has frequently been attributed to the perceived safety and controllability of the virtual environment. The findings of a more recent meta-analysis on VRET for specific phobias, which reported no significant differences between VRET and traditional in vivo exposure at posttreatment or follow-up [16], together with the inclusion of VRET in the American Psychological Association guidelines for empirically supported treatments, demonstrate the effectiveness of VRET for specific phobias [17].

Furthermore, VRET can expand access to treatment for individuals who face barriers to traditional methods, such as limited facilities, practitioner shortages, cultural disparities, or life circumstances [18]. One possible approach is the use of automated interventions, also referred to as guided self-help [19], which have recently been shown to achieve substantial symptom reductions [20]. However, findings from internet-based mental health interventions have suggested that maintaining treatment adherence can be challenging, particularly in unguided approaches [21]. This raises the question of how therapeutic guidance, psychoeducation, and emotional support can be delivered in the absence of a physically present therapist, for example, through virtual therapists that serve as a functional analog to human therapists during exposure exercises [22,23].

To date, several studies have explored automated VRET interventions incorporating virtual therapists, coaches, or health agents to provide guidance, psychoeducation, and user support throughout the treatment process [19,20,24,25]. Although these studies primarily focused on treatment efficacy and the implementation of exposure procedures, later work increasingly evaluated users' experiences and perceptions of virtual therapists and automated VRET systems [26-28]. A recent systematic review highlighted the increasing use of virtual humans in VR mental health research and the diversity of their implementations [29].

The implementation of these virtual therapists varied substantially across studies. Donker et al [24] incorporated a 2D animated avatar named *Tara* to deliver background information on phobias via a smartphone app. Miloff et al [25] incorporated voice-over psychoeducation, treatment instruction, and guidance through the exposure tasks in a VRET application for spider phobia primarily. Freeman et al [19], Lindner et al [20], and Wei et al [28] incorporated embodied conversational agents (ECAs) into automated VRET applications to provide psychoeducation, therapeutic guidance, and user support. While Lindner et al [20] used an animated holographic virtual therapist primarily for introductory psychoeducation, Wei et al [28] implemented an embodied virtual coach that conducted interactive psychoeducational consultations. Freeman et al [19] further extended this approach by using an embodied virtual coach not only for psychoeducation and assessment but also to guide participants throughout exposure exercises. ECAs are virtual agents designed to engage in interactive dialogues with users [30] and possess a virtual embodiment, enabling them to simulate real-world social interactions. Such interactions may enhance treatment adherence and efficacy through, that is, further immersion into the therapeutic context and a more natural way of conveying psychoeducation and reinforcing progress [31]. Consequently, the implementation of virtual therapists and ECAs has already been explored in various anxiety- and phobia-related applications beyond clinical trials of automated VRET interventions [32].

Given that the therapist-patient relationship can influence therapy outcomes [33] and human-agent interactions can evoke social and emotional responses similar to human-human interactions [34,35], examining how various aspects of ECA design impact users in clinical settings is crucial. Prior research suggests that the representation and embodiment of virtual interaction partners can influence users' sense of social presence and interpersonal engagement within immersive environments [36,37]. Social presence, defined as the sense of "being with another" [38], has been associated with factors such as empathy, involvement, and affective responses during social interactions [39]. Furthermore, trust and perceived listening behavior have been identified as important characteristics of virtual human interactions that can influence users' willingness to disclose personal information [40]. In therapeutic settings, these interpersonal and social interaction factors may further contribute to the development of therapeutic alliance between users and virtual therapists [41]. Factors such as visual appearance, social cues, and interactivity may therefore shape how users perceive and interact with virtual therapists during VRET interventions.

XSL•FO
RenderX

A scoping review by Provoost et al [42] further highlighted the importance and complexity of determining the optimal appearance of ECAs in a clinical context to enhance treatment success. As 1 key aspect of ECAs, appearance can be broken down into subfeatures such as degree of realism, species, gender, or clothing [43]. Previous research has examined how visual realism influences users' perceptions of virtual humans in social and conversational contexts. However, findings remain mixed, with studies suggesting that increased realism may affect aspects such as social presence, trust, and engagement, while others report little or no effect on these outcomes [36,44]. Beyond realism, the embodiment form of an artificial agent may influence how users perceive and interact with it. Previous research suggests that appearance establishes social expectations and can bias interactions, with human-, animal-, and machine-like embodiments potentially eliciting different responses [45]. Nevertheless, a recent systematic review revealed that extended reality–based ECA research has overwhelmingly focused on human-like embodiments, with only a small number of studies investigating robotic agents and virtually none exploring animal-like embodiments [46]. Consequently, little is known about how different embodiment species influence users' perceptions and experiences, and even less is known about their role in therapeutic extended reality applications such as VRET. Therefore, this study focuses on realism and species as 2 appearance dimensions that may shape users' perceptions of virtual therapists in VRET.

Degree of Realism of an ECA

A literature review by ter Stal et al [47] on the effects of ECAs design features in eHealth found mixed results regarding the impact of realism on user perception. While stylized ECAs have been rated more positively on friendliness, several studies indicated a preference for more realistic ECAs. However, findings regarding the optimal rendering style for ECAs varied, with significant differences in study design and participant demographics. The authors emphasized the need for further research to understand better how different rendering styles influence user experience.

McDonnell et al [48] explored how different rendering styles influence how people perceive animated virtual humans. They developed 10 virtual characters, ranging from abstract to realistic styles, and found that motion anomalies were perceived as less unpleasant in stylized virtual humans than in realistic virtual humans. This finding aligns with the Uncanny Valley effect described by Mori et al [49], which suggests that as a robot becomes more human-like, affinity toward it increases until it reaches a threshold where slight imperfections cause unease. Beyond this point, as realism continues to improve, affinity rises again. Research over the past decades has shown that Mori's observations extend to virtual characters as well [50]. Consequently, the Uncanny Valley effect must be carefully considered when designing ECAs for virtual environments to ensure positive user engagement and acceptance.

Species of an ECA

ECAs can take various forms beyond human representations, including animalistic, robotic, or fantastical creatures. A review by Mitrut et al [32] showed that virtual therapists and similar ECAs have been embodied in a wide range of forms, including male and female human avatars, photographic representations of clinicians, fantasy characters, and virtual animal characters. This diversity suggests that no consensus has yet emerged regarding the most suitable embodiment for virtual therapists. Animalistic ECAs have been studied in different contexts as digital conversation partners. For example, Chi et al [51] tested a tablet-based ECA system as a digital pet in the homes of 10 older adults. Although participants desired more extensive conversational capabilities, most developed a sense of companionship with the digital pet and appreciated its social support, health information, instant assistance, and entertainment features. A quantitative follow-up study by Bott et al [52] using the same pet avatar with older adults in a hospital found that the pet avatar helped alleviate symptoms of delirium and loneliness.

Although research on the Uncanny Valley effect in nonhuman virtual characters is limited [53], studies on the impact of realism in virtual animals suggest that an Uncanny Valley effect also exists for animalistic characters [53,54]. Notably, Sierra Rativa et al [54] investigated virtual pets with varying degrees of realism, examining dimensions such as familiarity, naturalness, attractiveness, and animateness. Their results showed that movement amplified the Uncanny Valley effect, while morbid features led to more negative perceptions of virtual pandas. Furthermore, differences between expert evaluations and participant responses regarding animal likeness underscored the nuanced ways virtual animals are perceived across various dimensions.

This Study

Previous work on VRET applications, including virtual therapists for psychoeducation or guidance, has rarely examined how appearance-related design dimensions influence users' socioemotional responses to these agents. While voice, behavior, and interactivity are also important design dimensions, realism and species were selected because they directly shape the visual embodiment of virtual therapists. Although realism has received considerable attention in prior research, findings remain inconclusive, whereas embodiment species have received comparatively little attention. At the same time, most prior work has examined these design dimensions in general interaction settings rather than therapeutic exposure scenarios, in which virtual therapists act as social interaction partners, providing psychoeducation and guidance prior to exposure. To address this gap, we selected dog phobia as an example application, as specific phobias are among the most established applications of VRET and provide a suitable context for investigating how users perceive virtual therapists in therapeutic exposure scenarios. Furthermore, dog phobia is particularly suitable for VRET, as encounters with dogs can be simulated in immersive virtual environments while retaining key characteristics of real-world exposure situations. By implementing interactive psychoeducational dialogues prior to exposure, we examined how ECA embodiments varying in realism and species influenced users' experiences and perceptions of a virtual therapist in a simulated VRET intervention for specific phobias. Our findings aim to assist interdisciplinary teams in making

XSL•FO
RenderX

informed decisions when integrating ECAs into research and clinical applications.

Methods

Experimental Design

We developed a VRET application for dog phobia and conducted a quantitative study in which each participant interacted with 4 distinct virtual therapist characters across 4 different interventions. The interventions varied in their psychoeducational content and exposure scenarios to maintain participant engagement. As the psychoeducational content was designed to build upon content presented in previous interventions, the intervention scenarios followed a fixed order for all participants, while virtual character assignment to sessions was randomized. Each intervention consisted of 2 main components: a virtual therapist providing psychological education and exposure to a virtual dog in the controlled VR environment. We measured several dependent variables related to the virtual characters, including user experience, therapeutic relationship skills, and trust as perceived by the participants. Additionally, we defined 4 different intervention scenarios as independent variables to analyze their effects on user experience, cybersickness, and presence. Our goal was to assess the potential of the application for follow-up studies with patients and evaluate the comparability of the virtual therapists across different intervention scenarios.

To ensure that participants could meaningfully assess the virtual therapist within the given context, we designed interactive dialogues between the participant and the virtual therapist. A Wizard-of-Oz approach was used, in which a human operator simulated the virtual therapist's responses [55]. This approach was chosen to ensure consistent therapeutic content and interaction quality across all virtual therapist embodiments while maintaining a natural conversational experience. The dialogue structure was predefined, with participants verbally responding to questions posed by the virtual therapist. Depending on the participant's response, the operator triggered corresponding scripted responses in real time. By relying on predefined dialogue paths, variability in the delivered content was minimized and experimental control was maintained across conditions. The same researcher operated the system throughout all study sessions to further ensure consistency in dialogue delivery.

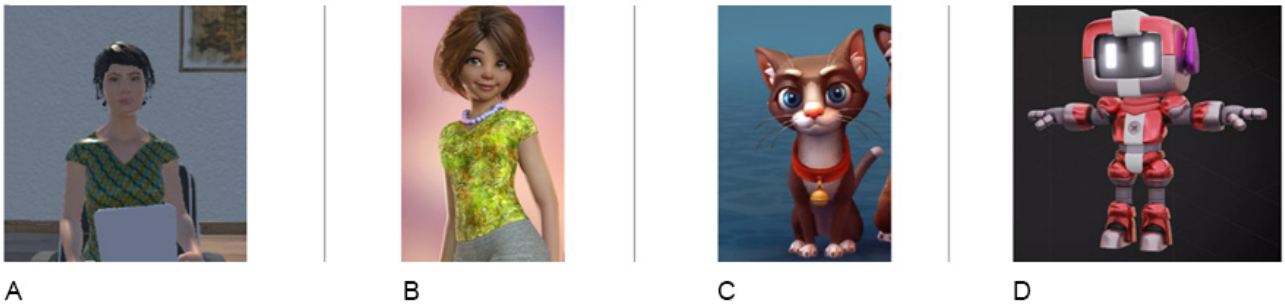
Recruitment

This study is an initial exploratory evaluation of different virtual character types in a VRET application, aiming to generate insights for further interventions with patients with dog phobia. Subsequently, 40 participants from a convenience sample of healthy adult volunteers were recruited via the university's student portal and on-campus notices. No specific inclusion criteria were required for participation, and each participant received a €15 (€1=US $1.12 as of July 17, 2023) compensation. The study took place in our Immersive Reality laboratory over a period of 2 weeks in July 2023.

Virtual Character Types

The 4 virtual character types chosen for the study were a realistic adult woman, a stylized adult woman, a stylized cat, and a stylized robot (Figure 1). The realistic and stylized human characters represented 2 levels of visual realism while maintaining the same species, gender, and clothing to facilitate comparison. The stylized cat was included to investigate an animal-like embodiment, whereas the stylized robot represented a machine-like embodiment. Together, these character types were selected to investigate how variations in realism and species influence users' perceptions of virtual therapists within a simulated VRET intervention.

Figure 1. Virtual character types used in the study (from left to right): (A) realistic adult woman, (B) stylized adult woman, (C) stylized cat, and (D) stylized robot.



To enhance realism, we implemented lip synchronization and eye movement for the first 3 characters using the Unity SALSA Lip Sync Suite package (Crazy Minnow Studio). The 3D models for the realistic and stylized humans were sourced from Daz3D (Daz 3D), while the cat and robot models were obtained from the Unity Asset Store.

Stimulus Material

Participants used a Meta Quest 2 VR headset during the experiment. They had no avatar representation, remained seated throughout the psychological education phase, and stood up for the exposure phase. The application was developed using Unity 3D (Unity Technologies), Daz Studio (Daz 3D), Visual Studio (Microsoft), and Audacity (The Audacity Team). We created 2 separate environments: 1 for the psychological education and 1 for the exposure.

A specialist wrote a separate dialogue for each intervention based on the psychological education principles related to specific phobias and exposure therapy. We prerecorded the virtual therapists' dialogues using a female voice, ensuring consistency with the 2 humanoid virtual therapists. The dialogues included interactive questions, and the experimenter could trigger appropriate responses based on participant input.

XSL•FO
RenderX

Participants went through these interactive dialogues at their own pace, which typically took around 3.5 minutes.

During the exposure phase, participants faced a shepherd dog standing, sitting, or lying on a sidewalk in an urban neighborhood (Figure 2). We set up 4 different exposure scenarios with varying walking directions and typical nonaggressive behaviors by the dog to keep participants engaged during the experiment. The implemented behaviors were intended to reflect typical everyday encounters with dogs while remaining comparable across the different exposure scenarios. After 55 seconds, the dog entered an idle state, meaning no further active behaviors were performed. Participants were then asked to remove the headset, resulting in approximately 1 minute of exposure per scenario. To ensure comparability between the exposure scenarios, the dog was positioned at a similar distance from the participant in all intervention scenarios. Because the content of the exposure scenarios was similar, they were conducted in the same order for all participants, with the therapist randomized in each scenario.

Figure 2. Exposure scene.

The experimenter controlled the intervention via a web app. The web app displayed the participant's camera view, allowing the experimenter to trigger conversation content by clicking interface buttons. A local web server connected the web app with the Unity application running in the Unity Editor.

Measures

An overview of the independent and dependent variables and the questionnaires used can be found in Table 1.

XSL•FO
RenderX

Table . Questionnaires used in the study.

Questionnaire	Items	Dependent variables	Independent variables
Sociodemographic questionnaire	7	Age, gender, occupation, VR experience, presence of anxiety disorder or phobia, and presence of uncorrected sensory impairment	N/A[a]
Socioemotional questionnaire	48	Therapeutic relationship skills, emotions in therapeutic work	Virtual character type
Virtual therapist–patient relationship	2	Trust, willingness to share information	Virtual character type
User Experience Questionnaire Short Version	8	Pragmatic quality, hedonic quality, and overall quality	Virtual character type, intervention scenario
Affinity for technological interaction scale	9	Affinity for technological interaction	N/A[a]
Igroup presence questionnaire	1	Presence	Virtual character type, intervention scenario
Short cybersickness questionnaire	1	Cybersickness	Intervention scenario
Postquestionnaire	4	Preference for virtual character type, empathy for dog phobic	Virtual character type

[a]N/A: not applicable.

Therapeutic Relationship Skills and Emotions in Therapy Work

The socioemotional questionnaire used in this study was developed in a theory-informed manner based on the systematic review by Heinonen and Nissen-Lie from 2019 [56], which summarizes professional and interpersonal therapist characteristics associated with therapeutic effectiveness. The selected items were adapted for use in virtual therapist interactions during VR exposure therapy to capture participants' perceptions of different virtual character types. As no standardized instrument specifically addressing these interaction characteristics in immersive VR therapeutic settings was available, a custom exploratory questionnaire was considered appropriate for this study's aims.

The questionnaire consists of 3 sections. The first 2 sections assess the therapist's relational skills, while the third section evaluates emotions in therapeutic work. The questionnaire was inspired by the "Questionnaire of Supervision, Intervision, and Self-awareness" from the Medical School Hamburg and the State Chamber of Psychotherapists Baden-Württemberg and the "Development of Psychotherapists Common Core Questionnaire" by Orlinsky et al [57]. Responses were recorded on a scale from 0 (do not agree at all) to 3 (completely agree), with an option for "kA" (no response).

Virtual Therapist–Patient Relationship

To adapt the questionnaire to a virtual therapy context, 2 additional questions were included: first regarding the willingness to share information with the virtual therapist, adapted from DeVault et al [58], and the second on the trust in the virtual therapist, based on the importance of trust in user interactions [59]. Both items were rated on a 5-point scale (1=not at all, 5=absolutely).

User Experience

As each intervention consisted of a psychoeducational dialogue with the virtual therapist followed by an exposure scenario, the User Experience Questionnaire Short Version (UEQ-S) [60] was administered after each intervention to assess the overall user experience. Because virtual therapist assignment was randomized across intervention scenarios, UEQ-S ratings were later analyzed with respect to both the virtual therapist embodiment and the intervention scenario. The UEQ-S uses a 7-point scale, with positive terms alternating on the left and the right.

Affinity for Technological Interaction

The participants' technical affinity was assessed using the Affinity for Technology Interaction (ATI) scale [61]. This scale measures an individual's propensity to engage with technological products, a factor that could significantly affect user experience. The questionnaire consists of 9 items on a 7-point Likert scale, ranging from 1 (completely disagree) to 6 (completely agree).

Presence

Effective virtual exposure therapy requires a natural and realistic environment for immersion [62]. To assess the sense of presence in the virtual environment, we used the "Igroup Presence Questionnaire" [63], specifically focusing on spatial presence. The scale ranges from 1 (no impression of being there) to 7 (strong impression of being there).

Cybersickness

Given the potential for cybersickness in VR applications [64], participants were asked about dizziness or nausea experienced during the intervention scenario. The scale ranges from 1 (no dizziness or nausea) to 5 (severe dizziness or nausea).

XSL•FO
RenderX

Postquestionnaire

After completing the experiment, participants indicated their preferred virtual therapist and provided justifications for their choice. These open-ended responses were collected to provide additional contextual information regarding participants' preferences and were not subjected to formal qualitative analysis. They also assessed their ability to empathize with a dog-phobic individual during the simulation on a 7-point Likert scale (1=not at all, 7=very well). Participants could also provide qualitative feedback on their overall experience in an open-ended question.

Procedure

The experiments were conducted in a university laboratory, ensuring a consistent and controlled environment throughout the study. At the start of each experiment, participants were given information about the study and asked to complete an informed consent form. They then filled out the sociodemographic questionnaire and the ATI scale. Afterward, participants were briefed on dog phobia and VRET to familiarize themselves with the experimental setting. Any questions from participants were addressed at this stage, and a brief text was provided to help immerse them in the role of a dog-phobic individual and encourage engagement during the interactions with the virtual therapists.

While participants read the provided information, the experimenter established the experimental conditions for the first intervention scenario. Participants were then instructed to sit in a chair and wear the VR headset. Once the experimenter triggered the "Start" button in the web app, the virtual therapist for that intervention appeared, and the dialogue between the virtual therapist and the participant began. When the dialogue ended, participants were instructed to stand up as the scene transitioned to the exposure environment.

After the exposure ended and participants removed the VR headset, they were given the socioemotional questionnaire, the virtual therapist–patient relationship questionnaire, the UEQ-S, the spatial presence item of the Igroup Presence Questionnaire, and the short cybersickness questionnaire. Once they had completed the questionnaires, the next intervention scenario began with another virtual character type embodying the therapist. This process was repeated for all 4 intervention scenarios. After the fourth intervention scenario, participants completed the postquestionnaire.

Ethical Considerations

The planning, conduct, and reporting of the study were in accordance with the World Medical Association Declaration of Helsinki. All procedures were performed in compliance with relevant laws and institutional guidelines and were approved by the local ethics committee on July 13, 2023 (reference number EL20230701). The privacy rights of human subjects have been observed and informed consent was obtained from all participants before the start of the experiment.

Statistical Analysis

Data were analyzed using SPSS (version 29; IBM Corp). Descriptive statistics are reported as means and standard deviations, consistent with the UEQ-S's recommended reporting and with common practice in prior VR and human-computer interaction research using comparable questionnaire scales. Because the study used a repeated-measures design, analysis of significant differences was conducted using Friedman ANOVA. When a Friedman test indicated a significant difference, post hoc tests were conducted using the Wilcoxon signed-rank test. All statistical tests were 2-sided, and a significance level of α=.05 was applied. Effect sizes were reported as Kendall coefficient of concordance (Kendall *W*) for Friedman tests and matched-pairs rank-biserial correlations (rrb) for Wilcoxon signed-rank tests. The Benjamini-Hochberg correction was applied to adjust the *P* values obtained from the Wilcoxon signed-rank tests to account for multiple comparisons. This adjustment controlled for type I errors, reducing the likelihood of false-positive results. UEQ-S and presence scores were analyzed across intervention scenarios and virtual character types. Cybersickness scores were analyzed across intervention scenarios only. Scores of the socioemotional questionnaire and the virtual therapist–patient relationship questionnaire were analyzed across virtual character types. To evaluate the UEQ-S in comparison with the user experience of other systems, we used the official UEQ data analysis tool [60]. Open-ended responses were considered descriptively and used to contextualize the quantitative findings.

Results

Participant Characteristics

The study was conducted with 40 participants over 2 weeks. Participants' ages ranged from 19 to 37 (mean 24.9, SD 3.86) years, with 25 (62.5%) participants identifying as male, 13 (32.5%) as female, and 2 (5%) as nonbinary. Out of 40 participants, 31 (77.5%) participants reported being students, 5 (12.5%) reported being research assistants, and 4 (10%) reported not working in a research-related field. Out of 40 participants, 34 (85%) participants had previous experience with VR, 15 (37.5%) participants reported having an anxiety disorder or phobia, 24 (60%) participants reported having no anxiety disorder or phobia, and 1 (2.5%) person declined to provide this information. The average ATI was mean 4.58 (SD 0.58).

User Experience, Presence, and Cybersickness Across Intervention Scenarios

Table 2 presents the means, SD, and Friedman ANOVA results for user experience, presence, and cybersickness across the intervention scenarios. Analysis using Friedman ANOVA revealed no statistically significant differences in pragmatic, hedonic, or overall quality (UEQ-S scores) across the intervention scenarios. Similarly, no significant differences were found in the presence or cybersickness scores across the scenarios.

XSL•FO
RenderX

Table . Descriptive statistics and Friedman ANOVA results of user experience, cybersickness, and presence scores across intervention scenarios.

Dependent value	Scenario 1, mean (SD)	Scenario 2, mean (SD)	Scenario 3, mean (SD)	Scenario 4, mean (SD)	Chi-square (*df*)	*P* value	Kendall *W*
User experience							
Pragmatic quality	1.93 (0.73)	1.72 (1.02)	1.74 (0.995)	1.74 (1.15)	1.19 (3)	.76	0.010
Hedonic quality	1.5 (1.095)	1.56 (1.220	1.56 (1.17)	1.67 (1.18)	3.93 (3)	.27	0.033
Overall quality	1.72 (0.84)	1.64 (0.99)	1.65 (0.96)	1.71 (1.06)	0.67 (3)	.87	0.006
Presence	5.23 (1.1)	5.3 (1.22)	5.1 (1.41)	5.3 (1.44)	2.92 (3)	.40	0.024
Cybersickness	1.38 (0.67)	1.33 (0.7)	1.4 (0.84)	1.45 (0.81)	0.76 (3)	.86	0.006

User Experience Across Virtual Character Types

Table 3 provides the means, SD, and Friedman ANOVA results for user experience, presence, and cybersickness across the intervention scenarios. The means and SDs of the UEQ-S are also presented in Figure 3 for a better visual overview. As illustrated in Figure 3, differences between the virtual character types were primarily observed in hedonic quality, whereas pragmatic and overall quality ratings were relatively similar across characters. In particular, the realistic human received the lowest hedonic quality ratings, while the cat received the highest. As shown in Table 3, Friedman ANOVA did not reveal statistically significant differences in the pragmatic or overall quality across the virtual character types. However, a statistically significant difference was observed in the hedonic quality. Subsequent post hoc Wilcoxon signed-rank tests indicated that the hedonic quality of the realistic human was significantly lower than that of the other virtual character types. Specifically, comparisons showed inferior hedonic quality for the realistic human compared with the stylized human (z=2.53; P=.01; Padj=.02; rrb=−0.54), the robot (z=2.726; P=.006; Padj=.012; rrb=−0.56), and the cat (z=3.772; P<.001; Padj=.004; rrb=−0.81). Additionally, a significant difference in hedonic quality was observed between the stylized human and the cat, with the cat demonstrating higher hedonic quality (z=2.528; P=.02; Padj=.02; rrb=−0.52).

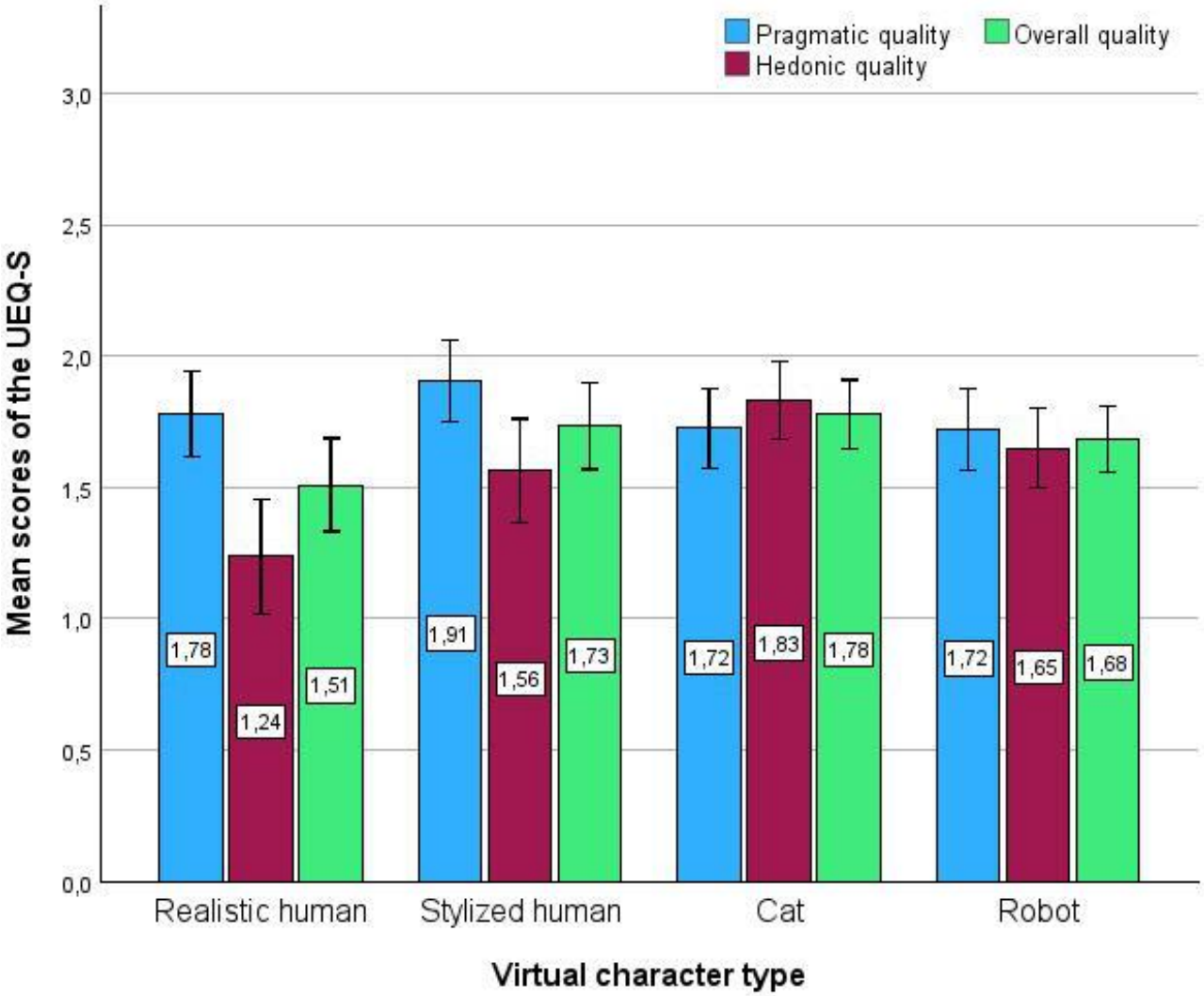
XSL•FO
RenderX

Table . Descriptive statistics and Friedman ANOVA results of user experience and socioemotional questionnaire scores across intervention scenarios.

Dependent value	Realistic human, mean (SD)	Stylized human, mean (SD)	Cat, mean (SD)	Robot, mean (SD)	Chi-square (*df*)	*P* value	Kendall *W*
User experience							
Pragmatic quality	1.78 (1.03)	1.91 (0.97)	1.73 (0.93)	1.72 (0.98)	0.51 (3)	.92	0.004
Hedonic quality	1.24 (1.38)	1.56 (1.26)	1.83 (0.94)	1.65 (0.96)	13.23 (3)	<.004	0.110
Overall quality	1.51 (1.13)	1.73 (1.04)	1.78 (0.82)	1.68 (0.8)	3.71 (3)	.30	0.031
Socioemotional questionnaire							
Therapeutic relationship skills 1	34.48 (8.515)	37.63 (8.227)	35.38 (9.502)	34.53 (9.454)	4.157 (3)	.25	0.035
Therapeutic relationship skills 2	28.28 (6.548)	30.13 (6.458)	28.68 (7.388)	27.13 (5.967)	10.108 (3)	.02	0.084
Emotions in therapeutic work	11.13 (3.924)	10.85 (4.105)	10.78 (4.264)	9.8 (3.428)	5.959 (3)	.11	0.050
Virtual therapist–patient relationship							
Trust	3.55 (1.06)	3.78 (0.97)	3.72 (1.09)	3.48 (1.13)	2.27 (3)	.52	0.019
Willingness to share information	3.33 (1.07)	3.7 (1.07)	3.6 (1.13)	3.45 (1.2)	3.37 (3)	.34	0.028
Presence	5.23 (1.21)	5.18 (1.3)	5.3 (1.34)	5.22 (1.35)	0.28 (3)	.96	0.002

XSL•FO
RenderX

Figure 3. Mean scores of the UEQ-S across the therapists (blue: pragmatic quality, red: hedonic quality, green: overall quality; error bars show standard errors). UEQ-S: User Experience Questionnaire Short Version.



Socioemotional Questionnaire and Virtual Therapist–Patient Questionnaire

While no significant differences were found among the virtual character types regarding the first and third sections of the socioemotional questionnaire (Table 3), a significant difference was observed in the second section assessing therapeutic relationship skills. Post hoc analysis revealed that the therapeutic relationship skills of the robot were significantly inferior to that of the stylized human ($z=-2.315$; $P=.02$; $P_{adj}=.02$; rbb=−0.42).

No significant differences were observed among the virtual character types regarding trust or willingness to share information. Additionally, presence scores did not vary significantly across the different virtual character types.

Postquestionnaire

Participant preferences for a virtual therapist in the VRET intervention showed distinct variations. The robot was the most preferred choice (12/40, 30% of participants), followed by the stylized human (11/40, 27.5%), the cat (10/40, 25%), and finally, the realistic human (7/40, 17.5%), which was the least favored. Participants also provided diverse insights when asked about their reasons for favoring certain characters. Those who chose the realistic human mentioned its perceived "professionalism," while those who preferred the stylized human highlighted a sense of "comfort." However, some participants found the realistic human "uncanny." The cat was described as having "playfully" loosened up the intervention and as appearing particularly "trusting." The robot, valued for its "objectivity" and "anonymity," appealed to some participants seeking a "trustworthy" therapeutic presence. Moreover, some participants also commented that its abstract nature facilitated immersion in the virtual environment and minimized distractions caused by a human character's unnatural appearance or behavior. Regarding empathetic engagement with dog phobia, most participants (31/40, 77.5%) expressed confidence, 10% (4/40) fell in the middle, and 12.5% (5/40) expressed lower confidence.

Discussion

Principal Findings

This study explored how different virtual therapist embodiments were perceived within a simulated VRET scenario by a convenience sample of healthy adults. We examined the effects of different realism levels and species of virtual therapists on user experience and perceived socioemotional relationship skills.

XSL•FO
RenderX

To ensure comparability among the virtual character therapist types, regardless of differing intervention scenarios, we also analyzed user experience, presence, and cybersickness across the intervention scenarios.

Our VRET application received above-average user experience scores, with no significant differences in user experience, presence, or cybersickness across the 4 intervention scenarios, which varied slightly in educational content and exposure settings. These findings allowed for the independent evaluation of the virtual character types without interference from the intervention scenarios.

Significant differences emerged in the hedonic quality of the virtual character types. The realistic human received significantly lower hedonic quality ratings than the stylized human, robot, and cat. In contrast, the cat received the highest hedonic quality ratings and scored significantly higher than both human embodiments. Contrary to our expectations, no significant differences were observed regarding trust, willingness to share information, or most socioemotional dimensions. However, the robot received lower ratings than the stylized human on 1 dimension of therapeutic relationship skills. Finally, participants' preferences were distributed across the stylized human, robot, and cat embodiments, with the realistic human rated as the least preferred and the robot as the most preferred.

Interpretation Within the Context of Prior Work

The most consistent finding of this study was the comparatively unfavorable evaluation of the realistic human embodiment. The realistic human received significantly lower hedonic quality ratings than all other virtual therapist embodiments and was also the least preferred character. Qualitative feedback further revealed that several participants perceived the realistic human as "uncanny." One possible explanation for these findings is the Uncanny Valley effect [49], according to which highly human-like but imperfect artificial agents can evoke discomfort or unease. This interpretation is consistent with previous research on virtual humans and social robotics. The perception of the realistic human as more uncanny than the stylized human supports McDonnell et al [48], who found that motion anomalies in anthropomorphic virtual humans were less disturbing than in realistic virtual humans. Similarly, Volonte et al [65] reported that subtle behavioral cues were more noticeable in realistic virtual humans than in stylized ones, which may have contributed to the lower preference for the realistic human in our study. Together, these findings suggest that increasing visual realism alone does not necessarily improve users' perceptions of virtual therapists and may even negatively affect user experience when realism is not sufficiently achieved.

Interestingly, participants responded positively to both nonhuman embodiments. The cat received the highest hedonic quality ratings, whereas the robot was the most frequently preferred virtual therapist. Qualitative responses indicated that participants often perceived the cat as playful and trustworthy, while the robot was described as objective, anonymous, and less distracting than the human embodiments. These findings are consistent with previous research suggesting that nonhuman agents can successfully fulfill social roles without adopting a human-like experience [43]. Research on animal-like agents has demonstrated that users can develop feelings of companionship and social connection with virtual pets [52,53], while research in social robotics [66,67] has shown that human-robot interactions exhibit social similarities to human-human interactions. The positive evaluations of both the cat and the robot indicate that virtual therapists do not necessarily need to resemble humans to be perceived favorably by users.

Contrary to our expectations, no significant differences were observed regarding trust or willingness to share information across the different virtual therapist embodiments. Previous research on automated VRET interventions has suggested that virtual therapists can contribute to therapeutic processes by providing psychoeducation, guidance, and supporting a relationship similar to a therapeutic alliance [27,31]. The present findings suggest that different embodiment species, including nonhuman embodiments, may be capable of supporting comparable levels of trust and willingness to share information. This interpretation is consistent with previous reviews showing that virtual therapists and ECAs have been embodied in a wide range of forms, including human avatars, photographic representations of clinicians, holographic therapists, fantasy characters, and virtual animals [32]. The positive evaluations of multiple embodiments suggest that there may not be a single universally preferred virtual therapist design. Future research should therefore investigate whether preferences for virtual therapist embodiments differ across user groups and application contexts. However, these findings should be interpreted with caution, as the relatively small convenience sample and exploratory nature of the study may have limited the ability to detect more subtle embodiment-related effects.

A similar pattern emerged in the socioemotional evaluations. Although the robot was the most frequently preferred embodiment, it received significantly lower ratings than the stylized human on 1 dimension of therapeutic relationship skills. This suggests that participants did not necessarily associate their preferred virtual therapist with the embodiment they perceived as most suitable for a therapeutic interaction. While the robot was often described as objective and anonymous, participants frequently characterized the stylized human as comforting and approachable. These findings indicate that different virtual therapist embodiments may be associated with different perceived strengths and that preference alone may not fully capture how users evaluate a virtual therapist. However, these findings should be interpreted with caution, as participants reported a relatively high affinity for technology, which may have increased the acceptance of technological embodiments such as the robot.

To our knowledge, this study is among the first to examine how realism and species influence users' perceptions of virtual therapists within a simulated VRET setting. Previous research has demonstrated the feasibility and effectiveness of incorporating virtual therapists, coaches, or health agents into automated VRET interventions to provide psychoeducation, guidance, and user support throughout treatment [19,20,24,25]. However, little attention has been given to how different embodiments are perceived by users. The present findings therefore provide an initial indication that embodiment

XSL•FO
RenderX

characteristics may influence user experience and socioemotional evaluations, while trust and willingness to share information appear less dependent on embodiment.

Limitations

While this study provides valuable insights, several limitations must be acknowledged. First, the sample consisted of healthy young adults who were not diagnosed with dog phobia, which limits the generalizability of the findings to clinical populations.

Second, the psychoeducational dialogues with the virtual therapists lasted only approximately 3.5 minutes, which is sufficient to obtain a first impression of different embodiments but unlikely to reflect the development of therapeutic processes that typically emerge over longer interactions. As a result, constructs such as trust, willingness to share information, and perceived therapeutic relationship skills may not have been fully captured within the brief simulated encounters.

Further, since the socioemotional questionnaire was developed specifically for this study and has not undergone formal psychometric validation, its results should be interpreted with caution. Although the questionnaire's items are based on theoretical findings and derive from constructs identified in the existing literature [56], the exploratory nature of the instrument limits the conclusions that can be drawn regarding construct validity and reliability. Future research should include formal psychometric validation procedures, such as assessments of internal consistency and construct validity, as well as comparisons with established socioemotional and therapeutic alliance measures.

Furthermore, the study relied on qualitative feedback to assess the Uncanny Valley effect, which limits the precision of these findings. Finally, while the Wizard-of-Oz approach increased experimental control and consistency across conditions, it may have reduced the ecological validity of therapist-participant interactions compared with fully autonomous conversational systems.

Conclusions

Despite the potential of VRET for treating specific phobias, it is not yet widely integrated into clinical practice. This underscores the need for further research into its key components, such as the role of virtual therapists, to improve its implementation and accessibility for individuals with specific phobias. This study represents an initial exploration of how the realism and species of virtual therapist character types may influence user experience and their socioemotional evaluation in a VRET application. The realistic human model received lower hedonic quality ratings and preference scores than the other embodiments, suggesting that increased visual realism does not necessarily result in more positive perceptions of virtual therapists and may, in some cases, lead to less favorable user responses. Additionally, the cat model's higher hedonic quality and the robot's popularity suggest that different user groups may respond differently to virtual therapist embodiments. Taken together, the findings indicate that virtual therapist design may influence user experience and socioemotional perceptions, highlighting the importance of considering embodiment characteristics when developing automated VRET interventions. Further research is needed to better understand the impact of virtual therapist design in VRET applications. Future studies should compare realistic and stylized virtual characters, as well as different species, across diverse user groups, longer interaction periods, and clinical settings.

Acknowledgments

During the preparation of this work, the authors used ChatGPT and Grammarly, in order to improve the readability and language of the manuscript. After using these services, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

The work was supported by funding from the Hamm-Lippstadt University of Applied Sciences.

Data Availability

Data from this study are available from the corresponding author upon reasonable request.

Authors' Contributions

MCH contributed to the conceptualization, methodology, investigation, formal analysis, visualization, writing—original draft, and writing—review & editing. NLP contributed to the conceptualization, methodology, and writing—review & editing. FV and JNV-A contributed to the conceptualization, methodology, writing—review & editing, and supervision. YS contributed to the conceptualization, methodology, and supervision.

Conflicts of Interest

None declared.

References

1. Wardenaar KJ, Lim CCW, Al-Hamzawi AO, et al. The cross-national epidemiology of specific phobia in the World Mental Health Surveys. Psychol Med 2017 Jul;47(10):1744-1760. [doi: 10.1017/S0033291717000174] [Medline: 28222820]

XSL•FO
RenderX

2. Rachman S. Verhaltenstherapie Bei Phobien [Behavior Therapy for Phobias], 2nd edition: Urban und Schwarzenberg; 1972.
3. Steinmann SA, Wootton BM, Tolin DF. Exposure therapy for anxiety disorders. In: Friedman HS, editor. Encyclopedia of Mental Health, 2nd edition: Academic Press; 2016:186-191. [doi: 10.1016/B978-0-12-397045-9.00266-4]
4. Leibetseder M, editor. Grundlagenbuch Verhaltenstherapie: Diagnostik, Methoden, Anwendungsbereiche, Sprachanalysen [Fundamentals of Behavior Therapy: Diagnostics, Methods, Areas of Application, Language Analyses]: Pabst Science Publisher; 2018.
5. Hamm A. Spezifische phobien [specific phobias]. In: Fortschritte Der Psychotherapie Manuale Für Die Praxis: Hogrefe; 2006, Vol. 27.
6. Schröder D, Wrona KJ, Müller F, Heinemann S, Fischer F, Dockweiler C. Impact of virtual reality applications in the treatment of anxiety disorders: a systematic review and meta-analysis of randomized-controlled trials. J Behav Ther Exp Psychiatry 2023 Dec;81:101893. [doi: 10.1016/j.jbtep.2023.101893] [Medline: 37453405]
7. Mackenzie CS, Reynolds K, Cairney J, Streiner DL, Sareen J. Disorder-specific mental health service use for mood and anxiety disorders: associations with age, sex, and psychiatric comorbidity. Depress Anxiety 2012 Mar;29(3):234-242. [doi: 10.1002/da.20911] [Medline: 22065571]
8. Rizzo AS, Bouchard S, editors. Virtual Reality for Psychological and Neurocognitive Interventions: Springer; 2019.
9. Boeldt D, McMahon E, McFaul M, Greenleaf W. Using virtual reality exposure therapy to enhance treatment of anxiety disorders: identifying areas of clinical adoption and potential obstacles. Front Psychiatry 2019;10:773. [doi: 10.3389/fpsyt.2019.00773] [Medline: 31708821]
10. Slater M, Sanchez-Vives MV. Enhancing our lives with immersive virtual reality. Front Robot AI 2016;3:74. [doi: 10.3389/frobt.2016.00074]
11. Slater M. Immersion and the illusion of presence in virtual reality. Br J Psychol 2018 Aug;109(3):431-433. [doi: 10.1111/bjop.12305] [Medline: 29781508]
12. Emmelkamp PMG. Technological innovations in clinical assessment and psychotherapy. Psychother Psychosom 2005;74(6):336-343. [doi: 10.1159/000087780] [Medline: 16244509]
13. Garcia-Palacios A, Botella C, Hoffman H, Fabregat S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. Cyberpsychol Behav 2007 Oct;10(5):722-724. [doi: 10.1089/cpb.2007.9962] [Medline: 17927544]
14. Garcia-Palacios A, Hoffman HG, See SK, Tsai A, Botella C. Redefining therapeutic success with virtual reality exposure therapy. Cyberpsychol Behav 2001 Jun;4(3):341-348. [doi: 10.1089/109493101300210231] [Medline: 11710258]
15. Powers MB, Emmelkamp PMG. Virtual reality exposure therapy for anxiety disorders: a meta-analysis. J Anxiety Disord 2008;22(3):561-569. [doi: 10.1016/j.janxdis.2007.04.006] [Medline: 17544252]
16. Morina N, Ijntema H, Meyerbröker K, Emmelkamp PMG. Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. Behav Res Ther 2015 Nov;74:18-24. [doi: 10.1016/j.brat.2015.08.010] [Medline: 26355646]
17. Carl E, Stein AT, Levihn-Coon A, et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. J Anxiety Disord 2019 Jan;61:27-36. [doi: 10.1016/j.janxdis.2018.08.003]
18. Kazdin AE, Blase SL. Rebooting psychotherapy research and practice to reduce the burden of mental illness. Perspect Psychol Sci 2011 Jan;6(1):21-37. [doi: 10.1177/1745691610393527] [Medline: 26162113]
19. Freeman D, Haselton P, Freeman J, et al. Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. Lancet Psychiatry 2018 Aug;5(8):625-632. [doi: 10.1016/S2215-0366(18)30226-8] [Medline: 30007519]
20. Lindner P, Miloff A, Bergman C, Andersson G, Hamilton W, Carlbring P. Gamified, automated virtual reality exposure therapy for fear of spiders: a single-subject trial under simulated real-world conditions. Front Psychiatry 2020;11:116. [doi: 10.3389/fpsyt.2020.00116] [Medline: 32210850]
21. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions—a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
22. Fehribach JR, Toffolo MBJ, Cornelisz I, et al. Virtual reality self-help treatment for aviophobia: protocol for a randomized controlled trial. JMIR Res Protoc 2021 Apr 12;10(4):e22008. [doi: 10.2196/22008] [Medline: 33843605]
23. González-Lozoya SM, Meza-Kubo V, Dominguez-Rodriguez A, Ramírez-Fernández C, Morán AL. A self-applied, virtual therapist assisted-based “thera” treatment for small animal phobia in Mexican population: a preliminary evaluation. 2021 Presented at: Technology, Mind, and Society; Nov 3-5, 2021. [doi: 10.1037/tms0000006]
24. Donker T, Cornelisz I, van Klaveren C, et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. JAMA Psychiatry 2019 Jul 1;76(7):682-690. [doi: 10.1001/jamapsychiatry.2019.0219] [Medline: 30892564]
25. Miloff A, Lindner P, Dafgård P, et al. Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: a randomized non-inferiority trial. Behav Res Ther 2019 Jul;118:130-140. [doi: 10.1016/j.brat.2019.04.004] [Medline: 31075675]

XSL•FO
RenderX

26. Lindner P, Rozental A, Jurell A, et al. Experiences of gamified and automated virtual reality exposure therapy for spider phobia: qualitative study. JMIR Serious Games 2020 Apr 29;8(2):e17807. [doi: 10.2196/17807] [Medline: 32347803]
27. Miloff A, Carlbring P, Hamilton W, Andersson G, Reuterskiöld L, Lindner P. Measuring alliance toward embodied virtual therapists in the era of automated treatments with the Virtual Therapist Alliance Scale (VTAS): development and psychometric evaluation. J Med Internet Res 2020 Mar 24;22(3):e16660. [doi: 10.2196/16660] [Medline: 32207690]
28. Wei S, Freeman D, Rovira A. A randomised controlled test of emotional attributes of a virtual coach within a virtual reality (VR) mental health treatment. Sci Rep 2023 Jul 17;13(1):11517. [doi: 10.1038/s41598-023-38499-7] [Medline: 37460586]
29. Wei S, Freeman D, Rovira A. Virtual humans in virtual reality mental health research: systematic review. JMIR XR Spat Comput 2025;2:e75087-e75087. [doi: 10.2196/75087]
30. Cassell J, Bickmore T, Campbell L, Vilhjálmsson H, Yan H. Human conversation as a system framework: designing embodied conversational agents. In: Cassell J, Sullivan J, Prevost S, Churchill EF, editors. Embodied Conversational Agents: MIT Press; 2000.
31. Lindner P. Better, virtually: the past, present, and future of virtual reality cognitive behavior therapy. J Cogn Ther 2021 Mar;14(1):23-46. [doi: 10.1007/s41811-020-00090-7]
32. Mitru O, Moldoveanu A, Petrescu L, Petrescu C, Moldoveanu F. A review of virtual therapists in anxiety and phobias alleviating applications. In: Chen JYC, Fragomeni G, editors. : Springer Presented at: Virtual, Augmented and Mixed Reality: 13th International Conference, VAMR 2021, Held as Part of the 23rd HCI International Conference, HCII 2021; Jul 24-29, 2021; Virtual event p. 72-84. [doi: 10.1007/978-3-030-77599-5_6]
33. Lilliengren P, Falkenström F, Sandell R, Mothander PR, Werbart A. Secure attachment to therapist, alliance, and outcome in psychoanalytic psychotherapy with young adults. J Couns Psychol 2015 Jan;62(1):1-13. [doi: 10.1037/cou0000044] [Medline: 25222907]
34. Pan X, Banakou D, Slater M. Computer based video and virtual environments in the study of the role of emotions in moral behavior. In: D'Mello S, Graesser A, Schuller B, Martin JC, editors. Affective Computing and Intelligent Interaction ACII 2011: Springer; 2011:52-61. [doi: 10.1007/978-3-642-24571-8_6]
35. Robb A, Kopper R, Ambani R, et al. Leveraging virtual humans to effectively prepare learners for stressful interpersonal experiences. IEEE Trans Vis Comput Graph 2013 Apr;19(4):662-670. [doi: 10.1109/TVCG.2013.35] [Medline: 23428451]
36. Oh CS, Bailenson JN, Welch GF. A systematic review of social presence: definition, antecedents, and implications. Front Robot AI 2018;5:114. [doi: 10.3389/frobt.2018.00114] [Medline: 33500993]
37. Wienrich C, Döllinger N, Hein R. Behavioral Framework of Immersive Technologies (BehaveFIT): how and why virtual reality can support behavioral change processes. Front Virtual Real 2021;2:627194. [doi: 10.3389/frvir.2021.627194]
38. Biocca F, Harms C, Burgoon JK. Toward a more robust theory and measure of social presence: review and suggested criteria. Presence Teleoperators Virtual Environ 2003 Oct;12(5):456-480. [doi: 10.1162/105474603322761270]
39. Kort YAW, IJsselsteijn WA, Poels K. Digital games as social presence technology: development of the social presence in gaming questionnaire (SPGQ). 2007 Presented at: Proceedings of the 10th Annual International Workshop on Presence; Oct 25-27, 2007; Barcelona, Spain p. 195-203 URL: https://research.tue.nl/en/publications/digital-games-as-social-presence-technology-development-of-the-so/ [accessed 2026-07-26]
40. You C, Ghosh R, Maxim A, Stuart J, Cooks E, Lok B. How does a virtual human earn your trust? guidelines to improve willingness to self-disclose to intelligent virtual agents. Presented at: IVA '22: Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents; Sep 6-9, 2022; Faro, Portugal p. 1-8. [doi: 10.1145/3514197.3549686]
41. Horvath AO. Research on the alliance: Knowledge in search of a theory. Psychother Res 2018 Jul;28(4):499-516. [doi: 10.1080/10503307.2017.1373204] [Medline: 28899230]
42. Provoost S, Lau HM, Ruwaard J, Riper H. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res 2017 May 9;19(5):e151. [doi: 10.2196/jmir.6553] [Medline: 28487267]
43. Straßmann C, Krämer NC. A categorization of virtual agent appearances and a qualitative study on age-related user preferences. In: Beskow J, Peters C, Castellano G, O'Sullivan C, Leite I, Kopp S, editors. Intell Virtual Agents: Springer; 2017:413-422. [doi: 10.1007/978-3-319-67401-8_51]
44. Kyrlitsias C, Michael-Grigoriou D. Social interaction with agents and avatars in immersive virtual environments: a survey. Front Virtual Real 2022;2:786665. [doi: 10.3389/frvir.2021.786665]
45. Fong T, Nourbakhsh I, Dautenhahn K. A survey of socially interactive robots. Robot Auton Syst 2003 Mar;42(3-4):143-166. [doi: 10.1016/S0921-8890(02)00372-X]
46. Yang FC, Acevedo P, Guo S, Choi M, Mousas C. Embodied conversational agents in extended reality: a systematic review. IEEE Access 2025;13:79805-79824. [doi: 10.1109/ACCESS.2025.3566698] [Medline: 40256415]
47. ter Stal S, Kramer LL, Tabak M, op den Akker H, Hermens H. Design features of embodied conversational agents in eHealth: a literature review. Int J Hum Comput Stud 2020 Jun;138:102409. [doi: 10.1016/j.ijhcs.2020.102409]
48. McDonnell R, Breidt M, Bülthoff HH. Render me real? Investigating the effect of render style on the perception of animated virtual humans. ACM Trans Graph 2012;31(4):91. [doi: 10.1145/2185520.2185587]
49. Mori M, MacDorman KF, Kageki N. The uncanny valley. IEEE Robot Autom Mag 2012;19(2):98-100. [doi: 10.1109/MRA.2012.2192811]

XSL•FO
RenderX

50. Stein C. Uncanny valley in virtual reality. In: Lee N, Lee N, editors. Encyclopedia of Computer Graphics and Games: Springer International Publishing; 2018:1-3. [doi: 10.1007/978-3-319-08234-9_177-1]
51. Chi NC, Sparks O, Lin SY, Lazar A, Thompson HJ, Demiris G. Pilot testing a digital pet avatar for older adults. Geriatr Nurs 2017 Nov;38(6):542-547. [doi: 10.1016/j.gerinurse.2017.04.002]
52. Bott N, Wexler S, Drury L, et al. A protocol-driven, bedside digital conversational agent to support nurse teams and mitigate risks of hospitalization in older adults: case control pre-post study. J Med Internet Res 2019 Oct 17;21(10):e13440. [doi: 10.2196/13440] [Medline: 31625949]
53. Schwind V, Leicht K, Jäger S, Wolf K, Henze N. Is there an uncanny valley of virtual animals? A quantitative and qualitative investigation. Int J Hum Comput Stud 2018 Mar;111:49-61. [doi: 10.1016/j.ijhcs.2017.11.003]
54. Sierra Rativa A, Postma M, van Zaanen M. The uncanny valley of a virtual animal. Comput Animation Virtual 2022 Mar;33(2):e2043. [doi: 10.1002/cav.2043]
55. Dahlbäck N, Jönsson A, Ahrenberg L. Wizard of Oz studies—why and how. Knowl Based Syst 1993 Dec;6(4):258-266. [doi: 10.1016/0950-7051(93)90017-N]
56. Heinonen E, Nissen-Lie HA. The professional and personal characteristics of effective psychotherapists: a systematic review. Psychother Res 2020 Apr;30(4):417-432. [doi: 10.1080/10503307.2019.1620366] [Medline: 31122157]
57. Orlinsky DE, Rønnestad MH, Collaborative Research Network of the Society for Psychotherapy Research. How Psychotherapists Develop: A Study of Therapeutic Work and Professional Growth: American Psychological Association; 2005. [doi: 10.1037/11157-000]
58. DeVault D, Artstein R, Benn G, et al. SimSensei kiosk: a virtual human interviewer for healthcare decision support. Presented at: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems; May 5-9, 2014; New York, New York p. 1061-1068. [doi: 10.65109/MXIV3169]
59. Rheu M, Shin JY, Peng W, Huh-Yoo J. Systematic review: trust-building factors and implications for conversational agent design. Int J Hum Comput Interact 2021 Jan 2;37(1):81-96. [doi: 10.1080/10447318.2020.1807710]
60. Schrepp M, Hinderks A, Thomaschewski J. Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. In: Marcus A, Marcus A, editors. Design, User Experience, and Usability Theories, Methods, and Tools for Designing the User Experience: Springer International Publishing; 2014:383-392. [doi: 10.1007/978-3-319-07668-3_37]
61. Franke T, Attig C, Wessel D. A personal resource for technology interaction: development and validation of the Affinity for Technology Interaction (ATI) Scale. Int J Hum Comput Interact 2019 Apr 3;35(6):456-467. [doi: 10.1080/10447318.2018.1456150]
62. Arnfred B, Svendsen JK, Adjourlu A, Horthøj C. Scoping review of the hardware and software features of virtual reality exposure therapy for social anxiety disorder, agoraphobia, and specific phobia. Front Virtual Real 2023;4:952741. [doi: 10.3389/frvir.2023.952741]
63. Schubert T, Friedmann F, Regenbrecht H. The experience of presence: factor analytic insights. Presence Teleoperators Virtual Environ 2001 Jun;10(3):266-281. [doi: 10.1162/105474601300343603]
64. Dörner R, Broll W, Grimm P, Jung B, editors. Virtual Und Augmented Reality (VR/AR): Springer Vieweg; 2019.
65. Volonte M, Babu SV, Chaturvedi H, et al. Effects of virtual human appearance fidelity on emotion contagion in affective inter-personal simulations. IEEE Trans Visual Comput Graphics 2016;22(4):1326-1335. [doi: 10.1109/TVCG.2016.2518158]
66. Hoffmann L. That robot touch that means so much: on the psychological effects of human-robot touch [doctoral dissertation]. : DuEPublico [University Duisburg-Essen]; 2017 URL: https://duepublico2.uni-due.de/receive/duepublico_mods_00043331 [accessed 2026-08-01]
67. Hoffmann L, Krämer NC. The persuasive power of robot touch. Behavioral and evaluative consequences of non-functional touch from a robot. PLoS One 2021;16(5):e0249554. [doi: 10.1371/journal.pone.0249554] [Medline: 33951034]

Abbreviations

ATI: Affinity for Technology Interaction
ECA: embodied conversational agent
UEQ-S: User Experience Questionnaire Short Version
VR: virtual reality
VRET: virtual reality exposure therapy

XSL•FO
RenderX

Edited by I Steenstra; submitted 30.Jul.2025; peer-reviewed by A Gupta, SS Kamath, S Mittal, Y Yang; revised version received 13.Jul.2026; accepted 13.Jul.2026; published 07.Aug.2026.

Please cite as:
Hallmann MC, Peperkorn NL, Vona F, Shiban Y, Voigt-Antons JN
User Experience and Socioemotional Evaluation of Virtual Therapist Character Types in a Simulated Virtual Reality Exposure Therapy: Quantitative Study
JMIR XR Spatial Comput 2026;3:e81537
URL: https://xr.jmir.org/2026/1/e81537
doi:10.2196/81537

© Michelle Celina Hallmann, Nicolina Laura Peperkorn, Francesco Vona, Youssef Shiban, Jan-Niklas Voigt-Antons. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 7.Aug.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Immersive Simulation Training in Health Informatics: Qualitative Descriptive Study

Jennifer Skelton, DHSc; Savannah Siebers; Alanna Pelmore

Department of Health Informatics and Information Management, University of Southern Indiana, 8600 University Blvd, Evansville, IN, United States

Corresponding Author:
Jennifer Skelton, DHSc
Department of Health Informatics and Information Management, University of Southern Indiana, 8600 University Blvd, Evansville, IN, United States

Abstract

This qualitative research study offers preliminary evidence that immersive simulation training can influence change management skills, confidence, and perceived usefulness among health informatics professionals.

(JMIR XR Spatial Comput 2026;3:e106961) doi:10.2196/106961

KEYWORDS

simulation training; virtual reality; qualitative research; change management; end user training; health informatics

Introduction

As health care technology success relies on people as much as design, preparing health informaticists to navigate stakeholder dynamics and sustain change is critical. However, when it comes to transformation in health care, the industry has an "implementation problem" because achieving success, meaning that desired outcomes are realized, has become difficult [1]. A systematic review of the change management practices and outcomes literature found that the choice of management practice and leaders' engagement with followers were important differentiators between project success and failure [1]. Commonly used models include Kotter's 8-Step Process, the Prosci Methodology, and the PEPE model [2-4].

Employers of health informaticists recognize this need for "soft skills" for health IT implementation. Flite et al [5] analyzed job advertisements for health care data analysts, a subset role within health informatics. The researchers found that 62% of postings desired applicants with interpersonal skills, and 6% specifically called for change management skills [5].

The development of virtual reality (VR) technology has enabled clinical skill development in health care education and practices [6]. Experiences with authentic simulations are considered a motivating factor in VR health care scenarios, and VR can also provide a safe environment for professionals to practice essential nonclinical skills [6]. Although simulation is used effectively in health care education, research on developing change management skills among health informaticists remains underexplored.

This study aimed to generate preliminary data on health informaticists' perceptions of how immersive simulation training influences change management skills, confidence, and decision-making.

Methods

Study Design

This qualitative descriptive study explored participants' experiences and confidence in skill development after completing an immersive simulation. Data were collected from February 13 to March 20, 2026. Of 14 people purposively recruited via snowball sampling, 9 health informatics professionals provided informed consent and completed either a desktop web-based or head-mounted display (Meta Quest 3) version of the simulation, custom-built using the Cornerstone Immerse CoPilot Designer, which enables emotionally realistic role-play with virtual human characters [7] (Figure 1).

XSL•FO
RenderX

Figure 1. Cornerstone Immerse interaction with Wanda, MD.

After simulation completion, participants concluded with a semistructured online Qualtrics interview. The three coauthors conducted reflexive thematic analysis by each familiarizing themselves with the data, iteratively summarizing responses into categories at weekly meetings, developing and reviewing themes, and contextualizing findings against the study aims, with disagreements resolved by consensus [8]. Weekly discussions focused particularly on responses to questions addressing confidence, decision-making, application to daily work, and usefulness, allowing the team to track the depth and consistency of responses across interviews. The team collectively decided to discontinue data collection after 9 participants, judging that continued interviews were enriching existing categories rather than substantively expanding them.

The change management module "Guiding Lasting Change" was created to allow real-time conversations with change-resistant health care staff in a realistic health IT change scenario. Informed by various change management models [2-4], the simulation provided structured practice in core interpersonal skills essential for health informatics professionals, including communication, conflict management, self-awareness, and understanding others' perspectives.

The module used instructional scaffolding during gameplay. Participants began by observing Jade, the nurse manager of a cardiac unit, delivering a discharge-planning project update. After a brief change management knowledge lesson from two virtual coaches, participants role-played as health informatics consultants negotiating one-on-one interactive exercises with Wanda, MD, and Keegan, PharmD. These change-resistant clinical team members allowed participants to practice active listening, empathy, and asking probing questions. Participants began the interactions by selecting from branching,

XSL•FO
RenderX

dialogue-driven exchanges accompanied by immediate feedback, then advanced to an open-ended, unscripted dialogue with a role-based prompted AI avatar. A final postassessment activity with James, RN, provided participants with scored feedback. The entire 5-lesson module was completed in an average of 12.3 minutes.

Ethical Considerations

Ethics approval was obtained from the institutional review board of the University of Southern Indiana, Evansville, Indiana (2411349‐1). Written informed consent was obtained from all participants before enrollment. Participant confidentiality was maintained by using deidentified data.

Results

Table 1 shows participant insights from the simulation intervention. Themes around confidence, decision-making, daily practice, and usefulness emerged. Health informaticists reported increased confidence in change management skills after the simulation. Approximately one-third of participants said that open-dialogue practice with the pharmacist AI avatar influenced their change management skills due to the critical thinking required and the way it reflected real-time conversations. Participants said the simulation influenced decision-making in change-related situations by reinforcing active listening, consideration of end user perspectives, and empathy. Participants described the experience as more engaging and realistic than traditional training methods, such as videos or readings, and said it prepared them for difficult conversations with end users of implementation projects. Several participants noted that verbal, free-response interactions were particularly effective at simulating real-world scenarios.

Table . Study themes and supporting participant quotes.

Theme	Participant quote
Confidence	“I feel more confident in my ability to manage the social aspects of change management... I’m used to the logistical concerns, but I tend to struggle with the social side.”
Usefulness	“Simulating a scenario where you practice working with frustrated end users makes you more prepared when you have to address this in real life... the verbal option more closely simulates real life and requires you to think on your feet.”
Decision-making	“It gave real-time interactions that forced me to think quickly... it emphasized empathy, active listening, and reinforced the need for open-ended questions.”
Daily practice	“I will continue to use what I was taught when it comes to listening, asking questions, and showing empathy... I want to make sure I handle every situation gracefully.”

Discussion

Principal Findings

Overall, these results suggest that immersive simulation training could be an effective skill-building tool for developing change management competencies and may serve as a valuable addition to professional development in health informatics. The immersive simulation enabled practiced engagement with followers, laying a foundation for project success. Participants expressed that the simulation allowed them to apply empathy and critical thinking to mitigate end user frustration and coformulate solutions, and they emphasized that the simulation helped them recognize the importance of de-escalation practices in handling change-related situations productively.

Participants reported intending to apply active listening strategies in daily practice to address concerns and problem-solve with resistant coworkers—interpersonal skills identified as desirable in health informatics job postings [5]. They also described the authentic hospital simulation as useful for entry-level professionals learning to acknowledge end user frustration, further supporting VR’s value as a safe practice environment [6]. These findings support future research into lower-cost immersive simulation and AI alternatives to help health informatics professionals build the change management skills needed for sustained health IT project success.

Although snowball sampling ensured reach among active health informatics professionals, nonprobability sampling may introduce sampling bias, limiting generalizability and the diversity of viewpoints reported [9]. Findings reflected immediate, self-reported perceptions and do not demonstrate long-term behavior change or skill acquisition.

Conclusions

This study has laid the groundwork for future research while highlighting the potential value of developing immersive simulation training modules for nonclinical, change management, and interpersonal skill development among health informatics professionals.

XSL•FO
RenderX

Acknowledgments

We acknowledge the contributions of the study participants and thank them for their support.

The authors declare the use of generative AI (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: defining the research objective; refining research questions and hypotheses; literature search and systematization; and proofreading and editing. The GAI tools used were: ChatGPT 5.2 (OpenAI), Claude Sonnet 5 (Anthropic). 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 by: Jennifer Skelton.

Funding

The development of the Guiding Lasting Change module utilizing Cornerstone Immerse technology was funded by the University of Southern Indiana: Early Career Faculty Grant. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The dataset analyzed during this study is available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JS, SS, AP
Data curation: SS, AP
Formal analysis: JS, SS, AP
Funding acquisition: JS
Investigation: JS, SS, AP
Methodology: JS, SS, AP
Project administration: JS (lead), AP (equal), SS (supporting)
Resources: JS
Software: JS
Supervision: JS
Validation: JS
Visualization: SS (lead); AP (supporting)
Writing – original draft: JS (equal), SS (equal), AP (equal)
Writing – review and editing: JS (lead), SS (supporting), AP (supporting)

Conflicts of Interest

None declared.

References

1. Hastings BJ. Guidance for successful healthcare transformation: a systematic review of change management practices and outcomes. Australian Journal of Management 2025 Aug;50(3):985-1005. [doi: 10.1177/03128962241307347]
2. Kotter JP. Leading change: why transformation efforts fail. Harv Bus Rev 2007;73(2):96-103 [FREE Full text]
3. Change management: driving change success by preparing, equipping, and supporting individuals to thrive through change. Prosci. URL: https://www.prosci.com/change-management [accessed 2025-06-21]
4. Vera T, Franklin M. Neuroscience for Change at Work: Practical Insights to Overcome Workforce Resistance to Organizational Change: Kogan Page; 2024.
5. Flite CA, Foster S, Houser SH, et al. Essential skill and knowledge required for health data professionals: a content analysis of job advertisements. Perspect Health Inf Manag 2024;21(2):1c. [Medline: 40134896]
6. Mäkinen H, Haavisto E, Havola S, Koivisto JM. User experiences of virtual reality technologies for healthcare in learning: an integrative review. Behav Inf Technol 2022 Jan 2;41(1):1-17. [doi: 10.1080/0144929X.2020.1788162]
7. Cornerstone Immerse. Cornerstone OnDemand. URL: https://www.cornerstoneondemand.com/platform/immersive-learning/ [accessed 2026-07-07]
8. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
9. Bhattacherjee A. Social Science Research: Principles, Methods, and Practices, 2nd edition: University of South Florida; 2012. URL: https://digitalcommons.usf.edu/oa_textbooks/3 [accessed 2026-01-17]

Abbreviations

VR: virtual reality

XSL•FO
RenderX

Edited by I Steenstra; submitted 13.Jul.2026; peer-reviewed by D Bouldrick, GA Diolaiuti, Z Du; revised version received 04.Sep.2026; accepted 07.Sep.2026; published 17.Sep.2026.

Please cite as:
Skelton J, Siebers S, Pelmore A
Immersive Simulation Training in Health Informatics: Qualitative Descriptive Study
JMIR XR Spatial Comput 2026;3:e106961
URL: https://xr.jmir.org/2026/1/e106961
doi:10.2196/106961

© Jennifer Skelton, Savannah Siebers, Alanna Pelmore. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 17.Sep.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perceived Usability, User Experience, and Technology Acceptance of Role-Specific Augmented Reality Decision Support Tools for Cardiac Arrest Resuscitation: Prospective Observational Pilot Study

Ryan Kang[1], MSc; Adam Cheng[2,3], MD; Yiqun Lin[2,3], MD, MHSc, PhD; Hyeongil Nam[1], PhD; Jennifer Davidson[2,3], RN; Donovan Curtis Duncan[2,3], MD; Johan N Siebert[4,5], MD, PD; Sergio Manzano[4,5], MD, PD; Alexandre De Masi[5], PhD; Ana Rajic[4,6,7], MS; Sharleen Kayne Olanka[4,6,7], MS; Frederic Ehrler[5,7], PhD; Kangsoo Kim[1], PhD

[1]Department of Electrical and Software Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada

[2]KidSIM Simulation Program, Alberta Children's Hospital, Calgary, AB, Canada

[3]Department of Pediatrics and Emergency Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[4]Department of Pediatric Emergency Medicine, Geneva University Hospitals, Geneva, Switzerland

[5]Faculty of Medicine, University of Geneva, Geneva, Switzerland

[6]Educational Technologies and Learning Sciences (TECFA), Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland

[7]Division of Computer Sciences, Geneva University Hospitals, Geneva, Switzerland

Corresponding Author:
Kangsoo Kim, PhD
Department of Electrical and Software Engineering, Schulich School of Engineering, University of Calgary, 2500 University Drive NW, Calgary, AB, Canada

Abstract

Background: Cardiac arrest is a critical medical emergency that requires strict adherence to clinical guidelines to achieve optimal outcomes. Deviations from these guidelines, often due to task complexity, can adversely affect patient outcomes. Augmented reality (AR) offers a way to deliver role-specific, in-view guidance, but evidence on its perceived usability, user experience, and acceptability in cardiac arrest resuscitation remains limited.

Objective: This study aimed to design, develop, and evaluate a role-specific AR decision support system for resuscitation team leaders and medication nurses. In this observational study, we assessed clinicians' perceived usability, user experience, and technology acceptance of the new AR system in a high-fidelity simulated cardiac arrest scenario.

Methods: We conducted a prospective observational pilot study using a high-fidelity simulated pediatric cardiac arrest scenario. A total of 10 clinicians were recruited from Alberta Children's Hospital, including 5 (50%) of 10 pediatric emergency physicians serving as team leaders (men: 3/5, 60%, and women: 2/5, 40%; median age 41, IQR: 40-42 y) and 5 (50%) of 10 emergency nurses serving as medication nurses (men: 1/5, 20%, and women: 4/5, 80%; median age 45, IQR: 42-46 y). Participants used role-specific AR decision support interfaces deployed on HoloLens 2 head-mounted displays. Following the simulation, perceived usability, user experience, and technology acceptance were assessed using validated questionnaires: the System Usability Scale, User Experience Questionnaire, and Technology Acceptance Model. Data were collected via postsimulation surveys and analyzed descriptively.

Results: Descriptive analyses were performed without inferential statistical testing. The mean System Usability Scale scores were 75.5 (SD 9.25, 95% CI 64.0 - 87.0) for team leaders and 82.0 (SD 11.20, 95% CI 68.0 - 96.0) for medication nurses. User experience was positive across roles, with mean User Experience Questionnaire scores indicating favorable attractiveness (team leaders: 1.87, SD 1.14, 95% CI 0.45 - 3.28; medication nurses: 2.43, SD 0.52, 95% CI 1.79 - 3.08), pragmatic quality (team leaders: 1.88, SD 0.87, 95% CI 0.80 - 2.97; medication nurses: 1.80, SD 0.69, 95% CI 0.94 - 2.66), and hedonic quality (team leaders: 2.40, SD 0.89, 95% CI 1.30 - 3.50; medication nurses: 2.28, SD 0.69, 95% CI 1.42 - 3.13). Technology acceptance was high, with mean combined Technology Acceptance Model scores of 5.92 (SD 0.46, 95% CI 5.35 - 6.49) for team leaders and 6.02 (SD 0.56, 95% CI 5.32 - 6.71) for medication nurses.

Conclusions: This study introduces a novel role-specific AR decision support system that delivers tailored, in-view guidance to resuscitation team leaders and medication nurses during cardiac arrest. Unlike prior cognitive aids that present uniform or device-agnostic information, this system explicitly adapts interface content and structure to distinct clinical roles and workflows. The findings contribute early empirical evidence on the perceived usability, user experience, and acceptability of role-tailored

XSL•FO
RenderX

AR support in high-acuity team settings and yield transferable design principles for developing role-aware AR interfaces. In real-world contexts, such systems may support protocol adherence and team coordination during resuscitation training and early-stage clinical deployment, informing future evaluations that incorporate objective performance and workflow outcomes.

(JMIR XR Spatial Comput 2026;3:e72013) doi:10.2196/72013

KEYWORDS

cardiopulmonary resuscitation; augmented reality; simulation training; wearable electronic devices; digital health; user-computer interface; decision support systems; clinical guideline adherence; technology acceptance; user-centered design

Introduction

Cardiopulmonary resuscitation (CPR) is administered to thousands of patients experiencing cardiac arrests (CAs) each year in North America [1]. Guideline-compliant basic life support and advanced life support guidelines significantly improve patient outcomes following CA [2,3]. However, health care providers often face challenges in consistently adhering to these guidelines during in-hospital CA events. Deviations, such as delays in epinephrine administration, defibrillation, and medication dosing errors, are commonly linked to poor patient outcomes [4]. These deviations are often attributed to the high cognitive demands and mental workload experienced by resuscitation team members [5,6].

Cognitive aids, designed to assist in decision-making and information recall, have demonstrated improved adherence to resuscitation guidelines during simulated cardiopulmonary arrest events [7-11]. By reducing errors and improving the timing of key interventions, cognitive aids can enhance clinical performance [12]. However, traditional cognitive aids, such as pocket cards, sometimes introduce delays in initiating CPR or administering drugs due to their design limitations or complexity, highlighting the need for more efficient, role-specific decision-support solutions. Recent scoping and systematic reviews published in the past few years highlight a growing interest in immersive technologies, including augmented reality (AR), for resuscitation training and emergency care, while also identifying variability in system design, evaluation approaches, and integration with clinical workflows [13,14].

AR overlays digital content onto the physical environment, enabling real-time delivery of context- and role-specific prompts directly in the user's field of view [15]. AR systems have been explored in CPR and emergency care training contexts, with some evidence of improved engagement and task performance compared with conventional approaches, although results remain heterogeneous and context-dependent [16,17]. Previous AR-based work in resuscitation and safety-critical domains further suggests that spatially registered visual cues can support situational awareness and reduce reliance on external reference materials during time-sensitive tasks [15,18,19]. Despite this growing body of work, recent reviews emphasize that evidence regarding the usability, user experience, and acceptability of wearable AR systems in CA resuscitation—particularly from the perspective of end users—remains limited [13,14].

To address these gaps, this study presents the design, development, and formative evaluation of an AR-based decision support system tailored to the resuscitation team leader (physician) and medication nurse roles during CA resuscitation. The objectives of this study were to (1) describe how team leaders and medication nurses perceive the AR system's usability and user experience when used during a simulated resuscitation scenario and (2) describe how team leaders and medication nurses perceive the system's acceptability and its potential for future integration into clinical practice.

Methods

Ethical Considerations

Ethics approval was obtained from the University of Calgary (REB23-1007) and the University of Geneva Health Research Ethics Boards (Req-2023‐00162). Before participation, all participants were provided with written information describing the study purpose, procedures, potential risks, and data handling practices, and written informed consent was obtained. Participation was voluntary, and participants were informed that they could withdraw from the study at any time without consequence. Consent included permission to collect survey data and to use nonidentifiable data and images generated during the simulation for research and publication purposes. Privacy and confidentiality were ensured for all study participants. No images included in the manuscript or supplementary materials contain identifiable information about individual participants. Participants did not receive any compensation for their role in this study.

Study Design: Experimental Setting

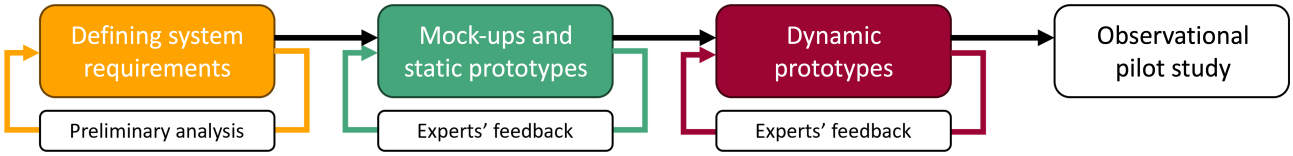
This study was designed as a prospective observational pilot study conducted in a high-fidelity pediatric CA simulation setting. The following section describes the overall process of the AR system design and development, which was used in the study.

Iterative Design and Development Process of the AR System

For our study, role-specific AR decision support systems for team leaders (physicians) and medication nurses were developed following a 4-phase, iterative prototyping process grounded in user-centered and clinician-informed design practices (Figure 1). Phases 1 to 3 focused on system design and development, whereas phase 4 evaluated the final prototype in a simulation-based clinical environment through an observational pilot study. The objective of this process was to progressively refine AR design concepts into a stable, simulation-ready system through iterative feedback and close collaboration with clinical domain experts.

XSL•FO
RenderX

Figure 1. Four-phase iterative development process.



Phase 1: Defining System Requirements

This phase focused on identifying the clinical, informational, and workflow requirements necessary to guide the content and design of the role-specific AR interfaces. A total of 30 health care professionals (15, 50%, emergency physicians and 15, 50%, emergency nurses) from Alberta Children's Hospital (ACH) and Geneva University Hospitals were surveyed to assess preferences for role-specific information, AR layout components, timer placement and behavior, and medication-related display features. Clinicians viewed role-specific, task-focused information as important elements of the AR system. Both physicians and nurses emphasized the utility of receiving targeted, step-relevant prompts through the AR headset. Real-time updates regarding current and upcoming tasks ("next steps") were perceived to enhance workflow by reducing the need to reference external materials visually. Both groups rated time-based cues highly, with the integration of a CPR timer (for the team leader) and an epinephrine timer (for the team leader and medication nurse) described as highly important for the AR headset. A detailed list of resuscitation medications and associated dosages was also rated highly for both groups of providers. Physicians expressed the desire to be notified when medications were given. These insights directly informed the layout, information hierarchy, and alerting behavior of the static (phase 2) and dynamic (phase 3) prototypes, ensuring that the AR interface design aligned with clinicians' informational needs and workflow demands. Table 1 provides a summary of key insights from phase 1.

XSL•FO
RenderX

Table . Key insights from each phase of the iterative development process.

Phase	Key insights
Phase 1: Defining system requirements	• Role-specific, task-focused clinical prompts • Current tasks (prioritized) • Next steps (prioritized) • CPR[a] timer (team leader) • Epinephrine timer (team leader and medication nurse) • Medication reference • Medication given—notification for team leader
Phase 2: Mock-ups and static prototypes	• Interface separation by user role improves clarity and relevance of displayed information. • Visually simple, structured layouts are preferred for rapid information recognition. • Timers and other time-sensitive elements should be placed in the upper peripheral field of view to avoid obstructing the patient. • Current tasks should be listed on the left and next steps to the right. • Cardiac rhythm should be displayed in the physician's augmented reality headset. • Patient weight should be displayed on the medication nurse display. • Other UI[b] elements should be fixed in space (e.g., clinical algorithm, Hs and Ts, medication reference) to avoid interference when team members move through the field of view.
Phase 3: Dynamic prototype	• Functional CPR and epinephrine timers included escalating visual cues. • The UI was visually refined with higher contrast, low-profile components, reorganized medication content, and larger fonts. • Medication card (team leader): categorized drug details with interactive dose counters • Guideline algorithm panel (team leader): full cardiac arrest algorithm visualization with a stage-tracking arrow. • Hs and Ts reference (team leader): a structured list of reversible causes for rapid diagnostic review. • Medication card (medication nurse): categorized drugs with strength, dose, volume, and instructions, plus an interactive syringe counter for tracking prepared or administered doses. • UI elements were arranged to maximize visibility and minimize occlusion during dynamic resuscitation.

[a]CPR: cardiopulmonary resuscitation.

[b]UI: user interface.

Phase 2: Mock-Ups and Static Prototypes

In this phase, the requirements identified in phase 1 were transformed into static prototype designs. Initial mock-ups were created to visualize the AR layout, role-specific information elements, and overall display functionality. To optimize role-specific design, separate static layouts were developed for the team leader and medication nurse roles. The team leader interface focused on 4 key elements (i.e., the CPR timer, epinephrine timer, current task list, and next task list), whereas the medication nurse interface incorporated 3 core elements (i.e., epinephrine timer, current task list, and next task list). A total of 9 static prototypes were created for the team leader, and 5 prototypes were created for the medication nurse, exploring variations in spatial arrangement and visual hierarchy (Figures S1 and S2 in Multimedia Appendix 1).

In total, 5 emergency room physicians and 5 emergency room nurses from ACH were selected to provide feedback on static layouts for their corresponding profession. Participants were shown each static layout in sequence and asked to provide verbal feedback regarding spatial organization, information grouping, font and icon size, color and contrast of user interface (UI) elements, and the position of UI elements relative to equipment and providers in the clinical space. Physicians rated their top 3 display options, and nurses were asked to rate their top 2 options. Feedback was documented using annotated screenshots and meeting notes. On the basis of this feedback, 3 dynamic prototypes for the team leader and 2 dynamic prototypes for the medication nurse were developed to ensure that UI elements were easy to identify, accessible, and minimally intrusive within the AR field of view. Table 1 provides a summary of key insights from phase 2.

Phase 3: Dynamic Prototypes

This phase involved developing and iteratively refining dynamic AR prototypes for both the team leader and medication nurse roles. Dynamic interface layouts were implemented using Unity and deployed on the Microsoft HoloLens 2. The initial dynamic versions preserved the core components established during the static prototyping phase, while introducing functional timers,

XSL•FO
RenderX

refined visual elements, and interactive components. For the medication nurse interface, an adjustable epinephrine dose counter was implemented, allowing users to adjust the number of doses prepared or administered. To guide iterative refinement, the 10 participants from phase 2 returned to provide feedback on the dynamic prototypes. Participants were asked to evaluate layout preferences, timer behavior, visual clarity, and ease of interaction. Additional role-specific questions were directed to team leaders and medication nurses to capture feedback aligned with each role's clinical responsibilities.

Feedback sessions identified refinements to visual hierarchy, timer behavior, text sizing, and the placement of role-specific components. Experts also evaluated the positioning of fixed elements such as the CA algorithms, reversible causes (Hs and Ts), and the medication card, providing insight into potential visual obstruction during active resuscitation. Feedback informed key improvements to support clarity, usability, and workflow alignment. Functional CPR and epinephrine timers were revised to include escalating visual cues (yellow flash at 10 s and rapid red flash at 1 s). The UI was visually refined with higher contrast, low-profile components, reorganized medication content, and larger fonts. An interactive epinephrine dose counter was added for medication nurses. Participants also emphasized the need to reposition or hide large reference panels to prevent obstruction and maintain clear grouping of current and upcoming tasks. Table 1 provides a summary of key insights from phase 3. All recommended changes were incorporated into one final updated dynamic prototype for the team leader (Figure 2) and medication nurse (Figure 3).

Figure 2. Team leader display showcasing real-time CPR and medication timers, visual alerts for task progression, and stepwise guidance for current and upcoming guideline tasks during a simulated cardiac arrest (CA) scenario. When in use, the CA algorithm, reversible causes, and medication card are positioned out of view when the team leader is looking straight ahead. To view each of these items, the team leader must turn to the left (to see the algorithm), to the right (to see the reversible causes), or look slightly down (to see the medication card). CPR: cardiopulmonary resuscitation; PEA: pulseless electrical activity; pVT: pulseless ventricular tachycardia; ROSC: return of spontaneous circulation; TEP: Treatment Escalation Plan; VF: ventricular fibrillation.

XSL•FO
RenderX

Figure 3. Medication nurse display showing step-by-step guidance on drug dosages, preparations, and administration timing, with a real-time epinephrine timer for ensuring timely interventions.

Final System Architecture and Components

The AR system used in this study was developed using a server-client architecture to enable seamless, real-time synchronization between a web-based control system operated by the experimenter and the role-specific AR interfaces used by the team leader and medication nurse (Figure S3 in Multimedia Appendix 1). This architecture ensured that each user received only the information relevant to their role while maintaining consistent timing, event updates, and algorithm progression across devices.

- Web-based control system (server): A centralized web-based control system was implemented to manage scenario flow and synchronize data to both AR devices (Figure S4 in Multimedia Appendix 1). During the simulation, the experimenter used this interface to advance the CA algorithm, trigger event notifications, reset timers, and record medication administration (Figure S5 in Multimedia Appendix 1). All adjustments made on the server were immediately transmitted to the AR clients, enabling real-time display without perceptible delay.
- AR interfaces for team leader and medication nurse (client): Two separate AR client applications were deployed on the HoloLens 2 devices, one for each role. These interfaces displayed synchronized timers, role-specific prompts, algorithm guidance, medication information, and interactive elements (e.g., dose counters). The client applications integrated incremental refinements derived from clinician feedback during dynamic prototyping, ensuring that the displays aligned with each role's workflow and cognitive demands.

Together, the server-client architecture, real-time synchronization, and role-specific display features formed a cohesive system for supporting resuscitation team members during high-acuity pediatric CA scenarios.

Phase 4: Simulation-Based System Evaluation

Phase 4 consisted of a prospective, observational pilot study in which participants managed a simulated CA scenario using the final prototype of the AR system.

Participants and Sample Size

Participants were recruited from the pediatric emergency department at ACH. All participants had completed basic life support and pediatric advanced life support training. There were no specific exclusion criteria. A convenience sample of 10 health care professionals participated, consisting of 5 (50%) pediatric emergency physicians (team leaders) and 5 (50%) emergency nurses (medication nurses). The same 10 participants who provided feedback in phases 2 and 3 were paired into physician-nurse dyads, with each clinician assigned the AR interface corresponding to their respective profession.

Study Procedure: Simulated CA Scenario

The simulation scenarios took place in the KidSIM Pediatric Simulation Center at ACH using a high-fidelity pediatric manikin (Laerdal SimJunior). Each dyad (1 physician team leader and 1 medication nurse) was embedded within a larger clinical resuscitation team composed of 3 additional research actors playing the roles of airway provider, bedside clinician, and CPR provider to recreate an authentic team-based resuscitation environment. The 2 study participants wore HoloLens 2 devices displaying their respective role-specific AR interfaces.

The scenario simulated an in-hospital pediatric CA involving a 5-year-old boy who presented with pulseless ventricular tachycardia, progressing through ventricular fibrillation and

XSL•FO
RenderX

pulseless electrical activity, before achieving return of spontaneous circulation at the 18-minute mark. Participants, acting as team leader or medication nurse, were guided by visual prompts on their respective AR displays. The team leader guided overall clinical management, including airway management, CPR, defibrillation, and ordering medications. The medication nurse handled medication preparation and administration, following role-specific cues on the AR interface. Research actors were trained to function in their role as they would in a real CA.

Measures

To provide a comprehensive assessment of the AR support system's perceived usability, user experience, and acceptance, we used 3 well-established instruments. The System Usability Scale (SUS) was used to measure perceived usability. It consists of 10 statements that assess users' perceptions of system ease of use and overall usability [20,21]. Each statement is rated on a 5-point Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (5), capturing both ease of use and learnability. SUS scores are calculated by first adjusting responses: for odd-numbered items, 1 is subtracted from the user's rating, and for even-numbered items, the rating is subtracted from 5. The adjusted scores for each statement are summed, and the total is multiplied by 2.5 to convert the raw score to a range of 0 to 100. On the basis of empirical benchmarks reported by Bangor et al. [21], SUS scores above 68 are generally interpreted as above average, whereas scores around 80 or higher are commonly associated with excellent usability. These benchmarks provide a practical reference for interpreting system usability levels. The SUS has demonstrated strong psychometric properties across diverse systems and application domains, including high internal consistency and established construct validity. Prior validation studies have shown that SUS scores are robust and interpretable even in small-sample usability evaluations, making the instrument suitable for early-stage and pilot studies [20,21].

The User Experience Questionnaire (UEQ) evaluates multiple dimensions of perceived user experience, including *attractiveness*, *pragmatic quality*, and *hedonic quality* [22]. The UEQ consists of 26 items rated on a 7-point semantic differential scale ranging from −3 (most negative) to +3 (most positive), capturing users' subjective impressions of different aspects of system interaction.

- Attractiveness: reflects the overall appeal of the system and represents users' general impression.
- Pragmatic quality: captures perceived task-oriented aspects of the system use, focusing on how well users feel the system supports task accomplishment through three subdimensions: (1) perspicuity: ease of understanding and familiarization, (2) efficiency: perceived smoothness and effort associated with task execution, and (3) dependability: user's perceived sense of control and predictability during interaction.
- Hedonic quality: captures the emotional and experiential aspects of interaction, covering (1) stimulation: how engaging and motivating the system feels; and (2) novelty: perceived originality and creativity of the system.

UEQ scale values above 0.8 are commonly interpreted as indicating a positive experience, whereas higher values may be classified as above average or excellent when compared against UEQ benchmark distributions, depending on the specific scale [23,24]. By distinguishing between pragmatic and hedonic qualities, the UEQ provides insight into both task-oriented interaction perceptions and experiential aspects of system use. The distinction is particularly relevant for AR systems, where perceived interaction support and user engagement jointly shape overall user experience. Validation studies of the UEQ have demonstrated acceptable to good internal consistency across its subscales and established construct validity for distinguishing between pragmatic and hedonic dimensions of user experience across a wide range of interactive systems [22-24].

The Technology Acceptance Model (TAM) assesses user acceptance of new technologies based on the relationship between two main dimensions: (1) perceived usefulness (PU), which measures the extent to which users believe that using a given technology enhances their job performance; and (2) perceived ease of use (PEU), which evaluates the extent to which users believe that using a technology will result in less effort to perform their tasks, focusing on its intuitiveness and the learning curve involved [25]. For this study, TAM was adapted to include 12 items across 2 primary dimensions, each rated on a 7-point Likert scale, ranging from "strongly disagree" (1) to "strongly agree" (7). Scores are averaged for each dimension. High scores across both dimensions suggest that users view the system as both beneficial and user-friendly—key factors for ensuring sustained use [26]. The PU and PEU constructs within TAM have demonstrated strong reliability and predictive validity for technology adoption and use intention across numerous information systems and health care technology studies, supporting their use in evaluating acceptance of emerging technologies, such as AR [25,26].

Statistical Analysis

In this observational pilot study, there were no missing data for survey responses, and all analyses were descriptive in nature and aimed at characterizing perceived usability, user experience, and technology acceptance of the AR system across clinical roles. For each outcome measure, summary statistics were computed separately for the team leader and medication nurse roles. For the SUS, UEQ, and TAM measures, central tendency and variability were summarized using means and SDs. SEs and 95% CIs for the mean were calculated to indicate the precision of the estimates. Where appropriate, medians and IQRs were visualized using box plots to illustrate score distributions.

Given the small sample size and the exploratory nature of this pilot evaluation, no formal hypothesis testing or inferential comparisons between roles were performed. Instead, overlapping CIs were used to support cautious interpretation of observed differences, consistent with recommendations for early-stage usability and feasibility studies.

Results

Participant Demographics

A total of 10 health care professionals participated in the study, comprising 5 (50%) pediatric emergency physicians (team

XSL•FO
RenderX

leaders) and 5 (50%) emergency nurses (medication nurses). Participants varied in age and clinical experience, with medication nurses generally reporting longer durations of practice and greater exposure to CA events. Most participants had limited prior experience with AR technologies, particularly in professional clinical contexts. Table 2 provides an overview of the participants' demographic characteristics.

Table . Participant demographics.

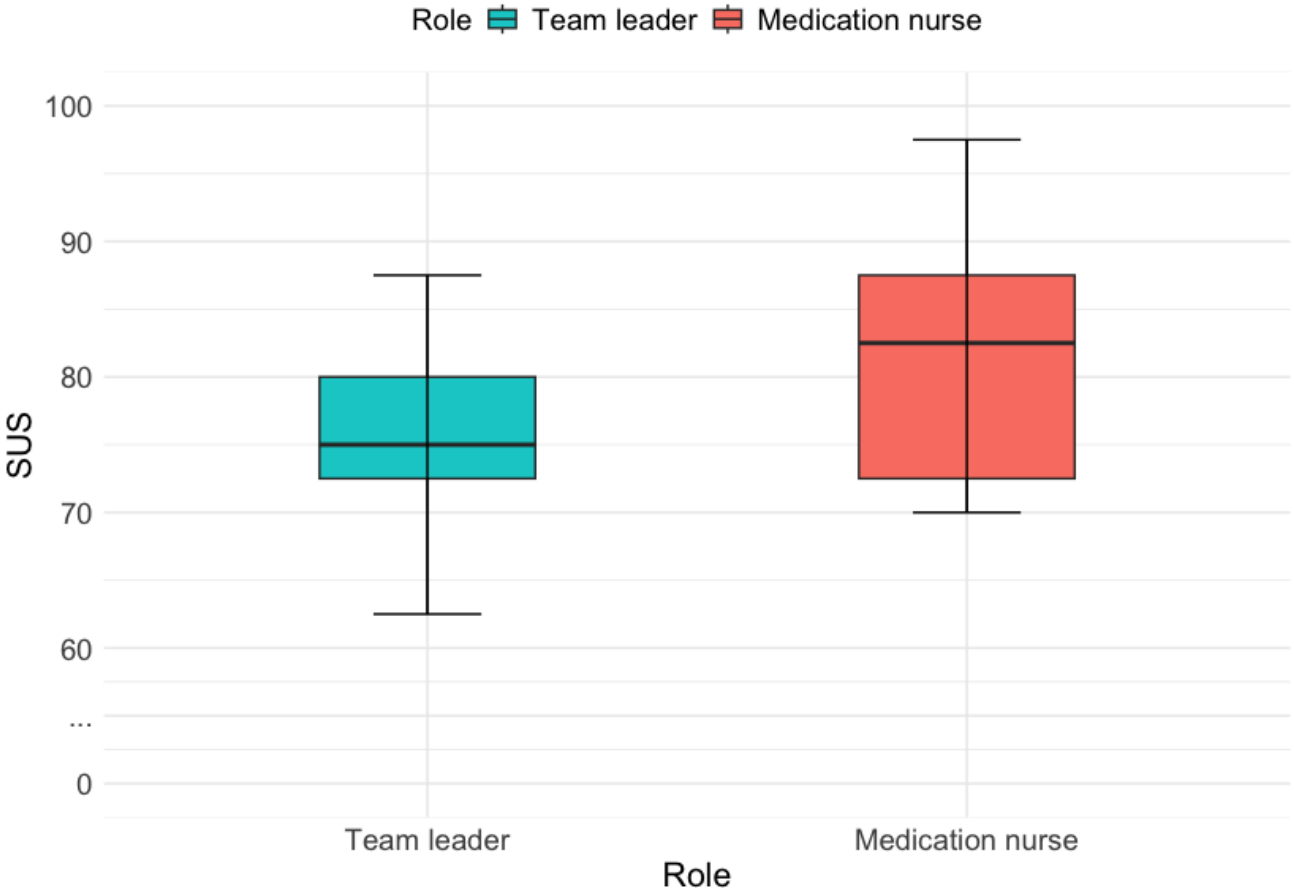
	Team leader (n=5)	Medication nurse (n=5)
Gender, n (%)		
Male	3 (60)	1 (20)
Female	2 (40)	4 (80)
Age (y), median (IQR)	41 (40 - 42)	45 (42 - 46)
Duration in practice (y), median (IQR)	12 (11 - 13)	20.5 (17 - 24.25)
How many times have you had to care for a child in cardiac arrest during a *real, live event* in the past 2 y?, median (IQR)	1 (1 - 4)	2 (2 - 3)
How many times have you had to care for a child in cardiac arrest during a *simulated event* in the past 2 y?, median (IQR)	4 (2 - 4)	5 (4 - 6)
Have you ever used any type of augmented reality device for *professional use*?	1 participant with prior experience (>10 times)	No prior experience
Have you ever used any type of augmented reality device for *recreational use* (e.g., gaming)?	3 participants (>10 times)	1 participant (1 - 4 times)

Perceived Usability

The AR system demonstrated favorable perceived usability for both roles (Figure 4); however, the precision of these estimates varied across roles. The SUS revealed that the team leader role scored a mean of 75.50 (SD 9.25, SE 4.14, 95% CI 64.00-87.00), corresponding to a "B" grade (74.10 - 77.10) on the SUS grading scale, categorized as "good" (Multimedia Appendix 2). The score generally suggests that team leaders perceived the system as usable and user-friendly; however, the relatively wide CI reflects uncertainty associated with the small sample size and indicates that this estimate should be interpreted cautiously.

XSL•FO
RenderX

Figure 4. Box plot displaying System Usability Scale (SUS) scores for the team leader and medication nurse roles using the augmented reality support system. Higher median SUS scores for medication nurses indicate greater ease of interaction and workflow support, reflecting an "excellent" grade compared to the "good" usability rating for the team leader role.



The medication nurse role achieved a higher mean score of 82.00 (SD 11.20, SE 5.02, 95% CI 68.00-96.00), corresponding to an "A" grade (80.80 - 84.00), which falls within the "excellent" usability range (Multimedia Appendix 2). Although the point estimate suggests a stronger perceived usability for medication nurses, the overlapping CIs between roles indicate that differences should not be interpreted as definitive in this pilot study. Overall, both roles reported favorable usability perceptions, with variability reflecting limited precision.

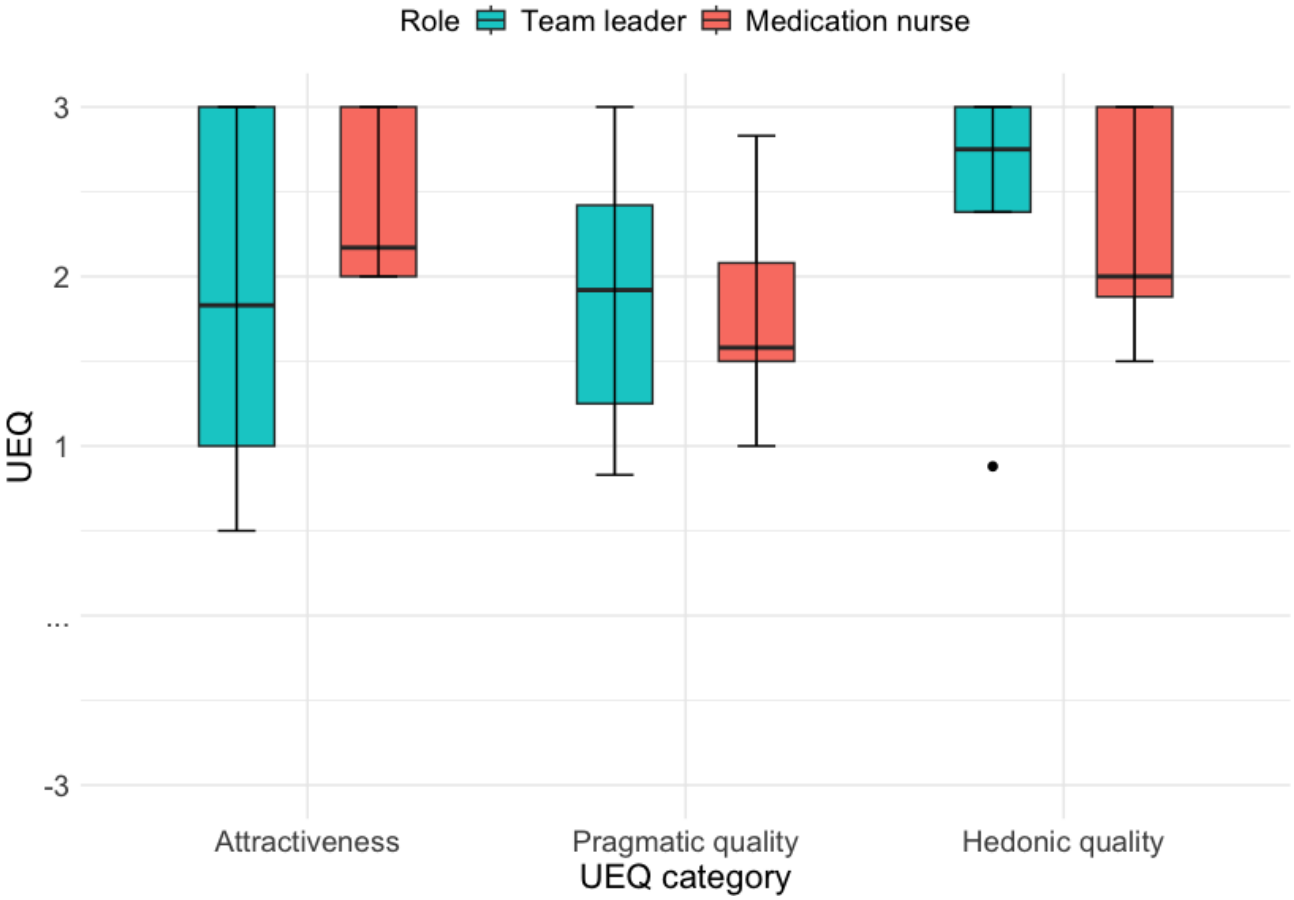
Perceived User Experience

High-Level Results: Attractiveness, Pragmatic Quality, and Hedonic Quality

Perceived user experience was assessed using the UEQ, capturing participants' subjective evaluations across attractiveness, pragmatic quality, and hedonic quality. The following results summarize mean scores and associated uncertainty for each dimension by clinical role (Figure 5).

XSL•FO
RenderX

Figure 5. Comparison of User Experience Questionnaire (UEQ) scores—attractiveness, pragmatic quality, and hedonic quality—across roles. Both groups rated the system well above benchmark levels.



- Attractiveness: The team leader role scored a mean of 1.87 (SD 1.14, SE 0.51, 95% CI 0.45-3.28), whereas the medication nurse role scored higher, with a mean of 2.43 (SD 0.52, SE 0.23, 95% CI 1.79-3.08). Both scores surpass the above-average benchmark, indicating a favorable overall impression of the system's appeal, but the wider CI for team leaders indicates greater variability in perceived appeal (Multimedia Appendix 3).
- Pragmatic quality: Pragmatic quality scores were similarly positive across roles, with team leaders reporting a mean score of 1.88 (SD 0.87, SE 0.39, 95% CI 0.80-2.97) and medication nurses reporting a mean score of 1.80 (SD 0.69, SE 0.31, 95% CI 0.94-2.66). The overlapping CIs suggest comparable perceived task support.
- Hedonic quality: The team leader role scored a mean of 2.40 (SD 0.89, SE 0.40, 95% CI 1.30-3.50), whereas the medication nurse role scored a mean of 2.28 (SD 0.69, SE 0.31, 95% CI 1.42-3.13). These high scores highlight that users perceived the system as engaging and stimulating, contributing to a positive user experience, but the width of the CIs underscores the preliminary nature of these findings.

Pragmatic Quality Subdimensions

The analysis of pragmatic quality subdimensions (Figure S6 in Multimedia Appendix 1) revealed similar patterns across roles.

- Perspicuity: Both the team leader and medication nurse roles reported a mean score of 1.80. The team leader's result (SD 1.14, SE 0.51, 95% CI 0.39-3.21) and the medication nurse's results (SD 0.76, SE 0.34, 95% CI 0.86-2.74) indicated greater variability in perceived ease of learning among team leaders.
- Efficiency: For task completion speed and support, both roles achieved high mean scores of 2.15. The team leader's score (SD 0.68, SE 0.30, 95% CI 1.31-2.99) and the medication nurse's score (SD 0.38, SE 0.17, 95% CI 1.68-2.62) suggest perceived efficiency benefits, although precision remains limited.
- Dependability: The team leader role achieved a mean score of 1.70 (SD 1.02, SE 0.46, 95% CI 0.43-2.97), whereas the medication nurse role scored slightly lower at 1.45 (SD 1.30, SE 0.58, 95% CI −0.17 to 3.07). These scores indicate that users felt a good level of control (predictable), but the CI spanning zero indicates uncertainty regarding perceived control, highlighting this dimension as an area requiring further investigation.

Hedonic Quality Subdimensions

The analysis of hedonic quality subdimensions (Figure S7 in Multimedia Appendix 1) focused on stimulation and novelty, capturing the emotional and experiential aspects of user interaction with the AR system.

- Stimulation: The team leader role achieved a mean score of 2.25 (SD 0.94, SE 0.42, 95% CI 1.09-3.41), whereas the medication nurse role scored similarly at 2.20 (SD 0.84, SE 0.37, 95% CI 1.16-3.24). These scores suggest that the system is engaging and helps sustain users' interest,

XSL•FO
RenderX

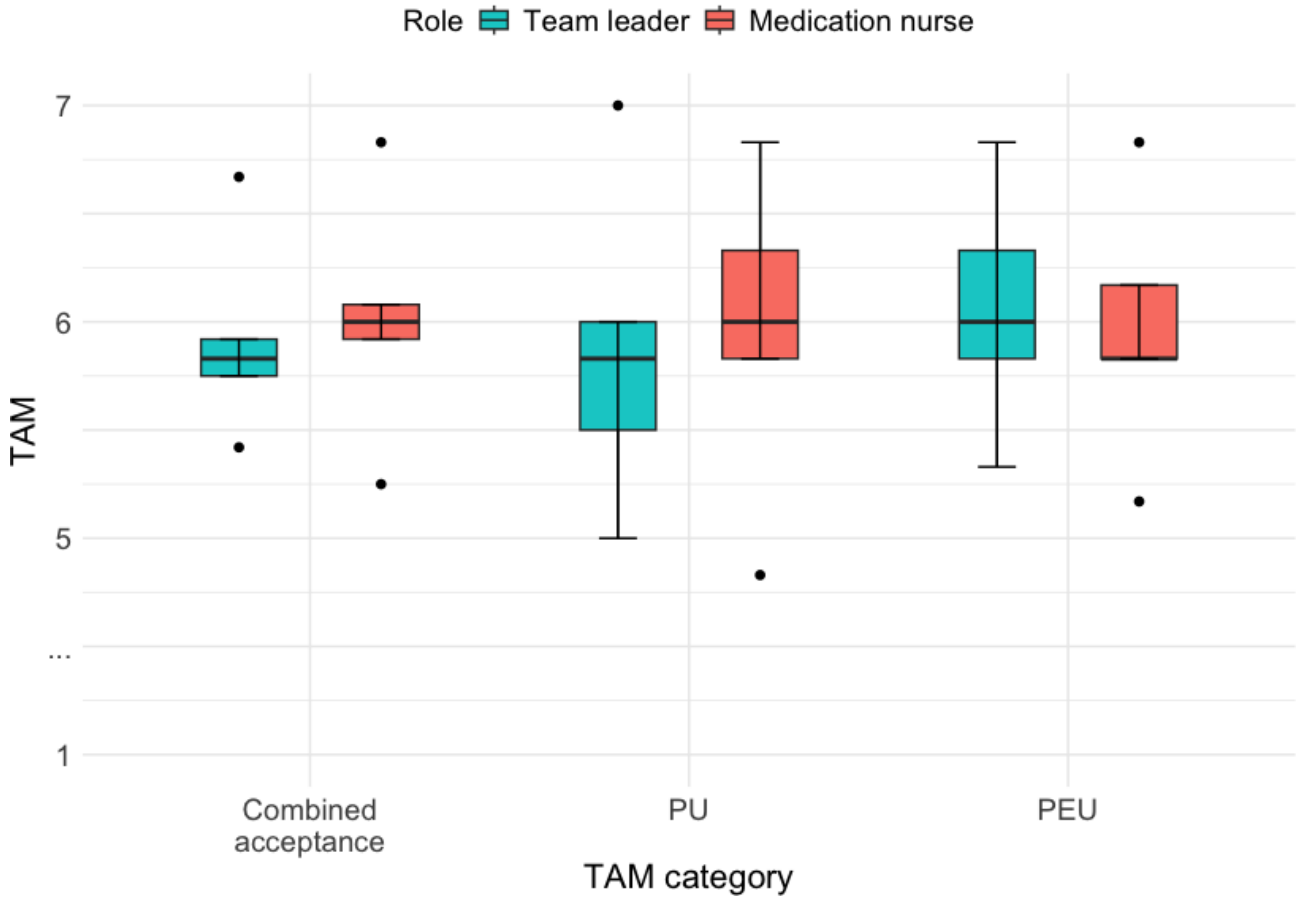
motivating them throughout its use, but overlapping CIs and moderate width reflect limited precision in this pilot evaluation.

- Novelty: This subdimension assesses the system's originality and innovative aspects. The team leader role scored a mean of 2.55 (SD 0.87, SE 0.39, 95% CI 1.47-3.63), whereas the medication nurse role scored a mean of 2.35 (SD 0.86, SE 0.38, 95% CI 1.28-3.42). These results indicate that users perceived the AR system as innovative, contributing to a unique and satisfying experience, but overlapping CIs again reflect limited precision.

Perceived Technology Acceptance

The TAM scores were evaluated across PU and PEU. Combined scores were also calculated to provide an overall measure of acceptance for each role (Figure 6).

Figure 6. Technology acceptance model (TAM) results showing high perceived usefulness (PU) and perceived ease of use (PEU) for both roles, suggesting strong intention to adopt the augmented reality system in clinical training or practice.



- Combined acceptance: The team leader role scored a mean of 5.92 (SD 0.46, SE 0.21, 95% CI 5.35-6.49), whereas the medication nurse role achieved a slightly higher score of 6.02 (SD 0.56, SE 0.25, 95% CI 5.32-6.71). The overlapping CIs suggest broadly comparable acceptance levels.
- PU: The team leader role achieved a mean score of 5.87 (SD 0.74, SE 0.33, 95% CI 4.95-6.78), whereas the medication nurse role scored slightly higher at 5.96 (SD 0.74, SE 0.33, 95% CI 5.05-6.88). While this indicates strong acceptance, CI width reflects uncertainty inherent to the small sample size.
- PEU: The team leader role scored a mean of 6.06 (SD 0.56, SE 0.25, 95% CI 5.37-6.76), whereas the medication nurse role scored similarly at 5.97 (SD 0.60, SE 0.27, 95% CI 5.22-6.72). These point estimates indicate strong acceptance, but CI width again reflects uncertainty.

Discussion

Summary of Main Findings

This study examined the feasibility and perceived usability, user experience, and acceptance of a role-specific AR decision support system designed for resuscitation team leaders and medication nurses. Consistent with the study objectives, clinicians generally perceived the system as usable, intuitive, and acceptable within a high-fidelity simulation context. Perceptions varied by role, reflecting differences in information needs, visual attention demands, and task responsibilities during CA management. These findings suggest that role-tailored AR interfaces are a potential tool for supporting cognitive work in resuscitation settings [15,27], while also underscoring that the present system represents an early-stage, proof-of-concept interface evaluated primarily through subjective measures.

XSL•FO
RenderX

Interpretation of Findings and Relation to Prior Work

Across instruments assessing perceived usability, user experience, and technology acceptance, participants reported favorable impressions of the AR system. These results indicate that clinicians were able to understand and interact with the interface with minimal difficulty and perceived the system as appropriate for use in a simulated resuscitation workflow. Differences in perceived usability and acceptance between team leaders and medication nurses likely reflect role-specific cognitive and visual demands, as team roles in dynamic, safety-critical environments impose distinct situation awareness requirements and attentional burdens depending on task responsibilities and information density [28]. In particular, the team leader interface presented a higher density of information intended to support situational awareness and decision coordination, which may have contributed to comparatively lower—but still positive—perceptions of ease of use.

Participants' responses suggest that the interface aligned with expectations for workflow support in emergent care contexts, where information must be rapidly accessible and interpretable at a glance. These findings are consistent with prior AR and mixed-reality research in clinical and safety-critical domains, which has shown that spatially anchored, role-relevant visual cues can be perceived as supportive when they reduce the need for external references and centralize task-critical information [29,30]. Importantly, these findings reflect perceived support rather than measured improvements in performance, workload, or coordination.

Several participants noted during postsimulation debrief discussions that the AR displays helped them maintain focus on the resuscitation process and reduced reliance on external reference materials. These observations represent subjective reflections elicited during informal debriefing rather than systematically collected performance data and should therefore be interpreted as experiential insights rather than evidence of objective benefit.

Ease of Use, Learnability, and PU

High PEU and learnability indicate that clinicians felt they could quickly become comfortable with the interface, an important consideration for emergency contexts where training time is limited [31]. The visual organization of information, use of glanceable timers, and limited interaction complexity appeared to align with clinicians' expectations for decision support during resuscitation [28,29].

Clinicians also viewed the system content as relevant and supportive of their respective roles, as reflected in ratings related to PU and pragmatic quality. These perceptions are consistent with the underlying design rationale of emphasizing medication-specific information for nurses and algorithmic pathway cues for team leaders. Although prior research suggests that highly usable systems can reduce cognitive load and support more fluid task execution [27], such perceptions should not be interpreted as evidence of improved task performance, guideline adherence, or efficiency. None of these outcomes were directly measured in the current study, and future evaluations must incorporate objective task-level metrics to determine whether perceived utility translates into measurable clinical benefits.

Novelty, Engagement, and Hedonic Experience

Participants rated the AR system highly on hedonic quality dimensions—novelty and stimulation—indicating that the interface was perceived as original, engaging, and distinct from existing tools. These responses reflect perceived innovativeness and experiential engagement rather than satisfaction or effectiveness. Such hedonic responses are encouraging for simulation-based training contexts, where engagement can influence motivation and willingness to adopt new tools [23]. Especially in AR, prior research demonstrated that spatially registered visual cues can increase engagement and perceived control [15,18,19].

At the same time, novelty effects are well documented in evaluations of emerging technologies, particularly during short-term exposure. Perceptions of engagement and stimulation may change with repeated use or prolonged deployment, emphasizing the need for longitudinal studies to assess sustained acceptance and experiential quality over time.

Role-Specific AR Design Implications

A central contribution of this study is the identification of actionable design principles for AR support during CA resuscitation. The iterative prototyping process revealed that AR interfaces should prioritize role-relevant information to minimize unnecessary visual load, use dominant and easily glanceable timers for actionable intervals such as CPR cycles and epinephrine dosing, maintain algorithmic transparency to allow clinicians to view the full pulseless arrest algorithm, and organize spatial layouts clearly by separating medication instructions, procedural steps, and timing cues. These principles provide practical guidance for developers of future AR support tools. While these design choices were intended to support coordination, anticipation, and situational awareness, their operational impact on team performance and guideline adherence remains to be empirically evaluated in future studies. These design considerations align with prior work on situation awareness, cognitive aids, and role-specific information presentation in safety-critical and resuscitation contexts [28-30].

Real-World Implementation Considerations

Although the system achieved promising perception-based results in a controlled simulation environment, translating AR decision support into real clinical workflows presents substantial challenges. Cost, hardware maintenance, device sterilization, and user training remain key considerations for AR deployment in clinical settings [13,30]. Furthermore, seamless interoperability with existing electronic health record systems, secure handling of patient data, and efficient user training are essential for successful integration. Although none of our participants reported discomfort related to the headset bulkiness or fatigue, future iterations should explore lightweight, cost-effective head-mounted devices and web-based synchronization frameworks that ensure data security and workflow continuity. Addressing these implementation barriers will be critical to realizing the clinical impact of AR-based decision support systems. Given that the current evaluation

XSL•FO
RenderX

involved standardized scenarios, conclusions about clinical applicability should be viewed as preliminary.

Limitations and Future Work

While the AR system demonstrated high usability, user experience, and technology acceptance, several limitations should be acknowledged. The most notable limitation is the small sample size (n=10), which restricts statistical generalizability and inferential power. Participants had prior exposure to an early prototype, which may introduce some bias in perceived usability and novelty but also provide more implementation-focused feedback due to their familiarity with the system. Future studies will distinguish between first-time and repeat users to maintain objectivity.

This study was designed primarily to assess initial technical and interaction viability and user experience rather than to test hypotheses or perform comparative statistics. Accordingly, future formal evaluations with larger and more diverse participant samples are planned to validate reproducibility and strengthen external validity. The current evaluation also relied primarily on subjective self-report measures. Incorporating objective performance metrics—such as time to defibrillation, time to epinephrine administration, adherence to CPR cycles, and error frequency—will be crucial in future work. These indicators, combined with physiological or behavioral measures (e.g., eye-tracking, gaze-based workload assessment, or speech-based coordination analysis), can provide richer evidence for the system's real-world effectiveness in improving team performance and reducing cognitive load. Additionally, the study's simulated pediatric CA scenario, while useful for evaluation, may not capture the full range of real-world situations that resuscitation teams might encounter. Expanding the system's evaluation to include a broader range of scenarios could improve its generalizability across diverse clinical environments.

To address these limitations, future research will involve testing the AR system in various CA simulation scenarios to assess its adaptability and reliability before clinical implementation. No major hardware stability issues were observed during testing, and participants, including those wearing corrective glasses, were able to use the device comfortably. Nonetheless, extended use may cause mild visual fatigue or vertigo in a small subset of users, as reported in prior AR literature [27], which warrants monitoring during longer clinical sessions. Plans include conducting an international multisite study with a larger, more diverse participant pool to gain broader insights. This study will also involve incorporating the AR tool into an expanded CPR support system, including additional tools such as a widescreen display for team information visualization, a tablet-based progress monitoring tool providing real-time clinical data, and advanced control interfaces. To gain deeper insights into user performance and behavior, follow-up studies will incorporate objective performance metrics, such as task completion time, gaze tracking, and speech analysis. These metrics will be instrumental in evaluating the system's effectiveness in real-world, high-stakes environments, with the ultimate goal of refining and enhancing its role-specific support functionalities for future clinical use.

Conclusions

This study demonstrates the feasibility and favorable perceived usability, user experience, and acceptance of a role-specific AR decision support system designed for pediatric resuscitation team leaders and medication nurses. Clinicians perceived the system as intuitive, clear, and appropriately tailored to their roles, supporting its potential use in simulation-based training and early-stage clinical exploration. Importantly, the present findings are limited to perception-based outcomes and do not provide evidence of improved performance, workload reduction, or guideline adherence. Rather, this work establishes a foundation for future evaluations that integrate objective measures and assess real-world impact. More broadly, the study illustrates how role-specific AR interfaces can be systematically designed and formatively evaluated as cognitive aids in high-stakes, team-based health care settings.

The innovation of this work lies in its explicit focus on role-specific, in-view AR decision support, which differs from prior studies that primarily evaluated role-agnostic cognitive aids delivered via tablets, posters, or nonadaptive AR displays. By empirically examining clinicians' perceptions across distinct team roles, the study contributes early evidence and practical design guidance for developing role-aware AR interfaces aligned with differing cognitive demands and workflows. In real-world contexts, such role-tailored AR systems may inform the design of next-generation simulation training tools and guide the integration of wearable decision support into clinical resuscitation environments, contingent on future validation using objective performance metrics.

Acknowledgments

All content in this manuscript was conceived and written by the authors. OpenAI ChatGPT-5.2 (2025) was used solely as an editorial aid to improve language quality, such as identifying grammatical issues and suggesting alternative phrasing to enhance clarity and readability of the authors' original text. No artificial intelligence–generated content was incorporated verbatim. All suggestions were critically reviewed, revised as necessary, and approved by the authors, who retain full responsibility for the content and accuracy of the manuscript. Use of generative artificial intelligence was limited to language refinement.

Funding

This study was funded by grant support provided by the Private Foundation of the Geneva University Hospitals, the Alberta Children's Hospital Research Institute, the Alberta Children's Hospital Foundation, the Department of Pediatrics at the University

XSL•FO
RenderX

of Calgary, the University of Calgary Clinical Research Fund Grant, and the NSERC Alliance – Alberta Innovates Advance Program.

Data Availability

The datasets generated and analyzed during this study are not publicly available but are available from the corresponding author upon reasonable request.

Authors' Contributions

The study's initial ideation was led by AC, JNS, and SM, with contributions to the overall study design from RK, YL, JD, DCD, ADM, AR, SKO, FE, and KK. RK was responsible for system development, while RK, AC, JD, YL, DCD, and KK conducted data collection. Data analysis was carried out by RK, HN, and KK. Manuscript writing was undertaken by RK, AC, HN, and KK, with YL, JD, DCD, JNS, SM, ADM, AR, SKO, and FE providing critical review and editing. All authors have reviewed and approved the final version of the manuscript and take responsibility for the integrity and accuracy of the research. The corresponding author affirms that the manuscript is an honest, accurate, and transparent account of the study and confirms that any deviations from the original study plan have been documented.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures illustrating augmented reality (AR) interface design prototypes, system architecture and control interfaces, and usability evaluation results for the role-specific AR decision support system.
[DOCX File, 7546 KB - xr_v3i1e72013_app1.docx]

Multimedia Appendix 2
Grading scale for System Usability Scale scores with corresponding percentile ranges, usability adjectives, and acceptability levels [32].
[XLSX File, 17 KB - xr_v3i1e72013_app2.xlsx]

Multimedia Appendix 3
Interpretation criteria for User Experience Questionnaire scores across different scales [33].
[XLSX File, 17 KB - xr_v3i1e72013_app3.xlsx]

References

1. Meaney PA, Bobrow BJ, Mancini ME, et al. Cardiopulmonary resuscitation quality: [corrected] improving cardiac resuscitation outcomes both inside and outside the hospital: a consensus statement from the American Heart Association. Circulation 2013 Jul 23;128(4):417-435. [doi: 10.1161/CIR.0b013e31829d8654] [Medline: 23801105]
2. Atkins DL, Berger S, Duff JP, et al. Part 11: Pediatric basic life support and cardiopulmonary resuscitation quality: 2015 American Heart Association Guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2015 Nov 3;132(18 Suppl 2):S519-S525. [doi: 10.1161/CIR.0000000000000265] [Medline: 26472999]
3. de Caen AR, Maconochie IK, Aickin R, et al. Part 6: Pediatric basic life support and pediatric advanced life support: 2015 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science With Treatment Recommendations. Circulation 2015 Oct 20;132(16 Suppl 1):S177-S203. [doi: 10.1161/CIR.0000000000000275] [Medline: 26472853]
4. Wolfe HA, Morgan RW, Zhang B, et al. Deviations from AHA guidelines during pediatric cardiopulmonary resuscitation are associated with decreased event survival. Resuscitation 2020 Apr;149:89-99. [doi: 10.1016/j.resuscitation.2020.01.035]
5. Brown LL, Lin Y, Tofil NM, et al. Impact of a CPR feedback device on healthcare provider workload during simulated cardiac arrest. Resuscitation 2018 Sep;130:111-117. [doi: 10.1016/j.resuscitation.2018.06.035] [Medline: 30049656]
6. Tofil NM, Cheng A, Lin Y, et al. Effect of a cardiopulmonary resuscitation coach on workload during pediatric cardiopulmonary arrest: a multicenter, simulation-based study. Pediatr Crit Care Med 2020 May;21(5):e274-e281. [doi: 10.1097/PCC.0000000000002275] [Medline: 32106185]
7. Crabb DB, Hurwitz JE, Reed AC, et al. Innovation in resuscitation: a novel clinical decision display system for advanced cardiac life support. Am J Emerg Med 2021 May;43:217-223. [doi: 10.1016/j.ajem.2020.03.007] [Medline: 32291164]
8. Hejjaji V, Malik AO, Peri-Okonny PA, et al. Mobile app to improve house officers' adherence to advanced cardiac life support guidelines: quality improvement study. JMIR Mhealth Uhealth 2020 May 19;8(5):e15762. [doi: 10.2196/15762] [Medline: 32427115]

XSL•FO
RenderX

9. Márquez-Hernández VV, Gutiérrez-Puertas L, Garrido-Molina JM, García-Viola A, Granados-Gámez G, Aguilera-Manrique G. Using a mobile phone application versus telephone assistance during cardiopulmonary resuscitation: a randomized comparative study. J Emerg Nurs 2020 Jul;46(4):460-467. [doi: 10.1016/j.jen.2020.03.015] [Medline: 32444161]
10. Siebert JN, Lacroix L, Cantais A, Manzano S, Ehrler F. The impact of a tablet app on adherence to American Heart Association guidelines during simulated pediatric cardiopulmonary resuscitation: randomized controlled trial. J Med Internet Res 2020 May 27;22(5):e17792. [doi: 10.2196/17792] [Medline: 32292179]
11. Corazza F, Snijders D, Arpone M, et al. Development and usability of a novel interactive tablet app (PediAppRREST) to support the management of pediatric cardiac arrest: pilot high-fidelity simulation-based study. JMIR Mhealth Uhealth 2020 Oct 1;8(10):e19070. [doi: 10.2196/19070] [Medline: 32788142]
12. Corazza F, Fiorese E, Arpone M, et al. The impact of cognitive aids on resuscitation performance in in-hospital cardiac arrest scenarios: a systematic review and meta-analysis. Intern Emerg Med 2022 Oct;17(7):2143-2158. [doi: 10.1007/s11739-022-03041-6] [Medline: 36031672]
13. Iqbal AI, Aamir A, Hammad A, et al. Immersive technologies in healthcare: an in-depth exploration of virtual reality and augmented reality in enhancing patient care, medical education, and training paradigms. J Prim Care Community Health 2024;15:21501319241293311. [doi: 10.1177/21501319241293311] [Medline: 39439304]
14. Harari RE, Schulwolf SL, Borges P, et al. Applications of augmented reality for prehospital emergency care: systematic review of randomized controlled trials. JMIR XR Spatial Comput 2025;2:e66222-e66222. [doi: 10.2196/66222]
15. Siebert JN, Ehrler F, Gervaix A, et al. Adherence to AHA guidelines when adapted for augmented reality glasses for assisted pediatric cardiopulmonary resuscitation: a randomized controlled trial. J Med Internet Res 2017 May 29;19(5):e183. [doi: 10.2196/jmir.7379] [Medline: 28554878]
16. Li X, Yin X, Huang G, Wang X. Effectiveness of extended reality technologies in cardiopulmonary resuscitation training: a Bayesian network meta-analysis. BMC Emerg Med 2025 Jun 7;25(1):94. [doi: 10.1186/s12873-025-01256-2] [Medline: 40483416]
17. Dubreucq E, Barlocco De La Vega S, Bouaoud J, Philippon AL, Thiebaud PC. Impact of virtual, augmented or mixed reality in basic life support training: a scoping review. Clin Simul Nurs 2025 Feb;99:101672. [doi: 10.1016/j.ecns.2024.101672]
18. Ehrler F, Del Zotto M, Rouyer F, Weinhold T, Lovis C, Siebert J. Design of InterFACE: a tool to improve collaborative work and decision making during rescucitation. Stud Health Technol Inform 2018;255:117-121. [Medline: 30306919]
19. Tsang KD, Ottow MK, van Heijst AFJ, Antonius TAJ. Electronic decision support in the delivery room using augmented reality to improve newborn life support guideline adherence: a randomized controlled pilot study. Simul Healthc 2022 Oct 1;17(5):293-298. [doi: 10.1097/SIH.0000000000000631] [Medline: 35102128]
20. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry, 1st edition: CRC Press; 1996, Vol. 194:189-194. [doi: 10.1201/9781498710411]
21. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
22. Laugwitz B, Held T, Schrepp M. Construction and evaluation of a user experience questionnaire. In: Holzinger A, editor. HCI and Usability for Education and Work USAB 2008 Lecture Notes in Computer Science: Springer; 2008, Vol. 5298:63-76. [doi: 10.1007/978-3-540-89350-9_6]
23. Schrepp M, Hinderks A, Thomaschewski J. Construction of a benchmark for the User Experience Questionnaire (UEQ). IJIMAI 2017;4(4):40. [doi: 10.9781/ijimai.2017.445]
24. Schrepp M, Thomaschewski J. Handbook for the modular extension of the User Experience Questionnaire version 5. UEQ+. URL: https://ueqplus.ueq-research.org/Material/UEQ_Handbook_V5.pdf [accessed 2026-02-05]
25. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
26. Sauro J. 10 things to know about the technology acceptance model. Measuring U. URL: https://measuringu.com/tam [accessed 2025-12-14]
27. Jeffri NFS, Awang Rambli DR. A review of augmented reality systems and their effects on mental workload and task performance. Heliyon 2021 Mar;7(3):e06277. [doi: 10.1016/j.heliyon.2021.e06277] [Medline: 33748449]
28. Endsley MR. Toward a theory of situation awareness in dynamic systems. Hum Factors 1995 Mar;37(1):32-64. [doi: 10.1518/001872095779049543]
29. 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] [Medline: 22833264]
30. Sielhorst T, Feuerstein M, Navab N. Advanced medical displays: a literature review of augmented reality. J Display Technol 2008;4(4):451-467. [doi: 10.1109/JDT.2008.2001575]
31. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 7;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
32. 5 ways to interpret a SUS score. Measuring U. URL: https://measuringu.com/interpret-sus-score/ [accessed 2026-02-03]
33. UEQ - User Experience Questionnaire. URL: https://www.ueq-online.org/ [accessed 2026-02-03]

XSL•FO
RenderX

Abbreviations

ACH: Alberta Children's Hospital
AR: augmented reality
CA: cardiac arrest
CPR: cardiopulmonary resuscitation
PEA: pulseless electrical activity
PEU: perceived ease of use
PU: perceived usefulness
SUS: System Usability Scale
TAM: technology acceptance model
UEQ: User Experience Questionnaire
UI: user interface

Edited by S Brini, T Leung; submitted 25.Mar.2025; peer-reviewed by A Adekola, K Kleinman, L Mommers, R Marcilly; revised version received 11.Jan.2026; accepted 13.Jan.2026; published 07.Apr.2026.

Please cite as:
Kang R, Cheng A, Lin Y, Nam H, Davidson J, Duncan DC, Siebert JN, Manzano S, De Masi A, Rajic A, Olanka SK, Ehrler F, Kim K Perceived Usability, User Experience, and Technology Acceptance of Role-Specific Augmented Reality Decision Support Tools for Cardiac Arrest Resuscitation: Prospective Observational Pilot Study
JMIR XR Spatial Comput 2026;3:e72013
URL: https://xr.jmir.org/2026/1/e72013
doi:10.2196/72013

© Ryan Kang, Adam Cheng, Yiqun Lin, Hyeongil Nam, Jennifer Davidson, Donovan Curtis Duncan, Johan N Siebert, Sergio Manzano, Alexandre De Masi, Ana Rajic, Sharleen Kayne Olanka, Frederic Ehrler, Kangsoo Kim. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 7.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Persuasive Gamified Virtual Reality Experience to Enhance Engagement and Focus in Young Adults With Mild Anxiety Symptoms: Randomized Pilot Experimental Study

Mahlet Misrak Argaw[1], MSc; Nuru Jingili[1], PhD; Solomon Sunday Oyelere[1,2], PhD; Markus B T Nyström[3]

[1]Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Forskargatan 1, Skellefteå, Sweden

[2]Department of Computer Science, University of Exeter, Exeter, United Kingdom

[3]Department of Health, Education and Technology Division, Luleå University of Technology, Luleå, Sweden, Luleå, Sweden

Corresponding Author:
Solomon Sunday Oyelere, PhD
Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Forskargatan 1, Skellefteå, Sweden

Abstract

Background: Anxiety-related symptoms are prevalent and can negatively affect concentration, motivation, and overall well-being. Traditional treatments such as cognitive behavioral therapy and medication work well for clinical anxiety disorders. However, individuals with anxiety often struggle with access, adherence, and staying engaged in treatment. Emerging technologies such as virtual reality (VR) and gamification offer new opportunities to enhance user engagement and motivational processes within digital mental health applications.

Objective: This study introduces *Cleanify*, a gamified VR cleaning simulation designed using the Octalysis framework and the persuasive system design model. The objective was to evaluate whether gamification elements improve user engagement, focus, and satisfaction compared to a nongamified version among individuals experiencing anxiety symptoms. We hypothesized that the gamified version would outperform the nongamified version in enhancing user engagement, immersion, and overall user experience.

Methods: A pilot experimental study was conducted with 50 participants aged 18 to 39 years recruited from the general population in northern Sweden. Participants were randomly assigned to either a gamified or nongamified version of the *Cleanify* VR application and completed a single 15-minute VR session using the Oculus Quest headset. Baseline anxiety symptoms were assessed using the Generalized Anxiety Disorder–7 scale for descriptive purposes only. Postintervention outcomes included focus and immersion measured using the Flow State Scale and user experience measured using the short version of the User Experience Questionnaire. Group differences were analyzed using 2-tailed independent-sample *t* tests.

Results: Participants using the gamified VR version demonstrated higher engagement and immersion than those using the nongamified version. The gamified group reached higher in-game levels overall, with a greater proportion of participants reaching level 3 (17/25, 68% vs 8/25, 32%), and reported higher recommendation scores (mean 4.20, SD 0.76 vs 3.36, SD 0.86). Significant group differences were observed for overall flow (t_{48}=3.87; P<.001), fluency (t_{48}=4.36; P<.001), and absorption (t_{48}=2.80; P=.008). User Experience Questionnaire results indicated higher pragmatic quality, hedonic quality, and overall user experience in the gamified condition.

Conclusions: Integrating gamification into a VR environment significantly enhanced user engagement, focus, and immersion in this pilot sample. These findings provide preliminary evidence that gamified VR design elements can positively influence user experience outcomes. Further research incorporating longitudinal designs and clinical outcome measures is needed to determine potential relevance.

(JMIR XR Spatial Comput 2026;3:e66713) doi:10.2196/66713

KEYWORDS

human-computer interaction; user-centered design; virtual reality; VR; game; gamification; anxiety; disorder; adult; young adults; generalized anxiety disorder; GAD; mental health; randomized; controlled trial; randomized controlled trial; RCT; questionnaire; persuasive technology

XSL•FO
RenderX

Introduction

Mental health conditions such as anxiety and depression affect nearly 1 in 8 people worldwide, with prevalence rising since the COVID-19 pandemic [1,2]. Anxiety-related symptoms, including excessive worry, restlessness, difficulty concentrating, and physical tension, are commonly reported across both clinical and nonclinical populations [3,4]. While treatments such as cognitive behavioral therapy and pharmacological interventions are effective for diagnosed anxiety disorders, many individuals experiencing mild or subclinical anxiety symptoms remain untreated due to stigma, limited access to care, and challenges with treatment adherence [5-7]. Research further indicates that anxiety is associated with difficulties in attentional control and emotional regulation, highlighting the importance of interventions that support focus, engagement, and adaptive coping strategies [8,9].

Virtual reality (VR) has emerged as a promising tool for mental health interventions, offering immersive, controlled environments for practicing mindfulness and exposure therapy [4,10,11]. Compared to traditional approaches, VR can increase motivation and simulate real-world contexts safely [12,13]. Applications such as *SnowWorld* have demonstrated VR's potential in stress and pain reduction, and VR-based exposure therapies have been successfully applied to disorders such as phobias and posttraumatic stress disorder [14-16].

Alongside VR, gamification—the integration of gamelike elements such as points, levels, and rewards—has been shown to improve motivation and sustain engagement in therapeutic contexts [17-19]. Recent studies combining gamification with VR have produced encouraging results for social anxiety, phobias, and public speaking interventions, reporting greater immersion and positive user experiences [20-22].

Building on this work, this study introduces *Cleanify*, a VR house cleaning simulation developed using principles of persuasive system design and gamification. Cleaning activities were chosen because they are familiar, accessible, and low barrier, whereas research indicates that they can reduce stress and restore a sense of control [23,24]. By embedding mindfulness into structured, repetitive cleaning tasks, the intervention offers a practical way to cultivate focused attention.

This early-stage pilot study compared gamified and nongamified versions of *Cleanify* to evaluate their impact on user engagement, immersion, and satisfaction. The research question was as follows: how do users perceive focus, immersion, and engagement in gamified vs nongamified versions of VR mindfulness applications? The findings aim to inform the design and optimization of future VR interventions targeting individuals experiencing mild anxiety symptoms, particularly by enhancing user motivation and sustained participation.

Methods

Theoretical Framework

A growing body of research highlights the connection between mental health and one's living environment. Cluttered spaces are associated with stress, confusion, and negative affect, whereas tidy spaces can promote clarity and well-being [25-27]. Cleaning and organizing activities are thought to reduce stress by restoring a sense of control and alleviating the cognitive burden of unfinished tasks [28]. Moreover, repetitive household activities such as dishwashing have been linked to increased mindfulness, reduced anxiety, and improved mental inspiration [29].

These findings suggest that cleaning tasks provide a natural context for mindfulness: they are structured, goal-oriented, and yield immediate feedback, which can foster a sense of accomplishment and calm [30]. Importantly, such tasks are low barrier and familiar, making them accessible across diverse populations.

However, the effectiveness of these interventions depends heavily on user engagement. Evidence shows that greater engagement in activities is linked to stronger reductions in anxiety and depression symptoms, with a dose-response relationship observed in cognitive behavioral therapy [31]. Therefore, gamification and persuasive design can play a critical role in sustaining motivation and maximizing the potential benefits of mindfulness-based cleaning activities delivered through VR.

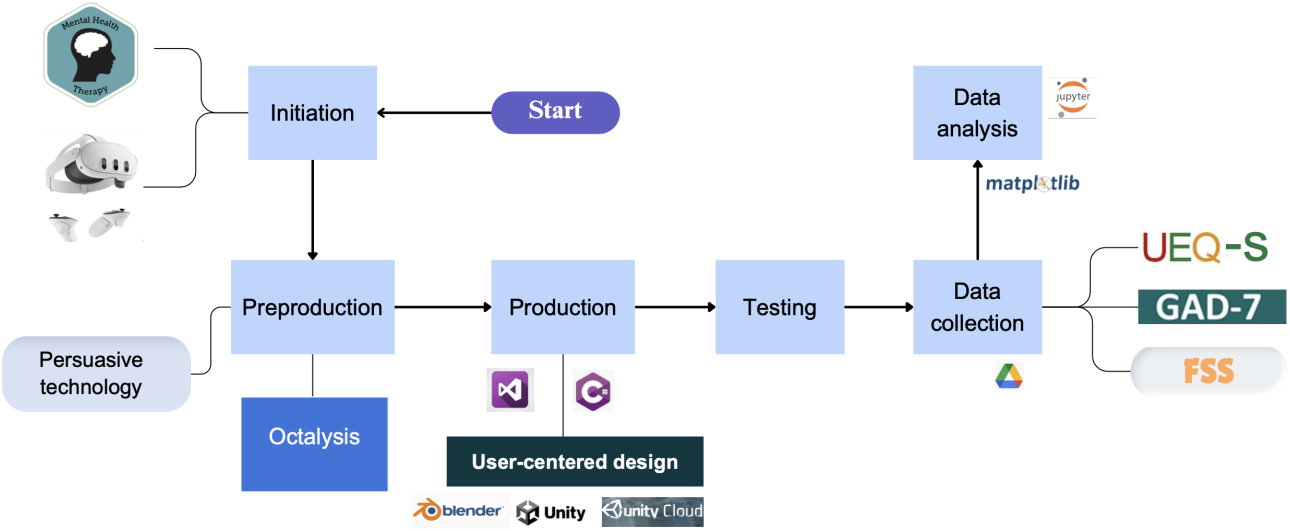
Game Design

Cleanify was developed as a gamified VR experience designed for the Oculus Quest headset (Reality Labs). The application features a 3-level house cleaning simulation structured to progressively engage players in immersive, goal-oriented tasks. The design was guided by the Octalysis framework, which incorporates 8 core motivational drivers to enhance user engagement and sustain motivation (Multimedia Appendix 1). In parallel, the persuasive system design model was applied to integrate persuasive elements that ethically support behavior change and improve user focus. This was particularly relevant given the attentional challenges often experienced by individuals with anxiety [32,33].

The system development process followed 5 phases (Figure 1): initiation (concept development), preproduction (design planning), production (game development with user-centered design), testing (pilot- and usability testing using the Generalized Anxiety Disorder–7 [GAD-7] scale, the short version of the User Experience Questionnaire [UEQ-S], and the Flow State Scale [FSS] tools), and data analysis (evaluation of outcomes).

XSL•FO
RenderX

Figure 1. Flow diagram illustrating the system development and randomized pilot experimental study process of the *Cleanify* virtual reality intervention conducted in Skellefteå, Sweden, in 2024, including initiation, preproduction, production, testing, data collection (Generalized Anxiety Disorder–7 [GAD-7], Flow State Scale [FSS], short version of the User Experience Questionnaire [UEQ-S]), and statistical analysis among young adults with mild anxiety symptoms.



Game Elements and Their Potential Effects on User Engagement and Well-Being

Overview

Cleanify integrates several game elements designed to enhance user focus, engagement, and immersive experience within a virtual environment. Rather than functioning as a clinical treatment, these elements aim to support attentional engagement and promote a calming, structured activity that may be beneficial for individuals experiencing mild anxiety symptoms. An overview of the game's components and their intended experiential functions is provided below. Screenshots of selected components are included in Multimedia Appendices 2 and 3.

Interactive Elements

The *Cleanify* application gives users access to a comfortable VR home with various interactive household necessities. Users can interact with these items in a natural way, which helps the virtual home feel more realistic. Instructions in the game offer different interactive tasks, which can help distract players from their worries and provide a sense of purpose [34]. In addition, gentle reminders to keep calm and relax are placed in different room areas. This helps maintain focus on the present and mindfulness [35].

Physical Movement

In the VR game, players must move around a room to clean up dirt stains, place trash in the trash can, and arrange objects. This engages players in physical activity that releases endorphins, leading to an improvement in anxiety and depression symptoms [36].

Organizing a Space

Cleanify enables players to experience a sense of relaxation and satisfaction through VR immersive technology. Organizing the space into a visually appealing environment increases feelings of well-being [29]. This feeling is further enhanced by different sound and visual effects, providing positive reinforcement that can improve mood and self-esteem [37]. Seeing dirt being removed or trash disappear followed by immediate positive feedback shows players the effectiveness of their efforts, which enhances motivation and satisfaction [38].

Level Progression

The VR game offers structured progression through levels with distinct objectives and benchmarks. This gives the entire gaming experience a sense of regularity and order [39]. Establishing these structured routines may help reduce perceived stress symptoms [40]. Moreover, completing the tasks under each level gives players a sense of achievement, which reduces stress [41].

Controlled Environment

Cleanify empowers players, providing a sense of control over their surroundings [42]. Additionally, the VR game's second and third levels include time-limited challenges. This helps players practice coping with pressure in a safe environment [43].

Game Scoring

Depending on the player's performance, the VR application offers several scoring systems, including stars, badges, and a scoreboard for high scores. These components increase focus on and engagement in tasks [44]. Displaying the top scorers enables players to view their rankings and compare their progress with that of others [45].

Participant Recruitment and Research Process

Participants were recruited between April 2024 and May 2024 from the general population in the municipality of Skellefteå in northern Sweden through multiple channels, including social media announcements, local community centers, and student Discord groups. Interested individuals completed an online screening survey that included demographic questions and the GAD-7 scale to assess anxiety symptom severity. Exclusion criteria were severe psychiatric disorders, epilepsy, or a history of severe motion sickness, ensuring suitability for VR

XSL•FO
RenderX

participation. A total of 54 individuals were recruited; 4 (7.4%) were excluded for not meeting the criteria, resulting in a final sample of 50 (92.6%) participants who were randomized. Studies suggest that 10 to 30 participants per group is usually acceptable for a pilot study [46], and therefore, our sample size aligns with established guidelines.

Prior to the main intervention, all participants underwent a structured onboarding and training session. This included an orientation briefing that explained the study purpose, VR procedures, and safety precautions. Participants were shown a short tutorial video and provided with guided, hands-on practice using the Oculus Quest headset and controllers. It was ensured that each participant demonstrated adequate understanding of headset operation, navigation, and interaction within the VR environment before beginning the intervention. This training minimized variability in participant readiness and helped establish comfort with the VR system.

Instruments

Several standardized instruments were used to assess participant characteristics and user experience outcomes. The GAD-7, a validated self-report measure of anxiety symptom severity, was administered at baseline to characterize participants' anxiety levels prior to the VR session [47]. The UEQ-S was used to evaluate participants' subjective experiences, focusing on usability, immersion, and satisfaction [48,49]. The short version was selected to efficiently capture key user experience dimensions while minimizing participant fatigue. The FSS was used to measure participants' perceived focus and immersion during the VR session [50]. Flow describes a psychological state in which perceived challenges are balanced with individual skills, resulting in deep concentration and intrinsic engagement. It is widely used to evaluate immersive experiences in domains such as gaming, training, and digital applications [51].

Experiment Design

The study began with an overview of its purpose, after which participants provided informed consent. Following demographic and eligibility screening, participants completed the GAD-7 assessment [47]. Screening ensured that participants met the inclusion criteria and had no contraindications for VR use, such as epilepsy or severe motion sickness.

Participants were randomly assigned to either the gamified or nongamified group using a simple randomization method with a 1:1 allocation ratio. No blocking or stratification was applied. Group allocation was revealed to participants only at the time of the intervention. Due to the visible differences between versions, participants were not blinded to group allocation. The researcher overseeing data collection was also aware of group assignments; however, data analysis was conducted without access to condition labels until completion of the primary analyses.

Each participant then engaged in a single 15-minute VR session. No external prompts or reminders were used to encourage participant engagement; all guidance and prompts were embedded within the VR application in the form of task instructions and in-environment reminders.

Questionnaire Design

A Google Forms questionnaire structured into 5 sections was used for data collection. The first section provided an explanation of the study purpose and procedures and included informed consent information, emphasizing voluntary participation, data confidentiality, and the right to withdraw at any time. The second section collected demographic information, including age, gender, and occupation.

The third section consisted of the GAD-7 scale, which assesses the frequency of anxiety-related symptoms over the preceding 2 weeks [52]. The fourth section included 8 items from the UEQ-S measured on a 7-point Likert scale, assessing usability, immersion, and overall user experience [48]. The fifth section comprised the FSS, consisting of 13 items measured on a 7-point Likert scale to assess focus and immersion during the VR experience [50].

In addition to these standardized instruments, 2 engagement-related measures were included. Level reached was used as an objective indicator of in-game engagement and referred to the highest level (levels 1 - 3) completed by each participant during the 15-minute VR session. The *Cleanify* VR application consists of 3 sequential levels of increasing task complexity. Recommendation score was used as a subjective measure of user satisfaction and was collected by asking participants to rate their likelihood of recommending the *Cleanify* VR application to others on a 5-point Likert scale (1="strongly disagree"; 5="strongly agree").

The complete set of questionnaire items used in this study, including the GAD-7, FSS, and UEQ-S items, is provided in Multimedia Appendix 4.

Data Analysis

All data were collected using Google Forms. All responses were collected anonymously, and no personally identifiable information was recorded. Access to the data was restricted to the research team, and data handling procedures complied with institutional data protection guidelines. Participants completed the questionnaires both before and after the 15-minute VR session by scanning a QR code with their smartphones and completing the forms on-site. The data were subsequently exported to Google Sheets for analysis. Analyses were performed using Python (version 3.0; Python Software Foundation) with standard libraries (pandas and NumPy). Descriptive statistics were calculated to summarize participant demographics, baseline anxiety levels (GAD-7 scores), and engagement metrics such as level reached and session duration. Means, SDs, and percentages were reported where appropriate.

Inferential statistical tests were conducted to compare outcomes between the 2 conditions (gamified vs nongamified). Two-tailed independent-sample *t* tests were used to assess differences in continuous variables, including flow, worry, fluency, and absorption scores from the FSS as well as user experience scores from the UEQ-S. Statistical significance was set at an α level of .05.

Assumptions of normality were considered when selecting statistical tests. Data distributions were assessed using

XSL•FO
RenderX

descriptive inspection and skewness values. Given the exploratory nature of this pilot study and the approximately balanced group sizes, independent-sample *t* tests were applied as they are robust to moderate deviations from normality.

Baseline GAD-7 scores were used for descriptive purposes only to characterize anxiety symptom severity within the sample, and no inferential statistical tests were applied to these variables. For ordinal variables such as level reached, descriptive comparisons were used instead of parametric statistical testing.

Ethical Considerations

This study was conducted in accordance with the World Medical Association's International Code of Medical Ethics (Declaration of Helsinki). The well-being of participants was prioritized before, during, and after the intervention. Participants were provided with clear information about the study's purpose and procedures and any potential risks, enabling them to make informed decisions about their involvement. Informed consent was obtained directly from all young adults prior to data collection, and they were informed of their right to withdraw at any time without consequence. Participants did not receive any compensation for their involvement in this study.

No personally identifiable information was collected; only limited demographic details (age, gender, occupation, and prior VR experience) were recorded. All responses were collected anonymously via Google Forms, stored securely, and accessible only to the research team, ensuring participants' privacy and confidentiality.

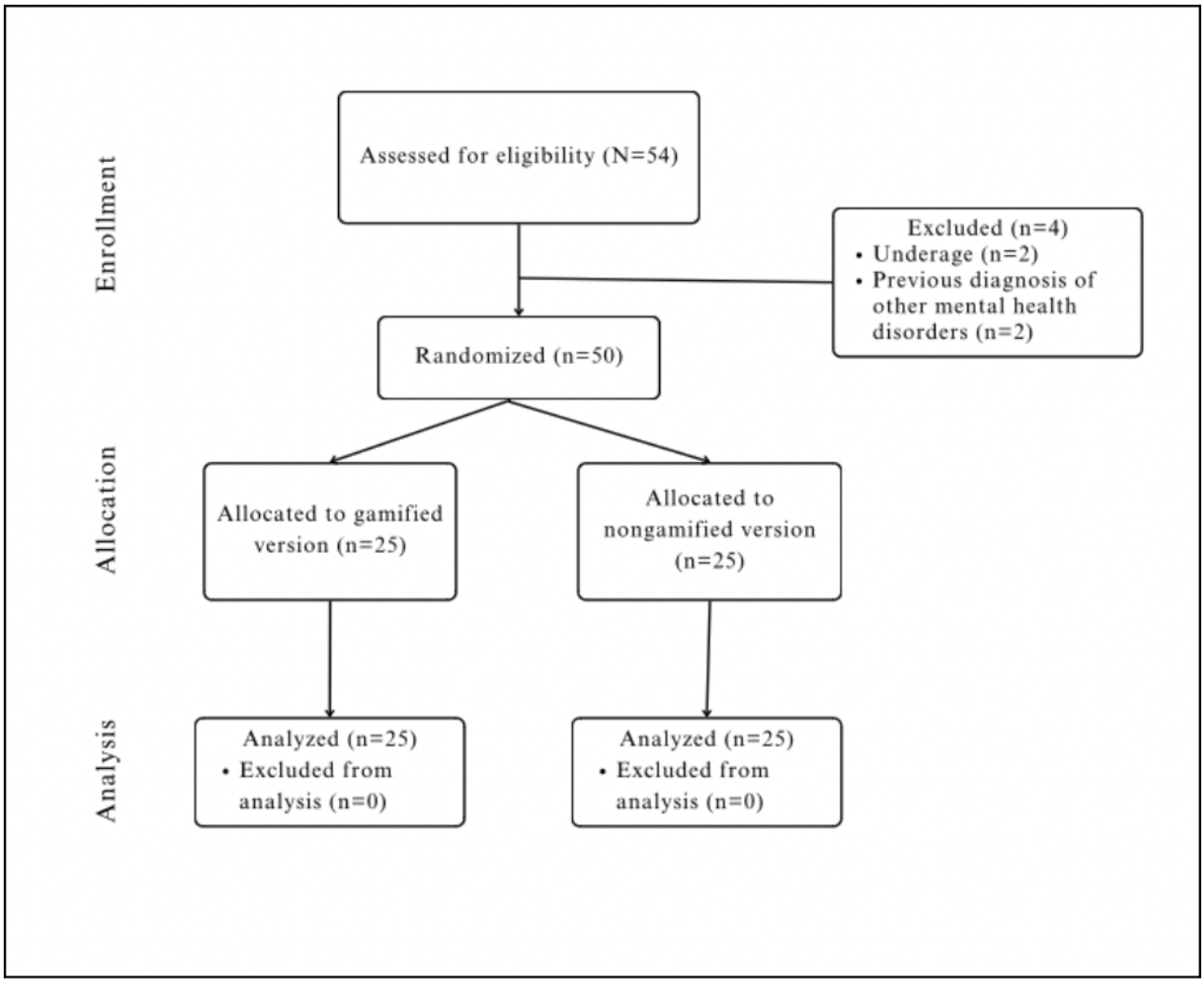
Given the nonclinical nature of the study, the use of anonymous data, and the minimal risk to participants, formal ethics approval was not obtained. However, the study adhered to standard ethical principles for human subject research.

Results

Participant Flow

Figure 2 illustrates the flow of participants through the study, including recruitment, exclusions, randomization, and final allocation to the gamified and nongamified conditions.

Figure 2. Participant flow diagram of the randomized pilot experimental study conducted in Skellefteå, Sweden, in 2024 showing screening, exclusions, randomization, and allocation to gamified and nongamified virtual reality intervention groups among young adults aged 18 to 39 years with mild anxiety symptoms.



XSL•FO
RenderX

Participant Demography

Table 1 summarizes the demographic characteristics of the study participants. A total of 50 individuals took part in the study, representing a diverse sample in terms of age, gender, occupation, and prior experience with VR.

Table . Demographic characteristics of participants enrolled in a randomized pilot experimental study conducted in Skellefteå, Sweden, in 2024 evaluating gamified vs nongamified virtual reality (VR) interventions (N=50).

Variable	Participants, n (%)
Gender	
Man	29 (58)
Woman	21 (42)
Nonbinary	0 (0)
Age (y)	
18-25	12 (24)
25-39	38 (76)
Occupation	
Student	28 (56)
Professional	22 (44)
Prior experience with VR	
Yes	30 (60)
No	20 (40)

Baseline Characteristics

Table 2 provides a comparison of baseline demographic and anxiety-related characteristics between the gamified and nongamified groups. Variables include age, gender, occupation, prior VR experience, and mean baseline GAD-7 scores. Independent-sample *t* tests and chi-square analyses were conducted to assess potential group differences.

Table . Baseline demographic and clinical characteristics of participants randomized to the gamified and nongamified groups in a pilot experimental study conducted in Skellefteå, Sweden, in 2024.

Variable	Gamified (n=25)	Nongamified (n=25)	Test statistic		*P* value
			t test (*df*)	Chi-square (*df*)	
Age (y), mean (SD)	29.1 (4.9)	29.0 (5.0)	0.09 (48)	—[a]	.93
Gender, n (%)			—	0.3 (1)	.56
Man	17 (68)	14 (56)			
Woman	8 (32)	11 (44)			
Occupation, n (%)			—	0.7 (1)	.39
Student	16 (64)	12 (48)			
Professional	9 (36)	13 (52)			
Prior VR[b] experience, n (%)	14 (56)	16 (64)	—	0.08 (1)	.77
GAD-7[c] score (0-21), mean (SD)	6.8 (5.9)	6.9 (5.3)	–0.10 (48)	—	.92

[a]Not applicable.

[b]VR: virtual reality.

[c]GAD-7: Generalized Anxiety Disorder–7.

The results indicated no statistically significant differences between groups on these baseline variables, supporting comparability prior to the intervention.

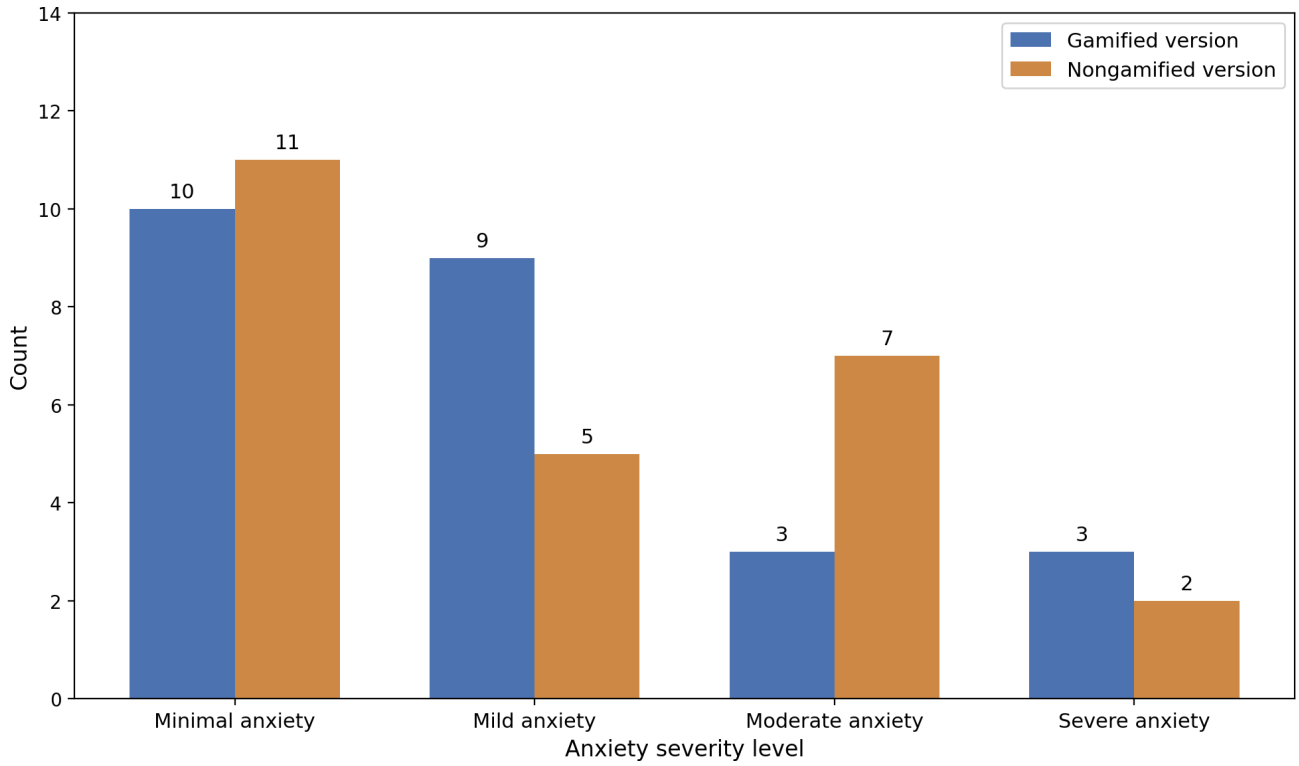
GAD-7 Score

The GAD-7 results were used for descriptive analysis to provide an overview of participants' anxiety levels before engaging with the *Cleanify* VR application. The data collected allowed for the assessment of baseline anxiety symptoms within the participant

XSL•FO
RenderX

group. This offered insights into the overall mental health status of the individuals involved. However, it is important to note that the study did not include a postintervention assessment, so the GAD-7 results primarily describe the participants' anxiety profiles rather than measuring changes resulting from the intervention. Figure 3 shows the distribution of anxiety levels between the 2 groups. It highlights the variations in minimal, mild, moderate, and severe anxiety levels among the participants. This visual representation underscores the diversity of anxiety experiences within the study population.

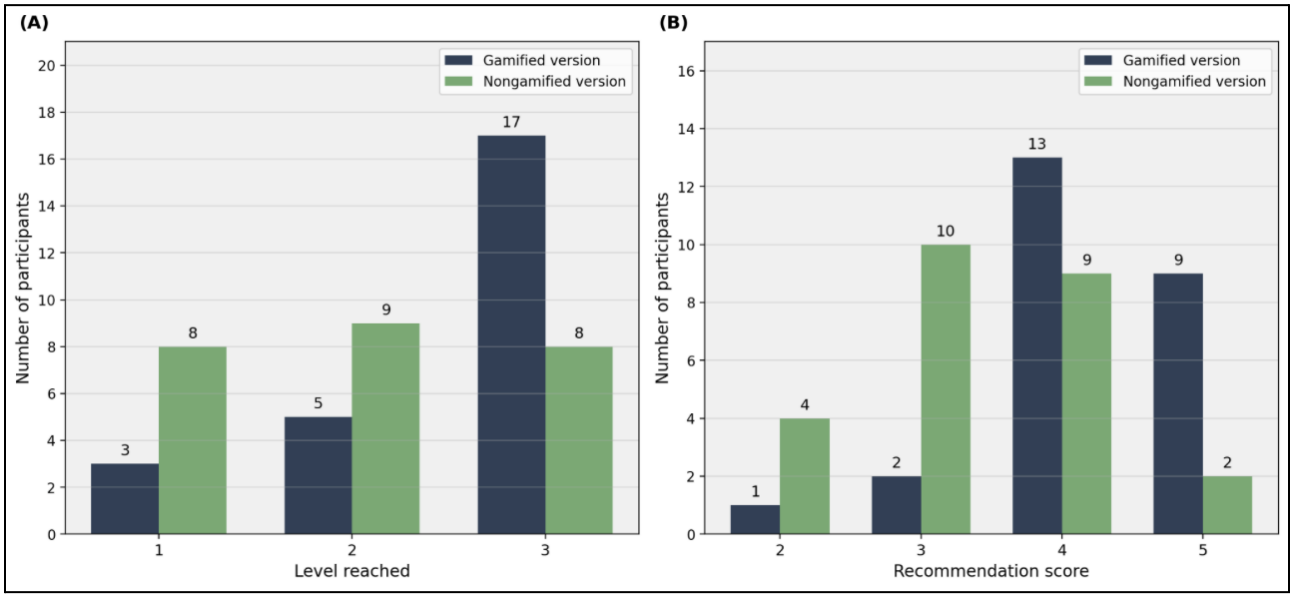
Figure 3. Distribution of baseline anxiety severity levels (minimal, mild, moderate, and severe) measured using the Generalized Anxiety Disorder–7 scale among young adults (N=50) in Skellefteå, Sweden, in 2024 randomized to gamified and nongamified virtual reality conditions.



User Experience

During the intervention, the level reached by participants was recorded as an indicator of engagement within the VR experience. As illustrated in Figure 4, differences in level progression between the 2 groups highlight variations in user engagement. A higher proportion of participants in the gamified group reached level 3 (17/25, 68.0%) than in the nongamified group (8/25, 32.0%), indicating greater progression and interaction with the game. This pattern suggests that gamification elements supported increased engagement during the session.

Figure 4. Comparison of engagement outcomes between the gamified and nongamified virtual reality groups in a pilot experimental study conducted in Skellefteå, Sweden, in 2024: (A) the distribution of the highest levels reached (levels 1 - 3) and (B) recommendation scores (5-point Likert scale).



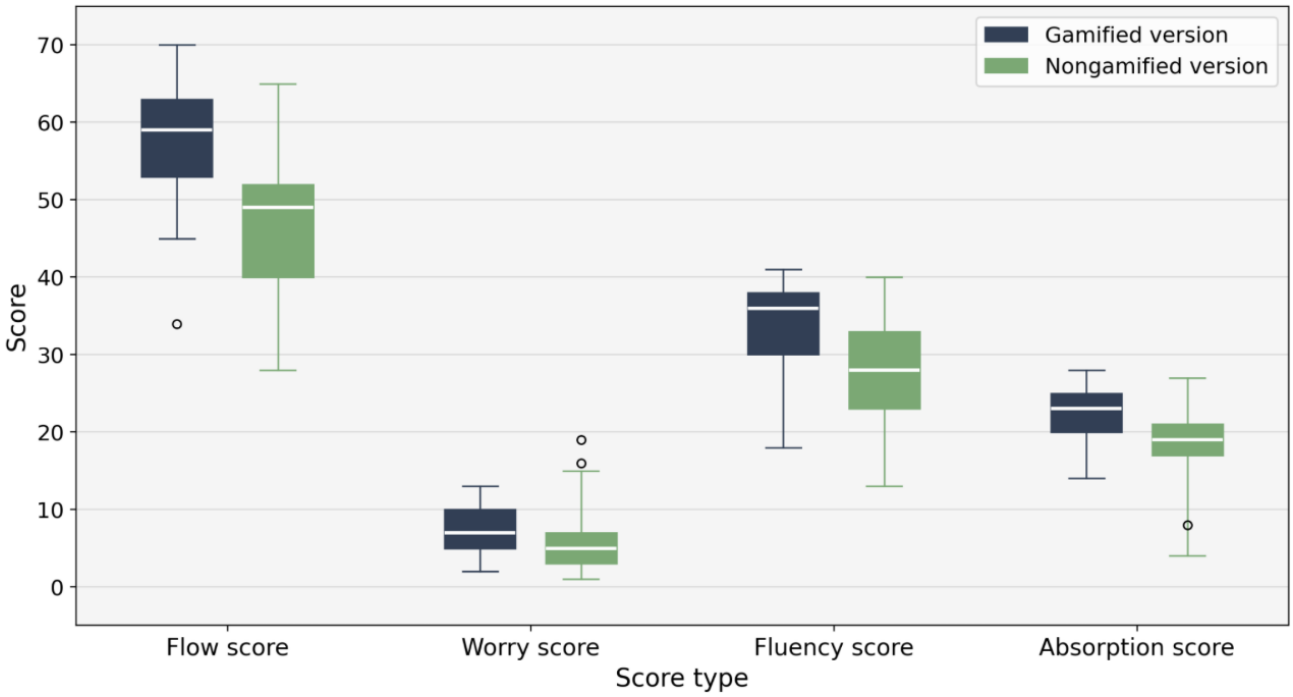
XSL•FO
RenderX

After performing the intervention, participants were asked whether they would recommend the application to others as an indicator of user experience. Participants in the gamified group reported a higher average recommendation score (4.20, SD 0.76) than the nongamified group (3.36, SD 0.86), with 36.0% (9/25)of the participants in the gamified group giving a top score of 5 compared to 8.0% (2/25) (n/N) in the nongamified group. These findings suggest that the gamified elements of *Cleanify* enhanced participants' enjoyment of the application, increasing their likelihood of recommending it to others.

User Focus and Immersion Analysis

The FSS scores showed significant differences between the 2 groups (Figure 5; α=.05; flow score: t_{48}=3.87 and P<.001; worry score: t_{48}=1.66 and P=.10; fluency score: t_{48}=4.36 and P<.001; absorption score: t_{48}=2.80 and P=.008).

Figure 5. Comparison of Flow State Scale subscale scores (flow, worry, fluency, and absorption) between the gamified and nongamified virtual reality (VR) groups in a randomized pilot experimental study conducted in Skellefteå, Sweden, in 2024. Box plots display the distribution of postintervention scores following a single 15-minute VR session.

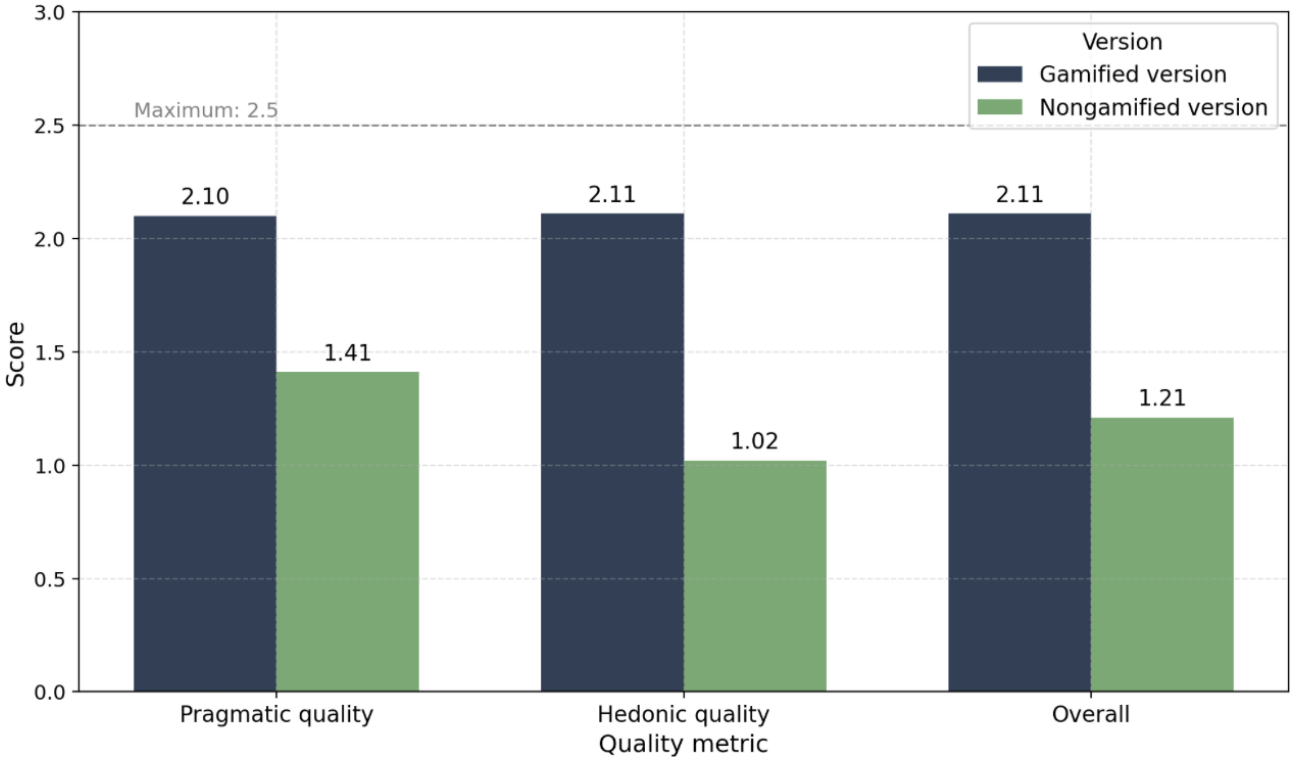


User Engagement and System Usability Analysis

UEQ-S data showed higher scores in the gamified version across pragmatic quality (mean 2.10, sd 0.92 vs mean 1.41 , sd 0.98), hedonic quality (mean 2.12 , sd 1.06 vs mean 1.01, sd 1.07), and overall quality (mean 2.11, sd 0.92 vs mean 1.21, sd 0.92 ; Figure 6). These differences indicate improved user experience in the gamified condition. However, as statistical significance testing was not conducted for UEQ-S subscales, these findings should be interpreted as descriptive comparisons.

XSL•FO
RenderX

Figure 6. Scores on the short version of the User Experience Questionnaire comparing pragmatic quality, hedonic quality, and overall user experience between gamified and nongamified versions of a virtual reality application in a randomized pilot study conducted in northern Sweden in 2024.



Discussion

Principal Findings

This study explored how gamification influences user engagement, focus, and immersion in a VR application designed for individuals with generalized anxiety symptoms. Results from the FSS indicated significant differences between groups: participants using the gamified version reported higher levels of flow, immersion, and absorption than those using the nongamified version. These findings suggest that gamified elements facilitated deeper engagement, with users becoming more absorbed in the activity and experiencing smoother, more cohesive interactions. Importantly, worry scores were low across both conditions, indicating that participants generally experienced the task as calming and focused.

User experience ratings (UEQ-S) were also higher in the gamified version across both pragmatic and hedonic dimensions, underscoring that gamification improved usability, satisfaction, and enjoyment. Collectively, these findings highlight that gamification enhanced the subjective quality of the VR experience.

Nevertheless, this study was not designed to evaluate therapeutic efficacy. While gamification clearly increases focus and engagement, the absence of postintervention measures of anxiety, stress, or mindfulness prevents conclusions about clinical benefits. Baseline equivalence across demographic characteristics and GAD-7 scores helps rule out preexisting differences as drivers of the results, but therapeutic impact remains untested.

Comparison to Prior Work

These results align with those of prior literature demonstrating that gamification enhances motivation, engagement, and user experience in nongame contexts [16,53]. The higher flow, fluency, and absorption scores in the gamified condition echo findings that game elements—such as challenges, feedback, and rewards—can deepen user involvement and sustain attention [54]. Similarly, the higher pragmatic and hedonic scores mirror earlier work linking gamification to both improved functionality and enjoyment [55,56].

Together, these findings extend existing knowledge by showing that gamified VR environments can support more engaging and satisfying experiences for individuals with generalized anxiety symptoms. They support the proposition that integrating game design elements into VR tools creates more compelling environments, although clinical benefits require further investigation.

Limitations and Future Work

Several limitations temper the interpretation of these findings. First, this study was exploratory and focused primarily on engagement, immersion, and usability rather than clinical outcomes. Therefore, the findings should not be interpreted as evidence of therapeutic efficacy.

Second, the observed levels of engagement and positive user experience may have been influenced by a novelty effect associated with the VR environment. Participants' responses could partly reflect the initial excitement or unfamiliarity with immersive VR technology rather than stable or sustained engagement attributable to the intervention itself. As the study design did not include repeated exposure or long-term follow-up, it is not possible to distinguish the effects from enduring

XSL•FO
RenderX

behavioral or experiential changes. Therefore, any implications regarding improved adherence or sustained engagement should be interpreted with caution.

Third, recruitment was resource constrained and limited to community participants in northern Sweden. Although random assignment helped mitigate selection bias, most participants did not have a clinical diagnosis of anxiety disorder. Future research should include clinically diagnosed populations to better evaluate the therapeutic relevance and applicability of the intervention.

Fourth, the intervention consisted of a single 15-minute session, which limits insights into sustained engagement, habituation effects, or long-term impact. Other studies incorporating repeated sessions are necessary to assess whether engagement and perceived benefits persist beyond initial exposure.

Fifth, the broad age range of the participants (18 - 39 years) may have influenced how individuals related to the cleaning-based tasks as perceptions of relevance and motivation can vary across life stages.

Finally, as the study did not include postintervention measures of anxiety, stress, or mindfulness, it is not possible to determine whether the observed increases in engagement translate into measurable psychological or therapeutic outcomes. To address this limitation, future research should incorporate both pre- and postintervention assessments using validated psychological instruments, ideally within a controlled clinical trial design. Longitudinal studies with repeated exposure sessions would further help determine the durability of engagement effects. Such designs would allow for a more comprehensive evaluation of both usability outcomes and potential clinical effectiveness.

Conclusions

This study aimed to develop and evaluate a gamified VR application designed to support individuals experiencing mild anxiety by enhancing engagement, focus, and immersion. The findings indicate that incorporating gamification and persuasive design elements significantly improved user experience metrics, including perceived flow and satisfaction, compared to the nongamified version.

As the study focused on engagement-related outcomes and did not include postintervention anxiety assessments, the results should not be interpreted as evidence of clinical effectiveness. Rather, the findings demonstrate that gamified VR environments can successfully increase user immersion and interaction quality in short-term exposure settings.

User feedback further highlighted the importance of audiovisual design quality and appropriate challenge levels in maintaining engagement. These insights contribute to the growing body of research exploring how immersive technologies can be optimized to enhance user experience in mental health–related applications.

Overall, this study provides empirical evidence that gamification strategies can positively influence engagement and immersion within VR-based applications targeting individuals with mild anxiety.

Acknowledgments

The authors sincerely thank all the participants, especially the students from Luleå University of Technology, for their time, openness, and valuable contributions to this study. The authors also acknowledge the use of ChatGPT (OpenAI) to assist with language editing and improving the clarity of this manuscript. All content was reviewed, revised, and approved by the authors, who take full responsibility for the final version of the manuscript.

Funding

This research received funding from PRECISE at Luleå University of Technology (grant LTU-3515-2022) for the “Personalized virtual reality and gamification intervention for addressing mental health” project.

Data Availability

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

Authors' Contributions

Conceptualization: MMA, SSO, NJ
Data curation: MMA
Funding acquisition: SSO, NJ, MBTN
Investigation: MMA, NJ, SSO
Methodology: MMA, NJ, SSO, MBTN
Project administration: MMA, NJ, SSO, MBTN
Resources: MMA
Software: MMA
Validation: MMA, NJ, SSO, MBTN
Writing—original draft: MMA, NJ, SSO
Writing—review and editing: MMA, NJ, SSO, MBTN
All authors contributed to the article and approved the submitted version.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Implementation of the Octalysis core drivers in the *Cleanify* virtual reality application.
[DOCX File, 8 KB - xr_v3i1e66713_app1.docx]

Multimedia Appendix 2
Level 3—space organization layout.
[PNG File, 1061 KB - xr_v3i1e66713_app2.png]

Multimedia Appendix 3
Badge and star reward system in the *Cleanify* virtual reality application.
[PNG File, 461 KB - xr_v3i1e66713_app3.png]

Multimedia Appendix 4
Full questionnaire items.
[DOCX File, 8 KB - xr_v3i1e66713_app4.docx]

References

1. SV G, RJ V, P I. An overview of psychosomatic diseases and homoeopathy. Int J Homoeopathic Sci 2022;6(4):147-149. [doi: 10.33545/26164485.2022.v6.i4c.650]
2. Moitra M, Santomauro D, Collins PY, et al. The global gap in treatment coverage for major depressive disorder in 84 countries from 2000-2019: a systematic review and Bayesian meta-regression analysis. PLoS Med 2022 Feb;19(2):e1003901. [doi: 10.1371/journal.pmed.1003901] [Medline: 35167593]
3. COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021 Nov 6;398(10312):1700-1712. [doi: 10.1016/S0140-6736(21)02143-7] [Medline: 34634250]
4. Navarro-Haro MV, Modrego-Alarcón M, Hoffman HG, et al. Evaluation of a mindfulness-based intervention with and without Virtual Reality Dialectical Behavior Therapy® mindfulness skills training for the treatment of generalized anxiety disorder in primary care: a pilot study. Front Psychol 2019;10:55. [doi: 10.3389/fpsyg.2019.00055] [Medline: 30745888]
5. Lieb R, Becker E, Altamura C. The epidemiology of generalized anxiety disorder in Europe. Eur Neuropsychopharmacol 2005 Aug;15(4):445-452. [doi: 10.1016/j.euroneuro.2005.04.010] [Medline: 15951160]
6. Allgulander C. Generalized anxiety disorder: what are we missing? Eur Neuropsychopharmacol 2006 Jul;16 Suppl 2:S101-S108. [doi: 10.1016/j.euroneuro.2006.04.002] [Medline: 16730165]
7. Tyrer P, Baldwin D. Generalised anxiety disorder. Lancet 2006 Dec 16;368(9553):2156-2166. [doi: 10.1016/S0140-6736(06)69865-6] [Medline: 17174708]
8. Roemer L, Lee JK, Salters-Pedneault K, Erisman SM, Orsillo SM, Mennin DS. Mindfulness and emotion regulation difficulties in generalized anxiety disorder: preliminary evidence for independent and overlapping contributions. Behav Ther 2009 Jun;40(2):142-154. [doi: 10.1016/j.beth.2008.04.001] [Medline: 19433145]
9. Hoge EA, Bui E, Marques L, et al. Randomized controlled trial of mindfulness meditation for generalized anxiety disorder: effects on anxiety and stress reactivity. J Clin Psychiatry 2013 Aug;74(8):786-792. [doi: 10.4088/JCP.12m08083] [Medline: 23541163]
10. Dellazizzo L, Potvin S, Luigi M, Dumais A. Evidence on virtual reality-based therapies for psychiatric disorders: meta-review of meta-analyses. J Med Internet Res 2020 Aug 19;22(8):e20889. [doi: 10.2196/20889] [Medline: 32812889]
11. Li Pira G, Aquilini B, Davoli A, Grandi S, Ruini C. The use of virtual reality interventions to promote positive mental health: systematic literature review. JMIR Ment Health 2023 Jul 6;10:e44998. [doi: 10.2196/44998] [Medline: 37410520]
12. Olenichenko I. Application of VR-technology methods in psychology and psychotherapy. Global Psychother 2023 Jul;3(2):89-95. [doi: 10.52982/lkj202]
13. Liu Z. Application of virtual reality technology in the psychological field. Appl Comput Eng 2023 Jun;6(1):533-539. [doi: 10.54254/2755-2721/6/20230853]
14. Jingili N, Oyelere SS, Nyström MB, Anyshchenko L. A systematic review on the efficacy of virtual reality and gamification interventions for managing anxiety and depression. Front Digit Health 2023;5:1239435. [doi: 10.3389/fdgth.2023.1239435] [Medline: 38026832]
15. Jingili N, Oyelere SS, Ojwang F, Agbo FJ, Nyström MB. Virtual reality for addressing depression and anxiety: a bibliometric analysis. Int J Environ Res Public Health 2023 Apr 24;20(9):5621. [doi: 10.3390/ijerph20095621] [Medline: 37174141]

XSL•FO
RenderX

16. Riches S, Jeyarajaguru P, Taylor L, et al. Virtual reality relaxation for people with mental health conditions: a systematic review. Soc Psychiatry Psychiatr Epidemiol 2023 Jul;58(7):989-1007. [doi: 10.1007/s00127-022-02417-5] [Medline: 36658261]
17. Hoffman HG, Doctor JN, Patterson DR, Carrougher GJ, Furness TA3. Virtual reality as an adjunctive pain control during burn wound care in adolescent patients. Pain 2000 Mar;85(1-2):305-309. [doi: 10.1016/s0304-3959(99)00275-4] [Medline: 10692634]
18. Botella C, Serrano B, Baños RM, Garcia-Palacios A. Virtual reality exposure-based therapy for the treatment of post-traumatic stress disorder: a review of its efficacy, the adequacy of the treatment protocol, and its acceptability. Neuropsychiatr Dis Treat 2015;11:2533-2545. [doi: 10.2147/NDT.S89542] [Medline: 26491332]
19. Gutiérrez-Maldonado J, Wiederhold BK, Riva G. Future directions: how virtual reality can further improve the assessment and treatment of eating disorders and obesity. Cyberpsychol Behav Soc Netw 2016 Feb;19(2):148-153. [doi: 10.1089/cyber.2015.0412] [Medline: 26378982]
20. Cheng VW. Recommendations for implementing gamification for mental health and wellbeing. Front Psychol 2020;11:586379. [doi: 10.3389/fpsyg.2020.586379] [Medline: 33365001]
21. Jingili N, Oyelere SS, Berghem SM, et al. A two-stage co-design process of Battleship-AST persuasive game for active school transportation in Northern Sweden. Int J Hum Comput Interact 2025 Apr 18;41(8):4888-4909. [doi: 10.1080/10447318.2024.2355395]
22. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining “gamification”. In: MindTrek ’11: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments: Association for Computing Machinery; 2011. [doi: 10.1145/2181037.2181040]
23. Páez-Quinde C, Arroba-Freire E, Espinosa-Jaramillo MT, Silva MP. Gamification as collaborative learning resources in technological education. In: 2023 IEEE Global Engineering Education Conference (EDUCON): IEEE; 2023. [doi: 10.1109/EDUCON54358.2023.10125264]
24. Sun A, Conde S, Elor A. Increasing sociability in a virtual world: a serious game for social anxiety disorder. In: 2021 IEEE 9th International Conference on Serious Games and Applications for Health (SeGAH): IEEE; 2021. [doi: 10.1109/SEGAH52098.2021.9551867]
25. El-Yamri M, Romero-Hernandez A, Gonzalez-Riojo M, Manero B. Designing a VR game for public speaking based on speakers features: a case study. Smart Learn Environ 2019;6(1). [doi: 10.1186/s40561-019-0094-1]
26. Miloff A, Lindner P, Dafgård P, et al. Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: a randomized non-inferiority trial. Behav Res Ther 2019 Jul;118:130-140. [doi: 10.1016/j.brat.2019.04.004] [Medline: 31075675]
27. Rogers CJ, Hart R. Home and the extended-self: exploring associations between clutter and wellbeing. J Environ Psychol 2021;73:101553. [doi: 10.1016/j.jenvp.2021.101553]
28. Bodrij FF, Andeweg SM, Prevoo MJ, Rippe RC, Alink LR. The causal effect of household chaos on stress and caregiving: an experimental study. Compr Psychoneuroendocrinol 2021;8:100090. [doi: 10.1016/j.cpnec.2021.100090] [Medline: 35757675]
29. Saxbe DE, Repetti R. No place like home: home tours correlate with daily patterns of mood and cortisol. Pers Soc Psychol Bull 2010 Jan;36(1):71-81. [doi: 10.1177/0146167209352864] [Medline: 19934011]
30. McMains S, Kastner S. Interactions of top-down and bottom-up mechanisms in human visual cortex. J Neurosci 2011 Jan 12;31(2):587-597. [doi: 10.1523/JNEUROSCI.3766-10.2011] [Medline: 21228167]
31. Hanley AW, Warner AR, Dehili VM, Canto AI, Garland EL. Washing dishes to wash the dishes: brief instruction in an informal mindfulness practice. Mindfulness 2015 Oct;6(5):1095-1103. [doi: 10.1007/s12671-014-0360-9]
32. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Int Syst 2009;24:485-500. [doi: 10.17705/1CAIS.02428]
33. Diagnostic and Statistical Manual of Mental Disorders: American Psychiatric Association; 2013. [doi: 10.1176/appi.books.9780890425596]
34. Pressman SD, Matthews KA, Cohen S, et al. Association of enjoyable leisure activities with psychological and physical well-being. Psychosom Med 2009 Sep;71(7):725-732. [doi: 10.1097/PSY.0b013e3181ad7978] [Medline: 19592515]
35. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J Altern Complement Med 2009 May;15(5):593-600. [doi: 10.1089/acm.2008.0495] [Medline: 19432513]
36. Carek PJ, Laibstain SE, Carek SM. Exercise for the treatment of depression and anxiety. Int J Psychiatry Med 2011;41(1):15-28. [doi: 10.2190/PM.41.1.c] [Medline: 21495519]
37. Wood AM, Tarrier N. Positive Clinical Psychology: a new vision and strategy for integrated research and practice. Clin Psychol Rev 2010 Nov;30(7):819-829. [doi: 10.1016/j.cpr.2010.06.003] [Medline: 20655136]
38. Fogg BJ. Persuasive Technology: Using Computers to Change What We Think and Do: Morgan Kaufmann; 2003. URL: https://books.google.co.in/books/about/Persuasive_Technology.html?id=gfiYh_BHj94C&source=kp_book_description&redir_esc=y [accessed 2026-06-15]
39. Fletcher-Watson S, Brook K, Hallett S, Murray F, Crompton CJ. Inclusive practices for neurodevelopmental research. Curr Dev Disord Rep 2021 Jun;8:88-97. [doi: 10.1007/s40474-021-00227-z]

XSL•FO
RenderX

40. Gillihan SJ. Retrain Your Brain: Cognitive Behavioral Therapy in 7 Weeks: A Workbook for Managing Depression and Anxiety: Althea Press; 2016.
41. Korpela K, Kinnunen U. How is leisure time interacting with nature related to the need for recovery from work demands? Testing multiple mediators. Leis Sci 2010;33(1):1-14. [doi: 10.1080/01490400.2011.533103]
42. Roster CA, Ferrari JR, Jurkat MP. The dark side of home: assessing possession 'clutter' on subjective well-being. J Environ Psychol 2016 Jun;46:32-41. [doi: 10.1016/j.jenvp.2016.03.003]
43. Craven MP, Groom MJ. Computer games for user engagement in Attention Deficit Hyperactivity Disorder (ADHD) monitoring and therapy. Presented at: 2015 International Conference on Interactive Technologies and Games (ITAG); Oct 22-23, 2015. [doi: 10.1109/iTAG.2015.9]
44. Cheng C, Ebrahimi OV. Gamification: a novel approach to mental health promotion. Curr Psychiatry Rep 2023 Nov;25(11):577-586. [doi: 10.1007/s11920-023-01453-5] [Medline: 37801212]
45. Park S, Kim S. Leaderboard design principles to enhance learning and motivation in a gamified educational environment: development study. JMIR Serious Games 2021 Apr 20;9(2):e14746. [doi: 10.2196/14746] [Medline: 33877049]
46. Yang Z, Han S, Zhang L, et al. Dose response effects of patient engagement on anxiety and depression in a cognitive-behavioral intervention: secondary analysis of a pilot randomized controlled trial and a clinical controlled trial. AIDS Behav 2024 Jun;28(6):1923-1935. [doi: 10.1007/s10461-024-04290-6] [Medline: 38570384]
47. Swinson RP. The GAD-7 scale was accurate for diagnosing generalised anxiety disorder. Evid Based Med 2006 Dec;11(6):184. [doi: 10.1136/ebm.11.6.184] [Medline: 17213178]
48. Schrepp M, Hinderks A, Thomaschewski J. Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). Int J Interact Multim Artif Intell 2017;4(6):103-108. [doi: 10.9781/ijimai.2017.09.001]
49. Schwind V, Knierim P, Haas N, Henze N. Using presence questionnaires in virtual reality. In: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems: Association for Computing Machinery; 2019. [doi: 10.1145/3290605.3300590]
50. Engeser S, Rheinberg F. Flow, performance and moderators of challenge-skill balance. Motiv Emot 2008 Sep;32(3):158-172. [doi: 10.1007/s11031-008-9102-4]
51. Rheinberg F, Manig Y, Kliegl R, Engeser S, Vollmeyer R. Flow during work but happiness during leisure time: goals, flow-experience, and happiness [Article in German]. Zeitschrift Arbeitsund Organisationspsychologie 2007;51(3):105-115. [doi: 10.1026/0932-4089.51.3.105]
52. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
53. Behringer R, Gram-Hansen SB, Soosay M, et al. Persuasive technology for learning in business context. Int J Inf Syst Eng 2013 Apr;1(1). [doi: 10.24924/ijise/2013.04/v1.iss1/86.97]
54. Kothgassner OD, Goreis A, Bauda I, Ziegenaus A, Glenk LM, Felnhofer A. Virtual reality biofeedback interventions for treating anxiety: a systematic review, meta-analysis and future perspective. Wien Klin Wochenschr 2022 Jan;134(Suppl 1):49-59. [doi: 10.1007/s00508-021-01991-z] [Medline: 34989862]
55. Lin AJ, Cheng F, Chen CB. Use of virtual reality games in people with depression and anxiety. In: ICMIP '20: Proceedings of the 5th International Conference on Multimedia and Image Processing: Association for Computing Machinery; 2020. [doi: 10.1145/3381271.3381299]
56. Hoffman HG, Patterson DR, Carrougher GJ. Use of virtual reality for adjunctive treatment of adult burn pain during physical therapy: a controlled study. Clin J Pain 2000 Sep;16(3):244-250. [doi: 10.1097/00002508-200009000-00010] [Medline: 11014398]

Abbreviations

FSS: Flow State Scale
GAD-7: Generalized Anxiety Disorder–7
UEQ-S: short version of the User Experience Questionnaire
VR: virtual reality

Edited by I Steenstra, T Leung; submitted 20.Sep.2024; peer-reviewed by S Lord; revised version received 20.Apr.2026; accepted 30.Apr.2026; published 24.Jun.2026.

Please cite as:
Argaw MM, Jingili N, Oyelere SS, Nyström MBT
Persuasive Gamified Virtual Reality Experience to Enhance Engagement and Focus in Young Adults With Mild Anxiety Symptoms: Randomized Pilot Experimental Study
JMIR XR Spatial Comput 2026;3:e66713
URL: https://xr.jmir.org/2026/1/e66713
doi:10.2196/66713

XSL•FO
RenderX

© Mahlet Misrak Argaw, Nuru Jingili, Solomon Sunday Oyelere, Markus B T Nyström. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 24.Jun.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Improving Health-Related Symptoms and Behaviors in Children and Adolescents Diagnosed With Diabetes by Using a Virtual Reality–Gamified Self-Care Method: Randomized Controlled Trial

Ruhollah Solat[1], MSc; Reza Pourhosein[1], PhD; Hadi Bahrami Ehsan[1], PhD; Ata Pourabbasi[2], PhD; Amin Hashemi[1], PhD; Javad Hatami[1], PhD

[1]Psychology Department, Tehran University, Dr. Kardan Street, Tehran, Iran
[2]Neuroscience Department, Tehran University, Tehran, Iran

Corresponding Author:
Javad Hatami, PhD
Psychology Department, Tehran University, Dr. Kardan Street, Tehran, Iran

Abstract

Background: Self-care plays an important role in improving health symptoms in patients diagnosed with chronic illnesses such as diabetes. Gamification is among the most effective methods for enhancing health monitoring by applying principles of behavioral economics and motivation.

Objective: This study aimed to develop an innovative self-care system using a virtual reality (VR)–gamified intervention and to evaluate its effects on patients diagnosed with diabetes metabolic symptoms and health-related behaviors compared to other standard interventions.

Methods: This study was a randomized controlled trial with a parallel-group design, conducted between December 2024 and February 2025. A total number of 78 patients diagnosed with diabetes were recruited from The Children's Medical Center clinic in a closed, offline setting. Briefing, tutorials, and interventions were partly face-to-face. Eligible participants (n=68) were randomly assigned to the 4 groups using stratified block randomization based on age and sex, with allocation concealment ensured through a sequentially numbered process. Each group received a different type of self-care intervention for 6 weeks, which was health care provider-assisted. One group received the VR gamified intervention, while the other 3 groups received comparator interventions, including the MySugr (mySugr GmbH) monitoring app, traditional counseling, and medication only, respectively. The primary outcomes were fasting blood sugar and health-related behaviors, including physical activity and food intake, while secondary outcomes were long-term metabolic indicators, including hemoglobin A_{1c} (HbA_{1c}) and BMI. All outcomes were measured by self-assessed questionnaires. No blinding was implemented for participants or care providers. However, data analysis was conducted using anonymized group labels. Eight participants were lost to follow-up. Minor missing daily entries were handled by weekly data aggregation.

Results: Data from 60 participants were included in the final analysis (VR-gamified group n=16, application group n=14, counseling group n=15, and control group n=15). Linear mixed model showed significant time×group interaction effects for fasting blood sugar ($F_{15,280}$=2.046; P=.01; partial η^2=.099), physical activity ($F_{15,280.61}$=2.544; P=.001; partial η^2=.120), and food intake ($F_{15,336}$=2.794; P<.001; partial η^2=.111), indicating greater improvement over time in the VR gamification group compared with the other parallel groups. No statistically significant differences were observed for BMI or glycated hemoglobin. No serious adverse events related to the VR intervention were reported.

Conclusions: VR gamified self-care interventions may potentially contribute to better management of diabetes-related symptoms and behaviors compared with conventional self-care approaches. Such gamified VR systems show promise as alternative or complementary tools for enhancing health outcomes among children and adolescents diagnosed with diabetes.

Trial Registration: Iranian Registry of Clinical Trials IRCT 89347; https://www.irct.ir/trial/89347

(JMIR XR Spatial Comput 2026;3:e81402) doi:10.2196/81402

KEYWORDS

diabetes; gamification; health; self-care; virtual reality

XSL•FO
RenderX

Introduction

Background

Self-care plays an important role in managing the progression of chronic diseases. A high level of control and effective symptom management can improve both somatic and psychological well-being in patients [1]. Self-care systems use personal goal-setting, social support, self-monitoring, and education to help patients in their daily lives [2]. Among various self-care approaches, technological methods appear to be more effective than others. The global availability and relatively low cost of technologies such as computers and smartphones make them suitable for remote monitoring and self-management [3].

Diabetes mellitus (DM) is a chronic disease that is directly influenced by lifestyle, including habits and behaviors related to glycemic control, dietary adherence, and physical activity. Therefore, encouraging motivation for self-care is a crucial aspect of diabetes treatment [4]. Poor control of diabetic symptoms can lead to mortality, while increasing physical activity and improving glycemic control can help prevent morbidity [5]. According to the "Association of Diabetes Care and Education Specialists 7" framework, any diabetes self-care approach must have a measurable impact on specific factors. These include a healthy diet, physical activity, medication management, and regular blood glucose monitoring [6].

Studies have demonstrated the relative effectiveness of health monitoring applications in enhancing daily self-care among patients diagnosed with diabetes. Numerous programs utilize health tracking processes in the context of diabetes self-care [2]. Most apps focus on blood glucose monitoring, which has been identified as the most important metric for diabetes management [6]. One well-known example is the Glucose Buddy (Azumio Inc) app, which allows users to manually log and track diabetes-related metrics such as blood glucose levels [7]. Mobile self-monitoring apps should include multiple features, such as blood glucose tracking, physical activity monitoring, food intake logging, reminders, note-taking, and social support, to be fully functional [8]. However, it has been shown that relying solely on self-monitoring features without incorporating gamification can limit the program's effectiveness in improving self-care behaviors [9]. Platforms like SlimKicker (Henley Wing Chiu) are examples of systems designed for tracking health metrics and physical activity without incorporating engaging game design elements [10]. In contrast, MySugr (Roche diabetes care) is a diabetes management app that combines health monitoring with gamification elements to enhance user engagement and adherence [11].

Gamification is the process of using game elements to engage and motivate individuals in serious contexts unrelated to entertainment [10]. Applying gamification in health care is an effective tool for shaping patient behavior through user engagement [12]. Techniques such as self-monitoring, goal setting, reward structures, social incentives, level design, medal/badge systems, and narrative context are among the most well-known and widely used behavior change strategies [13]. Reward and incentive systems have a powerful impact on habit formation and can foster intrinsic motivation for users to repeatedly engage in desired behaviors [14].

It has been shown that gamification is more effective in establishing desirable healthy habits than health monitoring alone [9]. Evidence indicates that gamification in mobile health apps leads to measurable improvements in glycemic control and medication adherence [15]. As an example, the MySugr program allows users to track and monitor their blood glucose levels, medications, and calorie intake while using challenges, achievements, and a monster avatar to encourage regular data logging [14]. Gamification enhances user motivation and supports behavior change by making health-related tasks more enjoyable and goal-oriented [16]. Techniques like point scoring, social incentives, and predictable rewards for desired actions effectively facilitate habit formation and behavioral change by reinforcing target behaviors and shaping lasting routines [17]. One of the most notable examples is exergames, games designed to promote physical activity. Games like Bant (University Health Network) track users' physical activity using GPS and encourage improvement [10]. It has been shown that incorporating scoring and reward systems in exergames can effectively foster sustained physical activity in users [18].

However, sole reliance on points and rewards can undermine the essence of gamification, making the experience less entertaining and engaging. Scoring elements must be integrated into a compelling and enjoyable context [14]. Studies indicate that using narrative and engaging audiovisual designs, similar to features found in video games, can convey a sense of fun, immersion, and pleasure, helping to reduce the perception of pain or discomfort [6]. A major flaw in traditional gamification designs is the absence of essential game elements such as narration, high-quality graphics, and gameplay, resulting in emotionally unappealing experiences for children and adolescents [19]. Integrating visually engaging elements and fun designs helps maintain user interest and enhances the educational value of diabetes self-care apps [20]. By using compulsion loops, cycles involving anticipation of a reward, performing an action, and receiving the reward, gamified systems can reinforce desired behavioral habits [21]. Some studies describe gamification as an art form, emphasizing that the most effective way to foster behavioral change is through well-designed reward systems. Aesthetically appealing and strategically crafted rewards engage both cognitive and motivational systems to shape lasting habits [22].

Apart from using the mentioned game design elements, creating an immersive atmosphere is necessary to reach an engaging experience. Toward this goal, the use of virtual reality (VR) platforms has been increasingly discussed in recent years [23]. Jo and Park [24] found that VR's visual aesthetics, including shape, depth, 3D graphics, and visual design, significantly influence perceived enjoyment and adherence to the technology. VR promotes behavior change through technological novelty focused on embodied experience. By incorporating interactive designs, VR creates a sense of flow, presence, and immersion [25]. Studies show VR environments can encourage healthy behaviors, including regular exercise and proper nutrition [26].

XSL•FO
RenderX

Objectives

The main goal of this research was to design an innovative gamification program for diabetes self-care. To our knowledge, there is limited evidence of gamified VR programs that integrate similar audiovisual and gameplay elements for diabetes self-care. This study helps address this gap by providing an initial implementation and evaluation of such a system. Following the development of the app, the next objective of our research was to assess its effectiveness in improving diabetes symptoms.

To evaluate the efficiency of every self-care method, such as our VR-gamified app, there are some metrics used in previous literature. Blood glucose monitoring is recognized as the most important metric of diabetes self-care [6]. Fasting blood sugar (FBS) is considered a short-term marker of blood glucose levels, while hemoglobin A_{1c} (HbA_{1c}) serves as a long-term indicator [27]. Research has identified several key factors that influence blood glucose control, with dietary adherence and physical activity being among the most significant predictors of glycemic regulation [28]. In line with these findings, we used 5 key metrics to evaluate the effectiveness of our designed intervention. These metrics are categorized into two main groups: (1) health symptoms, including FBS, HbA_{1c}, and BMI, and (2) health-related behaviors, including physical activity and food intake.

We hypothesized that experiencing our VR gamified program would have a greater impact on improving the primary self-care outcomes for diabetes. Based on this expectation, the following 5 hypotheses were formulated:

- Hypothesis 1 (H1): the improvement in the weekly average of FBS over time differs between the VR gamified group and the other experimental groups.
- Hypothesis 2 (H2): the improvement in the weekly total physical activity over time differs between the VR gamified group and the other experimental groups.
- Hypothesis 3 (H3): the improvement in the weekly average of daily food intake over time differs between the VR-gamified group and the other experimental groups.
- Hypothesis 4 (H4): the improvement in the weekly calculated BMI over time differs between the VR-gamified group and the other experimental groups.
- Hypothesis 5 (H5): the change in HbA_{1c} from pretest to posttest observations differs between the VR-gamified group and the other experimental groups.

Methods

Trial Design

The study used a parallel-group randomized controlled trial comprising 4 arms, with participants allocated to the groups with respective sizes of 16, 14, 15, and 15 for final analysis. Each group received one kind of self-care intervention for a 6-week period.

Two methodological changes were implemented after trial commencement. First, the physical activity measurement approach was modified from objective pedometer-based tracking to structured self-report assessment to enhance feasibility and adherence in the clinical setting. Second, the Iranian diabetes-monitoring app Idia (Mehr Nutrition Technology and Knowledge Development) was initially planned as a comparator for the VR-gamified group; however, limitations in its expected feature set necessitated its replacement with the MySugr app as one of the comparators.

Participants

The target population for this study consisted of children and adolescents aged 7-15 years diagnosed with diabetes, living in Tehran, Iran. Inclusion criteria were (1) a formal diagnosis of diabetes, (2) aged between 7 and 15 years, (3) basic computer/VR literacy, and (4) willingness and ability to participate in a 6-week eHealth trial. Exclusion criteria included any medical conditions that would prevent safe use of VR headsets (eg, cybersickness or nausea). The research team was introduced to potential participants as specialists in health psychology affiliated with the PhD program at the University of Tehran in order to provide a clear professional context for the intervention. A total of 78 participants meeting the inclusion criteria were recruited from The Children's Medical Center in August 2024 to November 2024, Tehran, Iran, in an offline, closed manner. Following recruitment, participants were provided with face-to-face instructions and a briefing. Written informed consent was subsequently obtained from eligible participants (n=68) who enrolled in the trial. Eight participants lost to follow-up over the 6-week period, resulting in a final sample of 60 participants (mean age 11.8, SD 1.74 years; 33, 55.0% boys and 27, 45.0% girls). All outcomes were self-assessed by online questionnaires throughout the trial timeline.

Interventions

VR Gamified Group

Participants in this group engaged with the VR gamified program developed specifically for this study.

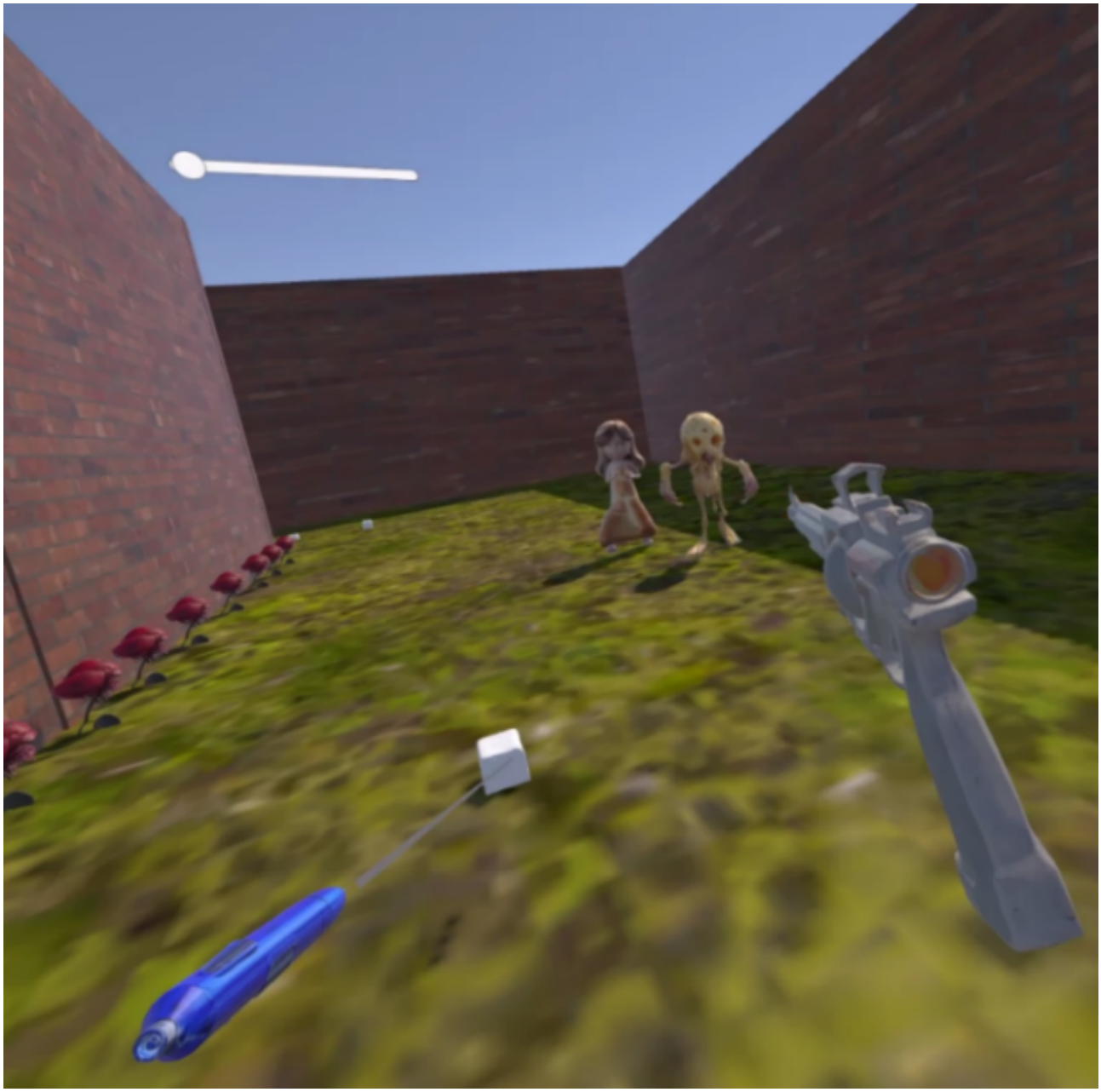
Participants engaged in one supervised VR gamified session per week for 6 consecutive weeks. Each session lasted approximately 5 - 15 minutes and took place in the clinical setting. In every session, participants played one level of the KalleGhandi (Ruhollah Solat) VR game, with the difficulty adjusted based on their weekly glycemic and behavioral data obtained prior to the session. Two care providers, including one health professional and one technical assistant, were present to provide onboarding, headset setup, and technical support throughout the intervention. Participants accessed the VR game at the medical clinic following their weekly physician visit. No payment was required from participants. The game was experienced using a VR headset provided by the university, and participants did not have personal access to the device outside the clinical setting. No prompts were used, but weekly Telegram (Telegram Messenger LLP) messages were sent to participants to remind them about the intervention and related self-assessments. Apart from slight differences in the level of training, no co-intervention was provided during the trial.

The VR game used in this study (see Figure 1) was designed and developed by the first author (RS). With part-time support

XSL•FO
RenderX

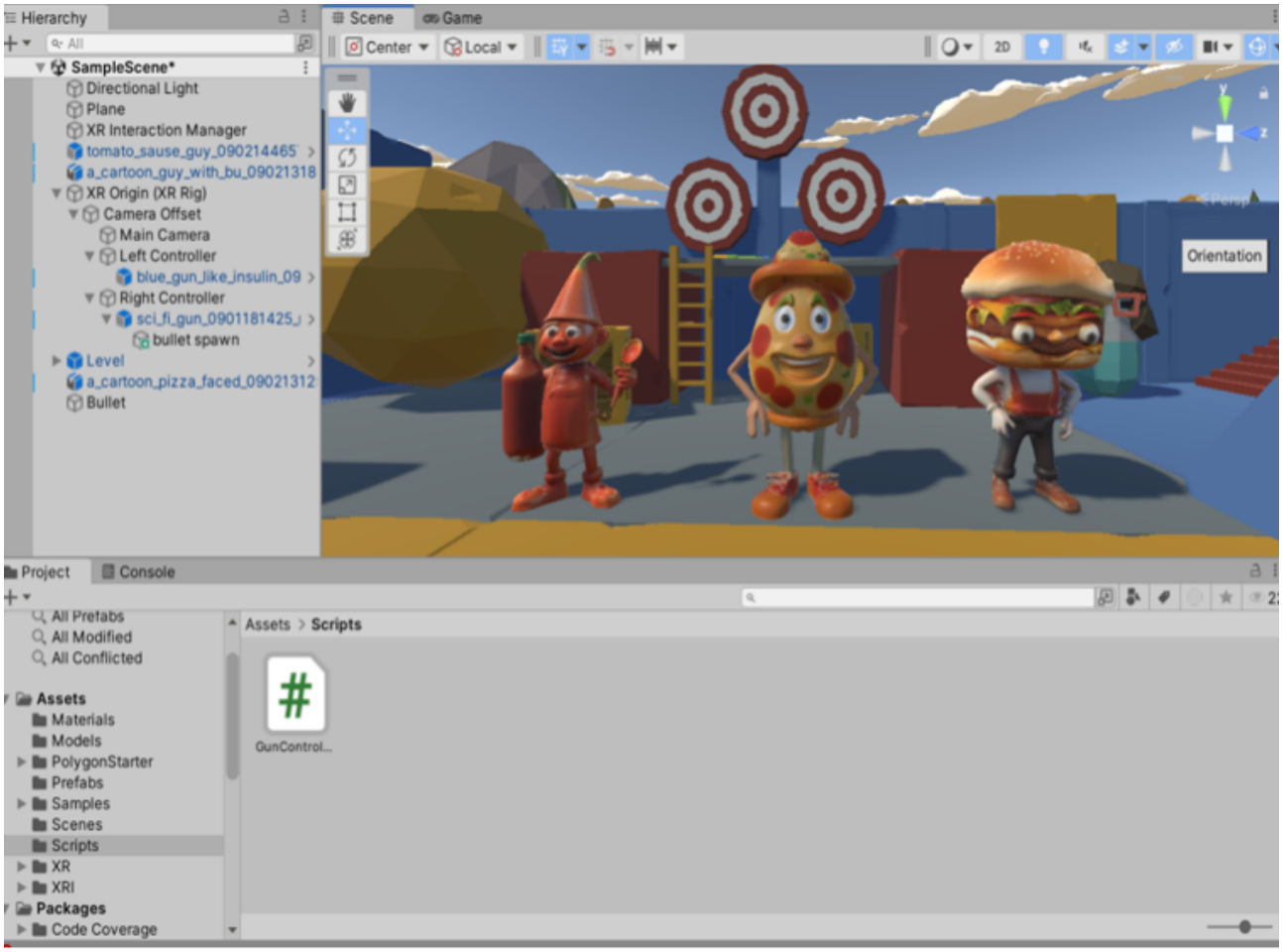
from a technical game development team in Iran, the product was developed and finalized through several iterative phases between November 2023 and October 2024, using the Unity engine (version 2022.3.44f1; Unity Technologies) and tested on the Oculus Meta Quest 2 (Meta Platforms) VR headset. Figure 2 shows one screenshot of the process of making the game in the Unity engine. For ensuring quality assurance and taking feedback about the experience, the game was pilot-tested multiple times with participants without diabetes prior to the main study. The game was designed in the first-person shooter genre. In the first core mechanic of the game, players use their VR right-hand controller to shoot enemies represented as unhealthy food items such as ice cream, pizza, and cheese. The second core mechanic involves using an insulin pen with the VR left-hand controller to break down and collect sugar cubes scattered throughout the environment. Each session lasts 5 minutes, and based on the number of sugar cubes collected, participants are awarded a medal: gold, silver, or bronze.

Figure 1. Visual overview of the "KalleGhandi" virtual reality gamification program used as an intervention in this randomized controlled trial study.



XSL•FO
RenderX

Figure 2. A screenshot taken from the process of making the game "KalleGhandi." Unity 2022.3.44fl game engine.



The gamification design links each gameplay element directly to real-life diabetes management behaviors. Specifically, the target number of sugar cubes is dynamically set based on the participant's weekly average FBS. Lower FBS averages reduce the target, making it easier to earn medals, while higher FBS averages increase the difficulty. Similarly, the player's movement speed in the game is determined by their physical activity level, and the number of enemies (unhealthy foods) is determined by the weekly average food calories. This integration of real-life health indicators into gameplay is intended to reinforce positive self-care behaviors.

To facilitate accessibility and ensure replicability, the source codes of the KalleGhandi game are provided in Multimedia Appendix 1, while 2 versions of the game have been uploaded to the Open Science Framework (OSF):

- The Windows build (for playing on a PC with a VR headset connection) is available [29].
- The Android build (for stand-alone headset use) is available [30].

Usability and Tolerability Assessment

Throughout the VR intervention, participants and caregivers were monitored for usability issues, cybersickness, or discomfort. Tolerability indicators (nausea, dizziness, and visual fatigue) were recorded via caregiver observations at each session. No adverse events or dropouts due to VR intolerance occurred. All sessions were conducted under adult supervision, with optional rest breaks. In addition to observational monitoring, participants completed standardized measures to capture their subjective experience of ease of use, perceived usefulness, and overall interaction quality with the VR system. System usability was measured using the System Usability Scale developed by Brooke [31], which consists of 10 items rated on a 5-point Likert scale. Technology acceptance was measured based on the Technology Acceptance Model (TAM) construct developed by Davis [32]. The TAM model's core constructs, perceived usefulness, and perceived ease of use were operationalized to assess users' perception of extended reality (XR) technology. After experiencing 6 consecutive sessions, participants completed the Persian-translated versions of both questionnaires to ensure clarity and language accessibility for children and adolescent participants.

Application Group

Participants in this group used the commercially available diabetes self-care app MySugr (Android/iOS, standard free version available during the study period 2024 - 2025) as a comparator to the VR gamified intervention.

At baseline, participants received a brief, structured training session (approximately 20 - 30 min) on how to:

- Create a personal user account and log in to the app;
- Log their blood glucose measurements obtained from their home glucometer;

XSL•FO
RenderX

- Record their dietary intake and approximate calorie consumption; and
- Review the summary feedback provided by the app.

Participants used the MySugr app to log and self-monitor their blood glucose values and dietary intake. They received automated reminders prompting them to record glucose measurements and diet-related information. Participants also interacted with the app's gamified "monster" avatar, which responded to completed logging tasks by providing visual feedback and rewards; this feature constituted the primary gamification component of the intervention.

Participants were instructed to use the app between 2 and 5 times per day, typically for several minutes per session, to enter their fasting and postprandial blood glucose values and main meals. App usage (frequency and duration) was based on self-report during follow-up visits/phone calls rather than backend log data, as we did not have access to the app's internal usage analytics.

Traditional Counseling Group

This group received diabetes self-care counseling provided by professional diabetes consultancy centers in Iran.

Participants were already enrolled in diabetes care programs at either Gabric Institute or the Iranian Diabetes Center, where they routinely received counseling and self-care support alongside medication services. Counseling interactions typically occurred multiple times per week via Telegram channels and phone calls. Each participant had a primary counselor familiar with their case and family, offering tailored guidance for diabetes self-care. Institutional practices are largely uniform but not scripted.

Control Group

Participants in this group received medical care by an endocrinology/child-adolescent diabetes/nutrition specialist on a weekly or monthly basis but did not receive any additional psychological health-related self-care intervention.

Outcomes

All participants (or their parents) were instructed to complete a structured data collection form throughout the 6-week study period, which is provided in Multimedia Appendix 2. The form required daily entries, including FBS, food intake, and physical activity, along with weekly entries for body weight. HbA_{1c} levels were recorded twice, once at baseline (pretest) and once after the final session (posttest). Demographic variables, including age, gender, and type of diabetes, were collected at the outset.

Primary Outcomes

Weekly Average Daily Monitored FBS

Participants measured their FBS levels every morning using a personal glucometer and self-reported the values. The recorded daily FBS values were then used by the experimenters to calculate a weekly average for each participant.

FBS values were measured by participants using personal home glucometers commonly available in community pharmacies. Because measurements were conducted in real-world settings, different commercially available glucometer brands like Accu-Check were used. Participants were instructed to measure blood glucose in the morning after an overnight fast (at least 8 h) following standard self-monitoring procedures recommended for home glucose testing. No continuous glucose monitoring devices were used in this study. Daily FBS values were recorded by participants and submitted through the study reporting forms.

Weekly Sum of Daily Calculated Physical Activity

Physical activity was quantified using a bar metric; one bar equaled either 30 minutes of vigorous activity (eg, gym workouts or swimming) or 60 minutes of moderate activity (eg, jogging). The weekly sum was calculated based on this definition.

As the study prioritized pragmatic feasibility in real-world settings, we intentionally relied on structured self-report measures rather than resource-intensive objective tools such as pedometers to record daily physical activity. Participants and parents received brief standardized training on reporting physical activity using the bar metric with examples of common daily exercises. This change in measurement methods happened shortly after the commencement of the trial.

Weekly Average Daily Calculated Food Intake

Participants either reported their total daily caloric intake directly or reported their food descriptions, from which caloric values were estimated. A weekly average of daily caloric intake was then computed.

To preserve pragmatic feasibility and participants' adherence, dietary intake was assessed using a self-assessment approach, rather than resource-intensive objective methods. Participants and parents received brief standardized training on estimating daily caloric intake using a nationally recognized Iranian food-calorie database (Dr Kermani's online nutrition resource [33]). Although self-report measures are inherently vulnerable to reporting bias, self-assessment contributed to operational feasibility, participant adherence, and real-world generalizability of the findings.

Secondary Outcomes

Weekly Calculated BMI

Participants self-reported their weight on a weekly basis. BMI was calculated using the standard formula: weight (kg) divided by height squared (m^2), which had been reported at first.

HbA_{1c} Pretest and Posttest

HbA_{1c} levels were measured twice via laboratory blood tests, once before the intervention (pretest) and once at the conclusion of the study (posttest), serving as a long-term indicator of blood glucose control.

Participants were instructed to undergo HbA_{1c} testing at a certified clinical laboratory of their choice using standard methods routinely used in clinical practice (eg, high-performance liquid chromatography or equivalent). All laboratories were accredited according to national standards. As HbA_{1c} measurements were collected from multiple

XSL•FO
RenderX

laboratories, the same laboratory was not used for all participants; however, HbA_{1c} is a standardized biomarker with minimal interlaboratory variability.

Sample Size

Sample size was calculated using G*Power (Heinrich Heine University Düsseldorf) software based on a repeated-measures ANOVA design with 4 groups. Assuming a medium effect size (f=.25), a significance level of α=.05, and statistical power of 0.80, the minimum required total sample size was estimated to be 52 participants. To account for potential attrition, additional participants were recruited. No interim analyses or stopping guidelines were defined or conducted.

Randomization

Participants were randomly assigned to one of the 4 groups using computer-generated stratified block randomization based on age and sex. The allocation sequence was generated by the technical assistant who was not involved in the study assessments. Allocation concealment was maintained through a sequentially numbered process, and the care provider responsible for briefing and tutorials remained unaware of group assignments until the moment of allocation.

Blinding

Due to the nature of the interventions, neither participants nor care providers could be blinded to group assignments. All outcome data were self-reported via online questionnaires, with no external outcome assessor involved. However, the data analyst was provided with anonymized group labels to minimize potential analysis bias. To reduce preconceived biases about the intervention of interest and comparators**,** the informed consent focused on the shared goal of diabetes improvement and participants' compensation, ensuring a neutral presentation of all study arms. No similarity of a placebo intervention was applicable here.

Statistical Methods

Given the repeated measurements of outcomes, a linear mixed model (LMM) was used. To investigate significant differences between groups, one-way ANOVA and paired-samples 2-tailed t tests were used. A correlation analysis was also used as an additional analysis. All statistical analyses were performed using SPSS Statistics (version 27; IBM Corp). The significance level was set at α=.05 for all statistical tests.

Handling of missing data: the variables BMI and HbA_{1c} contained no missing data, as BMI was derived from weekly weight measurements and HbA_{1c} was collected only at the pretest and posttest time points. The other 3 outcomes were constructed at the weekly level, as it is an established, accepted practice in mobile health and diabetes care studies. For FBS, physical activity, and food intake**,** missing values existed only in the daily records. According to the operational definitions of the study, these variables were aggregated to the weekly level using the mean or sum of the available daily entries. Weeks with partial daily logs were still computable, because weekly values can be calculated based on the days that were present rather than requiring a complete daily set. As a result of this aggregation process, the final analytic dataset contained no missing values for any of the weekly variables. No imputation procedure was used.

Reporting Guideline

The study was reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) 2010 statement. The preparation of the manuscript followed the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) reporting framework (Checklist 1) [34], and the abstract was structured according to the CONSORT guidance for abstracts [35], to ensure transparent and standardized reporting.

Ethical Considerations

Ethical approval was obtained from the ethical committee of Tehran University's psychology department, under the supervision of the Ministry of Health and Medical Education (MOHME), indexed as IR.UT.PSYEDU.REC.1404.106. All procedures were conducted in accordance with the ethical standards of the institutional research committee and the Declaration of Helsinki. Prior to participation, offline written informed consent was obtained from all participants or their parents. Participants were informed about the purpose of the study, trial procedures, and their right to withdraw at any time without any consequences. Participant privacy and confidentiality were strictly protected throughout the study. All collected data were deidentified and stored using coded identifiers, and no personally identifiable information was included in the analysis or reporting of the results. Access to the data was restricted to the research team and maintained in secure files in accordance with institutional ethical guidelines. No identification of individual participants in any image of the manuscript or supplementary material is possible. Participants received a small compensation to acknowledge the time and effort required to take part in the study. The type and amount of compensation were determined in accordance with the institutional ethical guidelines and were provided equally to all participants regardless of group allocation or study outcomes. No personal or health-related information was collected or transmitted through the VR application. In addition, XR safety procedures were implemented, including supervised use and monitoring for symptoms such as cybersickness, nausea, or any other bodily discomfort.

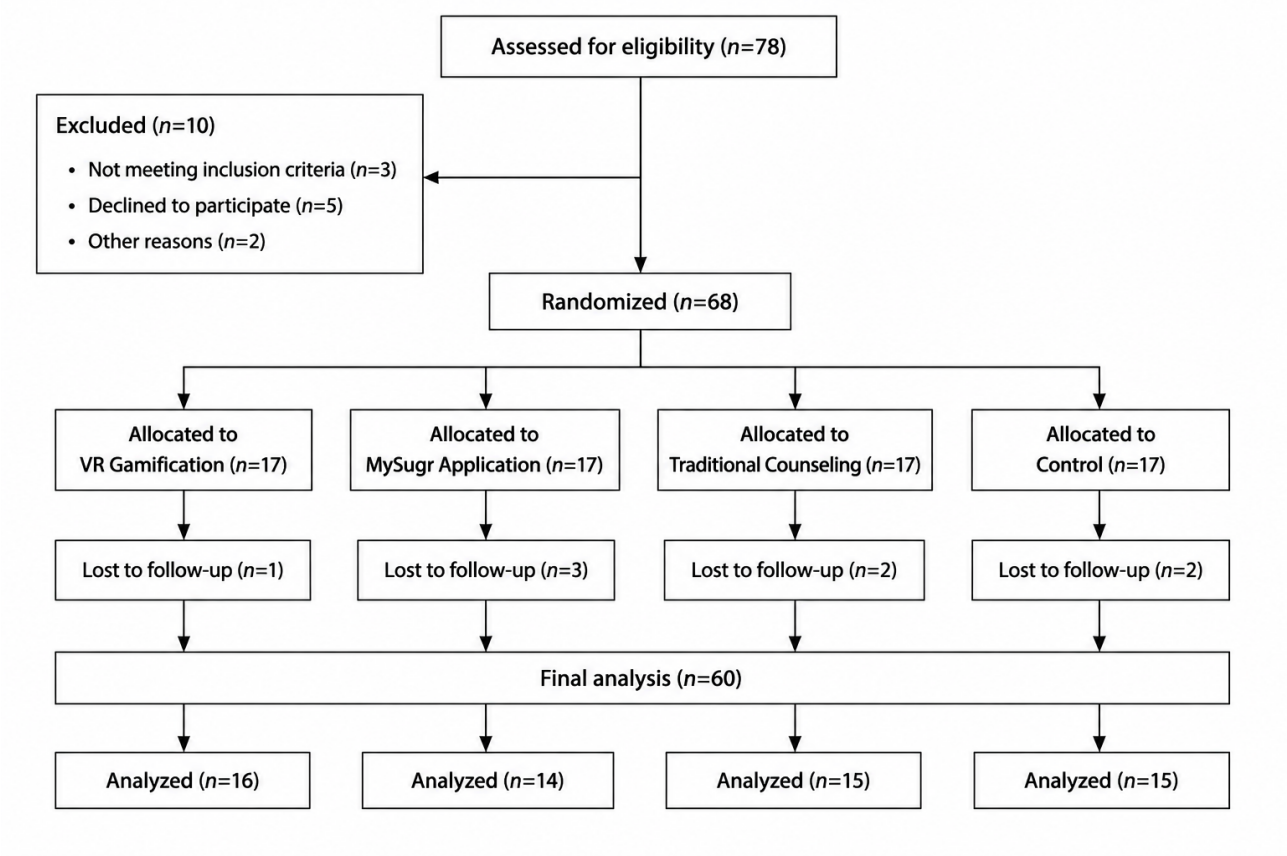
Results

Participant Flow

A total of 78 individuals were recruited and assessed for eligibility. After the initial briefing**,** 10 were excluded (3 did not fully meet the inclusion criteria, 5 declined to participate, and 2 for other reasons). A total of 68 participants were randomized using stratified block randomization into 4 parallel groups with equal-sized arms (n=17). During the 6-week study period, a total of 8 participants were lost to follow-up. Consequently, 60 participants were included in the final analysis. The detailed participant flow is illustrated in Figure 3.

XSL•FO
RenderX

Figure 3. Participant flow diagram of randomized controlled trial (RCT) program: "Improving diabetes health symptoms by using a specialized VR gamification self-care method: A parallel-group randomized controlled trial" according to CONSORT (Consolidated Standards of Reporting Trials) 2025. Eligibility assessment, enrollment, random allocation, follow-up status, and number of participants analyzed in each group. The study recruited children and adolescents diagnosed with diabetes in Tehran, Iran, from August 2024 to November 2024 and was conducted in December 2024-February 2025.



Recruitment

Participants were recruited and enrolled between August and November 2024. The intervention phase of the trial was conducted from December 2024 to February 2025, during which each participant completed a 6-week follow-up period. The trial concluded as planned without early termination.

Baseline Data

Table 1 presents the baseline demographic and clinical characteristics of participants in the 4 study groups prior to the intervention.

XSL•FO
RenderX

Table . Baseline characteristics of participants in each group. Baseline comparability can be interpreted. Gender is coded as 1=male and 2=female.

Characteristic	VR gamification (n=16)	MySugr app (n=14)	Traditional counseling (n=15)	Control (n=15)
Age (years)				
Mean (SD)	11.75 (2.05)	11.86 (1.66)	11.80 (1.74)	11.60 (1.84)
Range	7 - 15	9 - 15	9 - 15	8 - 15
Gender, mean (SD)	1.44 (0.51)	1.43 (0.51)	1.47 (0.52)	1.47 (0.52)
Diabetes type, mean (SD)	1.25 (0.44)	1.28 (0.46)	1.26 (0.45)	1.26 (0.45)
BMI (kg/m²)				
Mean (SD)	23.85 (5.55)	24.26 (6.62)	24.47 (6.74)	23.94 (6.18)
Range	18.3 - 36.2	16.2 - 36.2	16.5 - 35.3	18.8 - 37.1
FBS[a] (mg/dL)				
Mean (SD)	123.63 (18.95)	121.14 (18.19)	119.27 (20.22)	119.33 (18.23)
Range	93 - 163	83 - 155	81 - 158	84 - 152
HbA_{1c}[b] (%)				
Mean (SD)	7.94 (0.79)	7.66 (0.89)	8.13 (1.10)	7.51 (0.77)
Range	6.5 - 9.4	6.7 - 9.6	6.5 - 10.2	6.2 - 8.6

[a]FBS: fasting blood sugar.

[b]HbA_{1c}: glycated hemoglobin.

Numbers Analyzed

Data from 60 participants with complete outcome data were included in the final analysis. The denominator for all analyses corresponds to participants who completed the study and provided postintervention measurements. Analyses were conducted according to the originally assigned groups. The analyzed sample sizes were 16 in the VR gamified group, 14 in the MySugr app group, 15 in the traditional counseling group, and 15 in the control group.

Outcomes

The results of the LMM for all 5 outcomes are summarized in Table 2.

XSL•FO
RenderX

Table . Linear mixed model (LMM) results for trial outcomes. Model fit indices the Akaike information criterion (AIC), residual variance, and fixed-effects tests for time, condition, and their interaction are reported for all outcomes (fasting blood sugar [FBS], physical activity, food intake, BMI, and hemoglobin A_{1c} [HbA_{1c}]) in the parallel-group randomized controlled trial (RCT) evaluating a specialized virtual reality (VR)–gamification self-care program for children and adolescents diagnosed with diabetes. *P* values and effect sizes are reported, respectively.

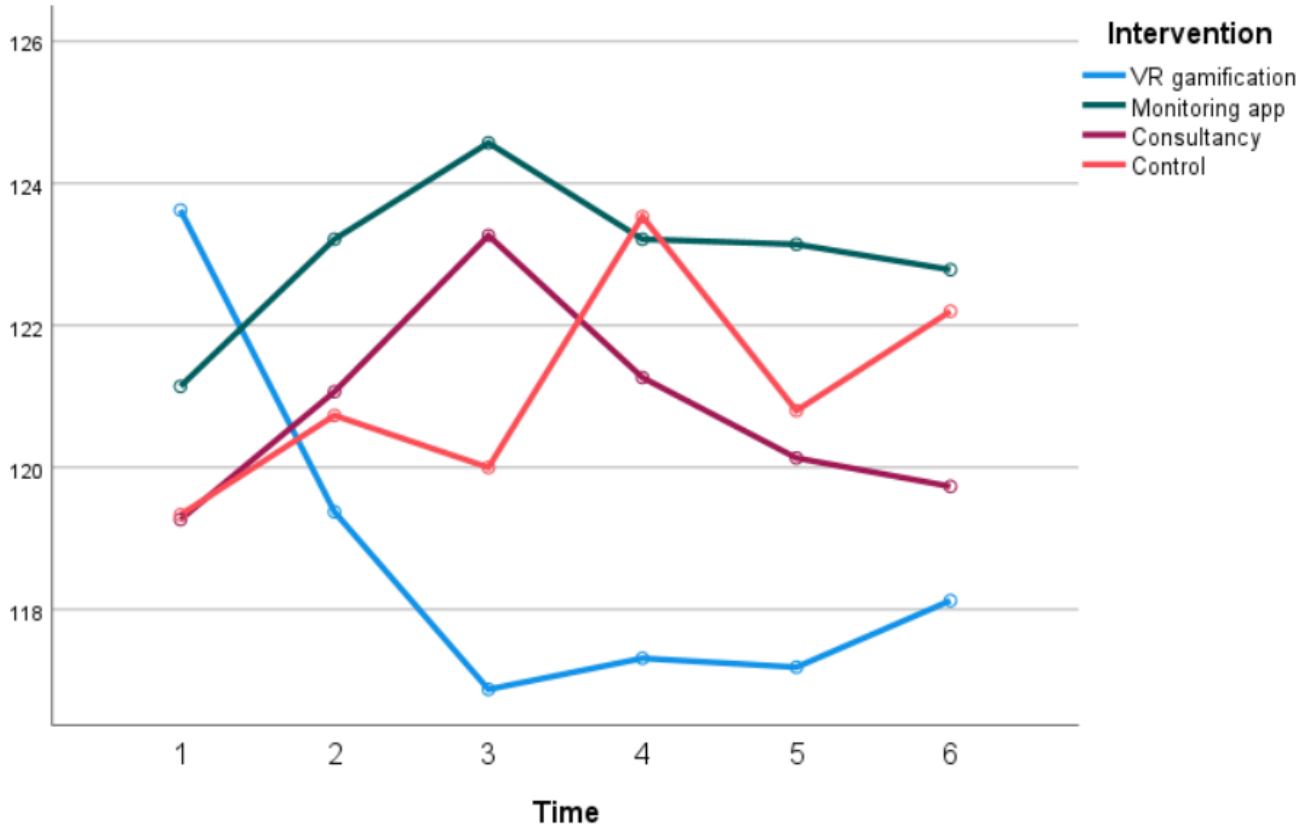
Outcome and source	*F* (df)	*P* value	Partial η^2	Model fit (AIC)	Residual variance
FBS				2383.29	29.05 (2.45)
Time	0.269 (5,280)	.93	.005		
Time×condition	2.046 (15,280)	.01	.099		
Condition	0.171 (3,39.91)	.91	.001		
Physical activity				1089.44	0.97 (0.08)
Time	1.580 (5,280.61)	.16	.027		
Time×condition	2.544 (15,280.61)	.003	.120		
Condition	14.399 (3.55.96)	<.001	.430		
Food intake				4138.80	4737.43 (400.39)
Time	3.487 (5,336)	.004	.059		
Time×condition	2.794 (15,336)	<.001	.111		
Condition	0.115 (3.336)	.95	.001		
BMI				429.61	0.03 (0.01)
Time	0.602 (5,280)	.69	.011		
Time×condition	0.660 (6.876,128.318)	.70	—		
Condition	0.021 (3,40.19)	.99	.034		
HbA_{1c}				306.57	0.71 (0.09)
Time	0.23 (1,112)	.63	.002		
Time×condition	0.064 (3,112)	.97	.002		
Condition	2.980 (3,112)	.03	.074		

FBS

An LMM was fitted with time, condition**,** and their interaction as fixed effects and a random intercept with autoregressive correlation of order 1 (AR1) covariance for participants. Model fit indices indicated adequate performance (Akaike information criterion [AIC]=2383.29), suggesting that adding time and the interaction improved overall model parsimony relative to the reduced model. The results of the LMM, as reported in Table 2, indicate that the time×condition interaction effect was statistically significant ($F_{15,280}$= 2.046; *P*=.01; partial η^2≈.099). However, no significant main effects were observed for time ($F_{5,280}$=0.269*;* *P*=.93; partial η^2≈.005) or condition ($F_{3,39.91}$=0.171; *P*=.91; partial η^2=.001), suggesting that overall changes over time and between-group differences were not statistically significant in isolation. Random-effects estimates indicated substantial between-subject variability (random-intercept variance=143.45, SE 29.92). The AR1 correlation parameter was negative (ρ=−.19), indicating weak inverse correlation between adjacent repeated measurements. Residual variance was 29.05 (SE 2.45). The temporal changes in the weekly average of FBS across the 4 experimental groups over 6 repeated measurements are illustrated in Figure 4. The significant interaction effect indicates that the pattern of FBS improvement over time varied by group. Specifically, participants in the VR gamified group showed a distinct trend in reducing FBS compared to other groups, with estimated marginal means decreasing from 123.625 mg/dL, 95% CI 114.147-133.103 at baseline to 118.125 mg/dL, 95% CI 110.387-125.863 at week 6. A paired-samples *t* test was conducted to examine the change in average FBS levels within the VR gamified group between the first and last sessions. The results revealed a significant decrease in FBS levels over time (t_{15}=−2.612; *P*=.02; mean difference=−5.500, 95% CI −1.012 to −9.987; Cohen *d*=0.653, 95% CI 0.102-1.186), indicating improved glycemic control among participants in this group, which is in line with recent studies [36]. This finding supports hypothesis 1.

XSL•FO
RenderX

Figure 4. Changes in fasting blood sugar (FBS) over time across intervention conditions. Trajectories show mean FBS values at 6 measurement points for the VR-Gamification, MySugr app, traditional counseling, and control group in the present randomized controlled trial study. Population: children and adolescents diagnosed with diabetes (age: 7 - 15 years), Tehran, Iran. Conducted in December 2024 to February 2025. VR: virtual reality.

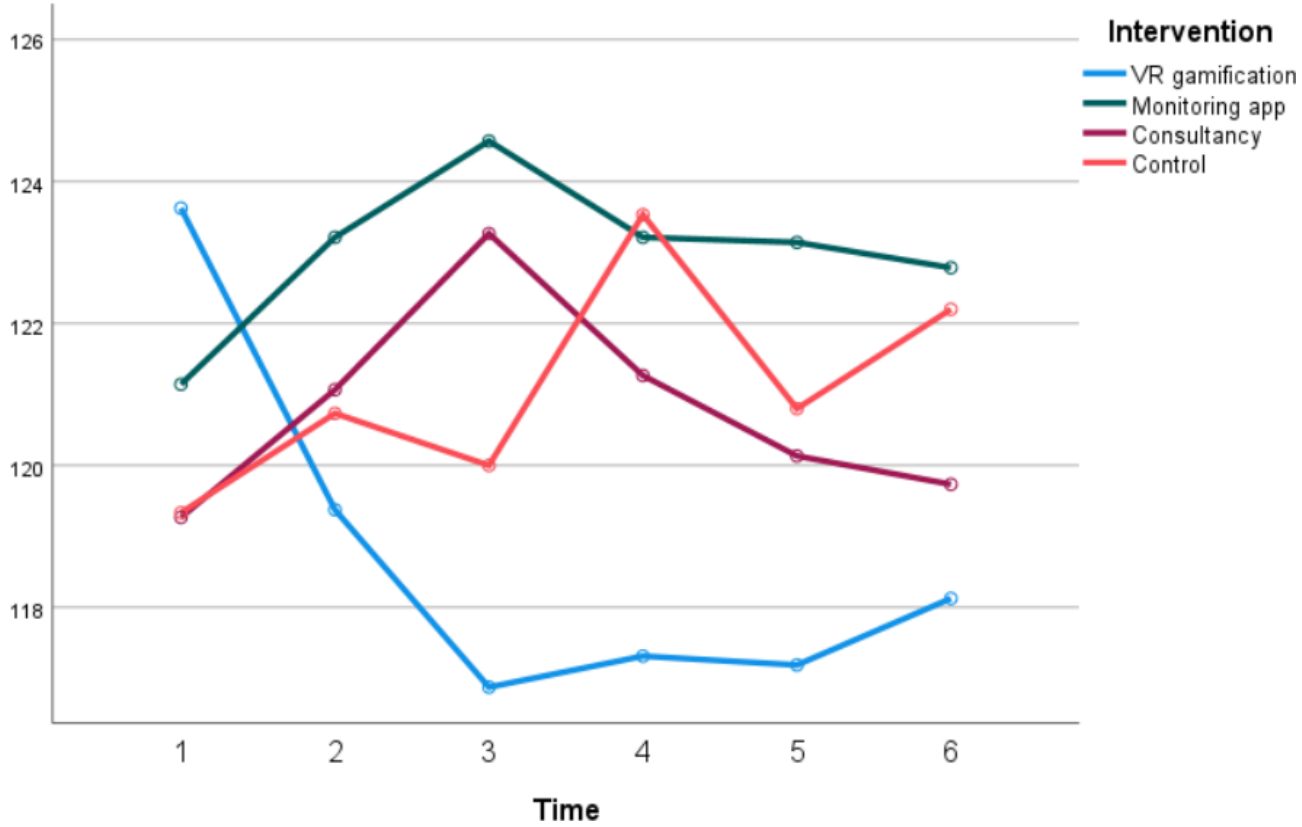


Physical Activity

An LMM was estimated with time**,** condition, and their interaction as fixed effects and a random intercept with an AR1 covariance structure for participants. Model fit indices (AIC=1089.44) indicated a relatively well-fitting model with adequate parsimony for the repeated-measures structure. According to Table 2, LMM results revealed a significant main effect of condition ($F_{3,55.96}$=14.399; P<.001; partial η^2≈.435) and a nonsignificant main effect of time ($F_{5,280.61}$=1.580; P=.16; partial η^2≈.027). The interaction effect of time×condition was significant ($F_{15,280.61}$= 2.544; P=.001; partial η^2≈.120), indicating that changes in physical activity over time differed significantly between the VR gamified group and the other experimental groups. Random-effects estimates showed modest between-subject variability (random-intercept variance=0.155; SE 0.0678). The AR1 correlation parameter was positive (ρ=0.656), indicating moderate positive correlation between adjacent repeated measurements. Residual variance was 0.979 (SE 0.083). Figure 5 illustrates the temporal changes in weekly physical activity across the 4 experimental groups. While physical activity levels remained relatively stable in the control, traditional, and app-based groups, the VR gamified group exhibited a noticeable upward trend over time. Thus, hypothesis 2 is supported.

XSL•FO
RenderX

Figure 5. Changes in physical activity (calculated as weekly sum of daily physical activity) over time across intervention conditions. One bar equaled either 30 minutes of vigorous activity (eg, gym workouts or swimming) or 60 minutes of moderate activity (eg, jogging). Trajectories show mean physical activity values at 6 measurement points for the VR-gamification, MySugr app, traditional counseling, and control group in the present randomized controlled trial study. Population: children and adolescents diagnosed with diabetes (age: 7 - 15 years), Tehran, Iran. Conducted in December 2024-February 2025. VR: virtual reality.



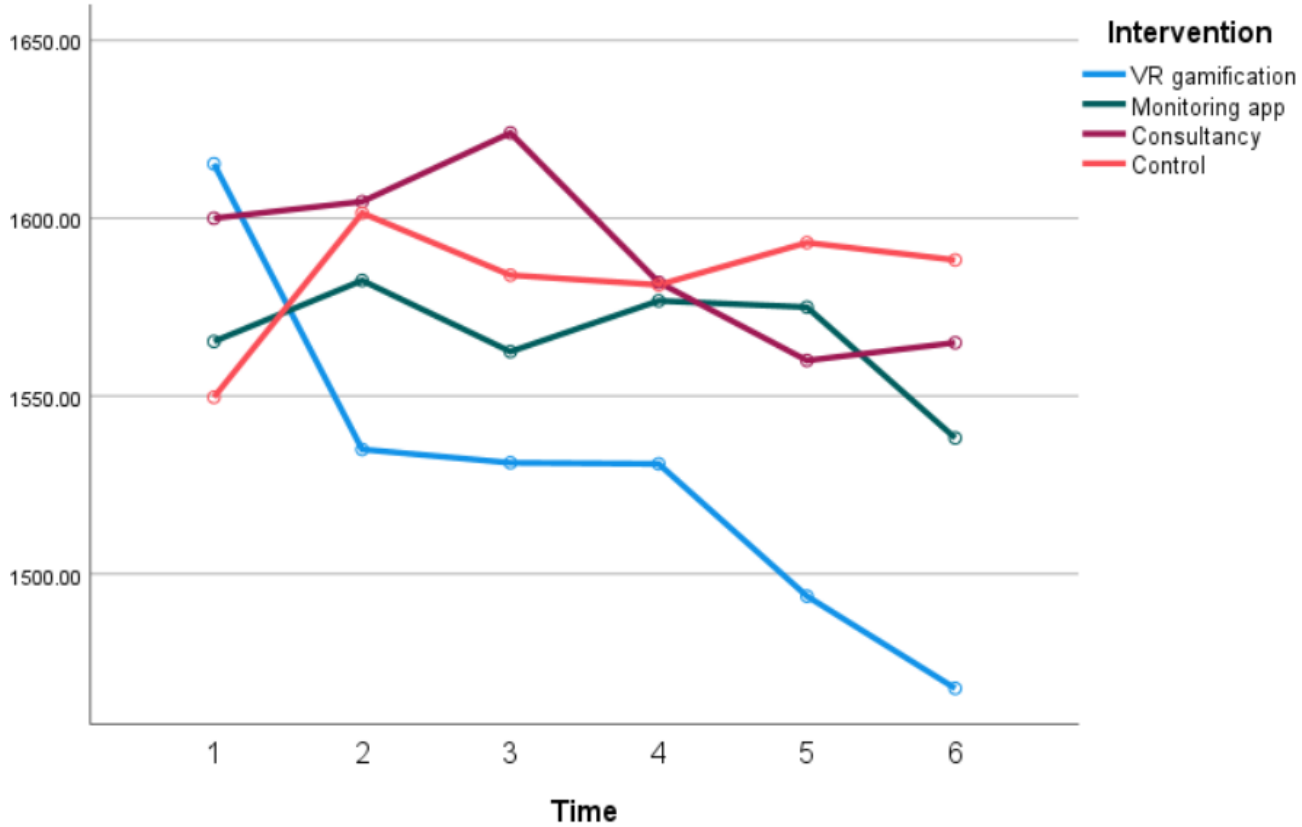
The one-way ANOVA analysis for physical activity during the last session revealed a significant difference in average physical activity between the VR gamified group and the other experimental groups ($F_{3,56}$=21.234; P<.001; partial η^2=.532). A paired-samples t test result showed improvement of physical activity in VR group between the first and last sessions (t_{15}=3.850; P=.002; mean difference=1.625, 95% CI 0.725-2.524; Cohen d=0.963, 95% CI 0.355-1.549). These findings align with recent meta-analytic evidence demonstrating a strong positive effect of exergames on physical activity levels [37].

Food Intake

An LMM was estimated with time**,** condition, and their interaction as fixed effects and a random intercept with an AR1 covariance structure for participants. Model-fit indices (AIC=4138.80) indicated acceptable fit for the repeated-measures design. According to the LMM results presented in Table 2, the main effect of condition was not significant ($F_{3,336}$=0.115; P=.95, partial η^2≈.001), while the main effect of time was significant ($F_{5,336}$=3.487; P=.004; partial η^2≈.059), indicating that the observed downward trend over time was unlikely due to chance. The time×condition interaction effect was significant ($F_{15,336}$=2.794; P<.001; partial η^2≈.111), suggesting that the pattern of change in food intake over time differed between the VR gamified group and the other groups. Random-effects estimates showed large between-subject variability (random-intercept variance=24,595.20,.20; SE 4685.39). The AR1 correlation parameter was reported as ρ=1.000; however, SPSS notes this value is redundant, typically indicating boundary estimation or insufficient information for precise AR1 correlation. Residual variance was 4737.43 (SE 400.39). Figure 6 illustrates the temporal changes in the weekly average calculated food intake across the 4 experimental groups. While food intake fluctuated irregularly in the control, traditional, and app-based groups, the VR gamified group showed a consistent downward trend. Thus, hypothesis 3 is supported.

XSL•FO
RenderX

Figure 6. Changes in food intake (calculated as weekly average of daily calories) over time across intervention conditions. Trajectories show mean daily calorie values at 6 measurement points for the virtual reality (VR)–gamification, MySugr app, traditional counseling, and control group in the present randomized controlled trial study. Population: children and adolescents diagnosed with diabetes (age: 7 - 15 years), Tehran, Iran. Conducted in December 2024-February 2025. VR: virtual reality.



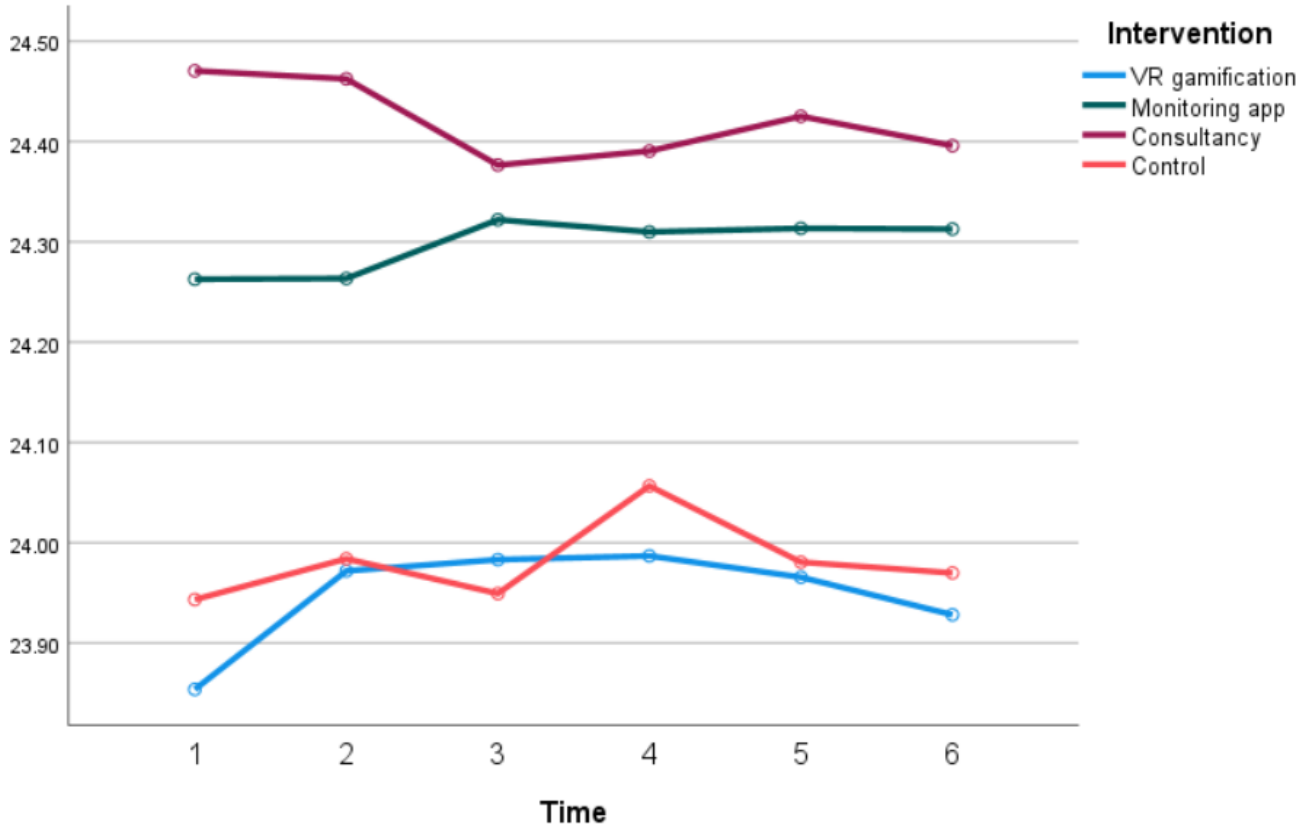
Results of the paired-samples *t* test for the food intake variable in the VR gamified group showed a significant decrease in average food intake from the first to the last session (t_{15}=−3.850; *P*=.002; mean difference=−147.500, 95% CI −0.725 to –2.524; Cohen *d*=0.963, 95% CI 0.355-1.549), which is in line with recent studies [38].

BMI

An LMM was fitted with time, condition**,** and their interaction as fixed effects and a random intercept with an AR1 covariance structure for participants. Model-fit indices (AIC=429.61) indicated that the model adequately captured the repeated-measures structure. According to Table 2, the LMM results for BMI showed no significant effects. For time effect ($F_{5,280}$=0.602; *P*=.69; partial η^2≈.011), for condition effect ($F_{3,40.19}$=0.021; *P*=.99; partial η^2≈.000), and for the time×condition interaction effect ($F_{15,280}$=0.660; *P*=.82; partial η^2=.034). Random-effects estimates indicated substantial between-subject variability (random-intercept variance=21.17; SE 4.32). The AR1 correlation parameter was slightly negative (ρ=–0.209), indicating weak inverse correlation between adjacent repeated BMI measurements. Residual variance was 0.0399 (SE 0.00337). Figure 7 illustrates that the weekly calculated BMI remains relatively stable and consistent across all 4 experimental groups throughout the study period. This suggests that the interventions did not have a significant impact on BMI changes over time. Therefore, hypothesis 4 is rejected.

XSL•FO
RenderX

Figure 7. Changes in BMI over time across intervention conditions. Trajectories show BMI mean values at 6 measurement points for the virtual reality (VR)–gamification, MySugr app, traditional counseling, and control group in the present randomized controlled trial study. Population: children and adolescents diagnosed with diabetes (age: 7 - 15 years), Tehran, Iran. Conducted in December 2024-February 2025. VR: virtual reality.

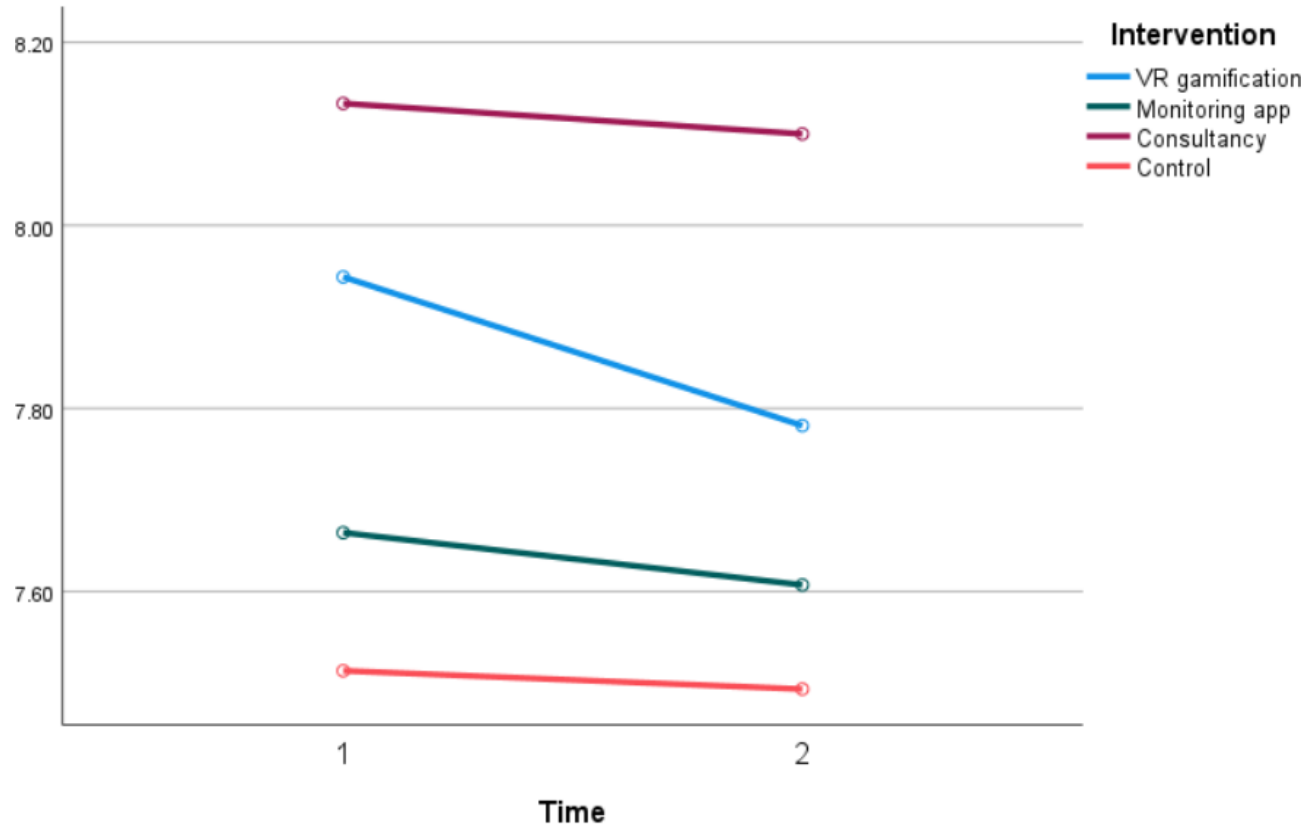


HbA_{1c}

An LMM was fitted with condition, time, and their interaction as fixed effects. A random intercept for participants was included, and the identity covariance structure was used. The model showed acceptable fit (AIC=306.57). The LMM results presented in Table 2 reveal significant effect for the condition effect ($F_{3,112}$=2.980; P=.03; partial η^2≈.074), the time effect ($F_{1,112}$=0.230; P=.63; partial η^2≈.002), and the time×condition interaction effect ($F_{3,112}$= 0.064; P=.97; partial η^2=0.002). Although the main effect of condition reached statistical significance in the mixed model, this effect reflects baseline differences between groups rather than any intervention-related change, as the related effect size is small (partial η^2≈.074). Therefore, no meaningful or clinically relevant effect of the intervention on HbA_{1c} can be inferred. As Figure 8 shows, the pattern of change between the 2 time points did not differ across conditions. Therefore, hypothesis 5 is rejected. Random intercept variance was very small (0.00098), reflecting minimal between-subject variability after accounting for fixed effects. Residual variance was 0.719 (SE 0.096).

XSL•FO
RenderX

Figure 8. Changes in glycated hemoglobin (HbA_{1c}) over time across intervention conditions. Trajectories show mean HbA_{1c} values at 2 measurement points: pretest and posttest for the virtual reality (VR)–gamification, MySugr app, traditional counseling, and control group in the present randomized controlled trial study. Population: children and adolescents diagnosed with diabetes (age: 7 - 15 years), Tehran, Iran. Conducted in December 2024-February 2025. VR: virtual reality.



Despite the strong correlation between HbA_{1c} and FBS (r_{58}=0.641; P<.001; 95% CI 0.469-0.770), our results showed no significant change in HbA_{1c} levels between the pretest and posttest. This finding contrasts with previous studies (eg, Kerfoot et al [39]), which reported significant changes in HbA_{1c} levels over a 12-month period (P=.048). The discrepancy may be attributed to the relatively short duration of the current intervention.

Ancillary Analysis

One exploratory analysis was performed, consisting of a correlation analysis to examine associations between behavioral and metabolic outcomes. As Table 3 shows, significant associations were observed between changes in FBS and physical activity (r_{58}=–0.422, 95% CI –0.610 to –0.188; P<.001) and between changes in physical activity and food intake (r_{58}= –0.292, 95% CI –0.509– to 0.041; P=.02), indicating linked behavioral and metabolic shifts. This pattern aligns with prior findings on the relationship between physical activity and blood glucose levels [40].

Table . Pearson correlations among change scores for trial outcomes. Correlation coefficients show the associations between changes in fasting blood sugar (FBS), physical activity, and food intake across 60 participants of the present parallel group randomized controlled trial study. Population: children and adolescents diagnosed with diabetes (age: 7 - 15 years), Tehran, Iran. Conducted in December 2024-February 2025.

Variable	Changes in FBS[a]	Changes in physical activity	Changes in food intake
Changes in FBS	1	–0.422	0.082
Changes in physical activity	–0.422	1	–0.292
Changes in food intake	0.082	–0.292	1

[a]FBS: fasting blood sugar.

Harms

All XR sessions were completed without any adverse events or early terminations. No instances of cybersickness, visual discomfort, dizziness, or excessive fatigue were reported, and no participant required a session to be stopped prematurely. Adherence to the intervention was exceptionally high, with all children attending and completing all scheduled sessions. Participants in the VR group (n=16) demonstrated high perceived usability, with a mean System Usability Scale score of 74.2 (SD 7.9). According to established usability benchmarks, this value falls within the Excellent range, indicating that the VR environment was experienced as intuitive, consistent, and

XSL•FO
RenderX

easy to navigate. No individual item showed unusually low ratings, suggesting that participants were able to interact with the system with minimal cognitive effort or technical difficulty. TAM responses further supported strong user acceptance of the VR system. Perceived usefulness yielded a mean score of 5.3/7 (SD 0.7), while perceived ease of use was rated 5.6/7 (SD 0.6). These values align with a moderately good acceptance profile, reflecting participants' perceptions that the VR platform was beneficial, motivational, and easy to learn.

Discussion

Principal Findings

The present study aimed to evaluate the effectiveness of a VR gamified program in improving metabolic indicators and health-related behaviors among children and adolescents diagnosed with diabetes. The intervention was compared with 3 comparators, a gamified mobile app (MySugr), traditional consultation, and standard medical care as a control group, in a longitudinal 6-week parallel-groups RCT design. Overall, the results indicate that the VR gamified intervention produced the most consistent improvements in behavioral outcomes, including physical activity and dietary control, compared to the other 3 parallel groups. Moderate yet promising improvements in FBS were also observed in the VR gamified group relative to the other intervention groups, suggesting that the VR gamified method can potentially yield an improved level of short-term glycemic control compared to other methods. In contrast, the intervention did not produce significant improvements in long-term metabolic indicators, including BMI or HbA_{1c}, in the VR gamified group compared to other interventions.

The behavioral outcomes including physical activity and dietary control are consistent with recent studies, indicating that interactive gamified environments can enhance the development of sustained health behaviors [41]. The significant improvement in short-term glycemic control (FBS) alongside a lack of significant improvements in long-term metabolic control markers, including HbA_{1c} and BMI, are in line with recent studies, as long-term metabolic indicators typically require sustained behavioral modification over longer periods to demonstrate measurable changes [42]. The overall improvement trends observed in the VR-gamified group highlight the importance of emotional engagement in self-care interventions. This finding is consistent with previous studies emphasizing that interactive gameplay and immersive audio-visual experiences are critical factors in the success of health gamification systems [43]. In contrast to conventional health-tracking programs that rely primarily on point/badge reward structures, the VR gamified intervention used in this study integrated immersive 3D environments with action-based gameplay that directly represented users' health symptoms and behaviors. Mentioned features may explain the stronger behavioral engagement observed in the VR gamified group and may also clarify why traditional point/badge gamification approaches may produce weaker effects, as reported in some studies [44].

Applications and Limitations

Real-world implementation of this VR intervention could align with existing pediatric diabetes education pathways in which families and educators work together to support self-care. In practice, parents can help supervise home-based sessions, while school health staff may provide complementary reinforcement during the day. Follow-up by diabetes educators or community nurses could be incorporated into routine clinical visits or remote check-ins, enabling the VR experience to serve as an adjunct to ongoing education rather than an isolated tool [45]. However, equitable access to VR remains a challenge. Variability in the availability of VR devices and limited technological resources in some families or schools may restrict adoption. To mitigate such barriers, future implementations may require shared devices in clinical settings or mobile-based alternatives to ensure broader accessibility. Due to the experimental design of this study, the external validity and generalizability of the findings should be interpreted with caution.

Several limitations should be considered when interpreting the findings of this study. First, the relatively small sample size (n=60) may limit the generalizability of the results. Second, the intervention period was relatively short (6 weeks), which may not have been sufficient to observe meaningful changes in long-term metabolic indicators, including BMI and HbA_{1c}. Third, outcomes were measured through self-reported data, which may have caused reporting bias. Future research should therefore incorporate larger samples, longer intervention durations, and objective behavioral tracking tools to provide more robust evidence regarding the effectiveness of VR gamified self-care interventions.

The VR gamified experience developed in this study appears to offer a promising complement or alternative to existing self-care methods, with preliminary indications of potentially enhanced engagement and outcomes. However, interpretations of the behavioral findings should be calibrated to the limitations of measurements, including reporting bias, limited sample size, short duration of the intervention, variations in adherence and fidelity, and the novelty effect of the VR game. Results should be understood as preliminary behavioral signals rather than evidence of clinical improvements. While the program effectively fostered behavioral engagement and adherence trends, no clinically causal inference can be drawn. Future studies with new game mechanics, more engaging designs, larger samples, longer follow-ups, and objective tracking tools can scale the usability of the present intervention and extend related findings.

Conclusion

This study combined gamification elements with immersive VR features to develop a novel self-care system for children and adolescents diagnosed with diabetes. The intervention was grounded in the behavioral economics framework, specifically the theoretical sequence of reward, motivation, behavior, and habit formation, which is central to modern gamification approaches. By offering engaging gameplay within an immersive VR environment, the game established a direct connection between real-life diabetes and in-game mechanics. The goal was to foster intrinsic motivation through in-game

XSL•FO
RenderX

rewards, thereby encouraging users to sustain desirable self-care behaviors. Findings support the potential effectiveness of VR gamified intervention, highlighting its power to serve as a complementary or alternative digital tool for enhancing diabetes self-management. However, clinical implications should be interpreted cautiously.

Acknowledgments

We sincerely thank Dr Aria Sotoudeh, Ms Alma Roshani, Dr Hamidreza Sabour, and Dr Mohsen Kazemi Moghaddam for their invaluable assistance in participant recruitment. We are also deeply grateful to the Children's Medical Center (Tehran University of Medical Sciences), Iran Diabetes Center, and the Gabric Diabetes Center for their essential support in study preparation and providing operational insights during the implementation phase of this project. Furthermore, we sincerely appreciate Mr Vahid Masoumi and Ms Roghina Pourmoghaddam, experienced videogame makers in Iran, for their expert consultations and valuable advice on our virtual reality game design. The study was registered in IRTC (Iranian registry of clinical trials) with Trial ID=89347, Registration date: 2026-04-11, and IRCT registration number: IRCT20260317068997N1. The full registration is provided at: https://www.irct.ir/trial/89347. Generative AI tools were used in a limited and supportive role during the preparation of this manuscript. Specifically, ChatGPT models (GPT-4o and GPT-5.2, OpenAI) were used to assist with grammatical editing, step-by-step guidance for conducting statistical analyses in SPSS, identifying potentially relevant academic papers during the literature search process, and adapting the manuscript to the reporting structures and formatting requirements of the target journal (JMIR) and reporting guidelines such as CONSORT, including reference formatting and structural organization. All scientific content, interpretations, analyses, and conclusions were determined by the authors. Any AI-generated text or suggestions were critically reviewed, edited, and, when necessary, rewritten by the authors to ensure accuracy, originality, and alignment with the authors' intended meaning. The authors take full responsibility for the final content of the manuscript.

The study protocol was developed as part of the first author's (RS) doctoral dissertation of Health Psychology at Tehran University, which is fully documented and accessible. While the designed VR game is available through OSF framework, the source codes and technical tools used during the interventions are provided as supplementary materials to this article.

Funding

This research received no external funding.

Data Availability

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

Authors' Contributions

RS, JH, and RP performed conceptualization, project administration, and methodology. RS and AH performed investigation. AH and AP performed validation and supervision. HBE and RP performed supervision and data curation. RS managed software. HBE conducted formal analysis. RS wrote the original draft. RS, AH, and JH reviewed and edited the manuscript.

Conflicts of Interest

No potential conflicts of interest relevant to this article were reported. The development team of the virtual reality VR game overlapped substantially with the research team conducting the study. The VR game evaluated in this trial was designed and developed mainly by the first author: RS, who was also involved in the trial implementation and manuscript writing. This relationship is disclosed to ensure transparency regarding the study team's connection to the evaluated intervention.

Multimedia Appendix 1
The game source codes.
[RAR File, 9 KB - xr_v3i1e81402_app1.rar]

Multimedia Appendix 2
Diabetes markers data collection form.
[DOCX File, 17 KB - xr_v3i1e81402_app2.docx]

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 1112 KB - xr_v3i1e81402_app3.pdf]

References

XSL•FO
RenderX

1. Puzhakkal S, Kavanagh S, Conway B, Kow CS, Hasan SS. Self-care and self-management in diabetes: concepts, theories and practices. Int J Clin Pharm 2025 Dec;47(6):2094-2100. [doi: 10.1007/s11096-025-01941-z] [Medline: 40489012]
2. Klaassen R, Bul KCM, Op den Akker R, van der Burg GJ, Kato PM, Di Bitonto P. Design and evaluation of a pervasive coaching and gamification platform for young diabetes patients. Sensors (Basel) 2018 Jan 30;18(2):402. [doi: 10.3390/s18020402] [Medline: 29385750]
3. Fontecha J, González I, Barragán A, Lim T. Use and trends of diabetes self-management technologies: a correlation-based study. J Diabetes Res 2022;2022(1):5962001. [doi: 10.1155/2022/5962001] [Medline: 35712029]
4. Asadzandi S, Sedghi S, Bigdeli S, Sanjari M. A systematized review on diabetes gamification. Med J Islam Repub Iran 2020;34(168):168. [doi: 10.47176/mjiri.34.168] [Medline: 33816367]
5. Sze WT, Waki K, Enomoto S, et al. StepAdd: a personalized mHealth intervention based on social cognitive theory to increase physical activity among type 2 diabetes patients. J Biomed Inform 2023 Sep;145:104481. [doi: 10.1016/j.jbi.2023.104481] [Medline: 37648101]
6. Brady VJ, Mathew Joseph N, Ju HH. Impact of gaming (gamification) on diabetes self-care behaviors and glycemic outcomes among adults with type 2 diabetes. Sci Diabetes Self Manag Care 2023 Dec;49(6):493-511. [doi: 10.1177/26350106231208153] [Medline: 37927059]
7. Veazie S, Winchell K, Gilbert J, Paynter R, Orleans T. Mobile Applications for Self-Management of Diabetes: A Systematic Review: Agency for Healthcare Research and Quality (US); 2018:1-187 URL: https://www.ncbi.nlm.nih.gov/sites/books/NBK518944/ [accessed 2026-05-28]
8. Tuah NM, Yoag A, Ahmedy F. MyDiabetes—the gamified application for diabetes self-management and care. Computers 2021;10(4):50. [doi: 10.3390/computers10040050]
9. Akker RO, Klaassen R, Bul K, Kato PM, Burg GJ, Bitonto P. Let them play: experiences in the wild with a gamification and coaching system for young diabetes patients. Presented at: In Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare; May 23-26, 2017; Barcelona, Spain p. 409-418. [doi: 10.1145/3154862.3154931]
10. King D, Greaves F, Exeter C, Darzi A. “Gamification”: influencing health behaviours with games. J R Soc Med 2013 Mar;106(3):76-78. [doi: 10.1177/0141076813480996] [Medline: 23481424]
11. Lansink C, Middelweerd A, Laverman GD, Jong NB, Velde SJ. Barriers to and facilitators of user engagement with digital self-monitoring tools for diabetes: qualitative study. Implement Sci Commun 2025;6(1). [doi: 10.1186/s43058-026-00904-3] [Medline: 41857778]
12. Höchsmann C, Infanger D, Klenk C, Königstein K, Walz SP, Schmidt-Trucksäss A. Effectiveness of a behavior change technique-based smartphone game to improve intrinsic motivation and physical activity adherence in patients with type 2 diabetes: randomized controlled trial. JMIR Serious Games 2019 Feb 13;7(1):e11444. [doi: 10.2196/11444] [Medline: 30758293]
13. Priesterroth L, Grammes J, Holtz K, Reinwarth A, Kubiak T. Gamification and behavior change techniques in diabetes self-management apps. J Diabetes Sci Technol 2019 Sep;13(5):954-958. [doi: 10.1177/1932296818822998] [Medline: 30762420]
14. Holmen H, Wahl AK, Cvancarova Småstuen M, Ribu L. Tailored communication within mobile apps for diabetes self-management: a systematic review. J Med Internet Res 2017 Jun 23;19(6):e227. [doi: 10.2196/jmir.7045] [Medline: 28645890]
15. 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] [Medline: 34982715]
16. AlMarshedi A, Wills G, Ranchhod A. Guidelines for the Gamification of Self-Management of Chronic Illnesses: Multimethod Study. JMIR Serious Games 2017 May 12;5(2):e12. [doi: 10.2196/games.7472] [Medline: 28500018]
17. Fortunato M, Harrison J, Oon AL, et al. Remotely monitored gamification and social incentives to improve glycemic control among adults with uncontrolled type 2 diabetes (iDiabetes): protocol for a randomized controlled trial. JMIR Res Protoc 2019 Nov 20;8(11):e14180. [doi: 10.2196/14180] [Medline: 31746765]
18. Peng W, Lin JH, Crouse J. Is playing exergames really exercising? A meta-analysis of energy expenditure in active video games. Cyberpsychol Behav Soc Netw 2011 Nov;14(11):681-688. [doi: 10.1089/cyber.2010.0578] [Medline: 21668370]
19. Johnson D, Deterding CS, Kuhn KA, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106. [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
20. Ghodousi Moghadam S, Mazloum Khorasani Z, Sharifzadeh N, Tabesh H. A mobile serious game about diabetes self-management: design and evaluation. Heliyon 2024 Sep 30;10(18):e37755. [doi: 10.1016/j.heliyon.2024.e37755] [Medline: 39364243]
21. 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] [Medline: 28536062]
22. Sharma D, Sharma J, Mehta N. The art of gamification for societal mind shifts: refining human behaviour through strategic engagement. ShodhKosh J Vis Per Arts 2023;5(1):621-XXX [FREE Full text] [doi: 10.29121/shodhkosh.v5.i1.2024.855]

XSL•FO
RenderX

23. Lønne TF, Karlsen HR, Langvik E, Saksvik-Lehouillier I. The effect of immersion on sense of presence and affect when experiencing an educational scenario in virtual reality: a randomized controlled study. Heliyon 2023 Jun;9(6):e17196. [doi: 10.1016/j.heliyon.2023.e17196] [Medline: 37360072]
24. Jo H, Park DH. Affordance, usefulness, enjoyment, and aesthetics in sustaining virtual reality engagement. Sci Rep 2023 Sep 12;13(1):15097. [doi: 10.1038/s41598-023-42113-1] [Medline: 37699934]
25. Liao CH. Exploring perceived learning effectiveness in virtual reality health communication through the lens of construal level theory. BMC Public Health 2024 Aug 29;24(1):2348. [doi: 10.1186/s12889-024-19827-6] [Medline: 39210293]
26. Vaughan N. Virtual reality meets diabetes. J Diabetes Sci Technol 2025 May;19(3):810-819. [doi: 10.1177/19322968231222022] [Medline: 38193465]
27. ElSayed NA, Aleppo G, Aroda VR, et al. Erratum. 2. Classification and diagnosis of diabetes: Standards of Care in Diabetes-2023. Diabetes Care 2023;46(Suppl. 1):S19-S40. Diabetes Care 2023;46(Supplement_1):S19-S40. [doi: 10.2337/dc23-er05]
28. Ayele K, Tesfa B, Abebe L, Tilahun T, Girma E. Self care behavior among patients with diabetes in Harari, Eastern Ethiopia: the health belief model perspective. PLoS ONE 2012;7(4):e35515. [doi: 10.1371/journal.pone.0035515] [Medline: 22530039]
29. OSF storage files. Open Science Framework. URL: https://osf.io/kn6z2/files/osfstorage [accessed 2026-05-29]
30. OSF storage files. Open Science Framework. URL: https://osf.io/bydrt/files/osfstorage [accessed 2026-05-28]
31. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability Evaluation in Industry: Taylor and Francis; 1996:189-194 URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke [accessed 2026-05-28]
32. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
33. Food calorie table. Kermany. URL: https://kermany.com/food-calorie-table [accessed 2026-05-28]
34. Eysenbach G, Group CE. 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] [Medline: 22209829]
35. Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration. PLoS Med 2008 Jan 22;5(1):e20. [doi: 10.1371/journal.pmed.0050020] [Medline: 18215107]
36. Liu K, Xie Z, Or CK. Effectiveness of mobile app-assisted self-care interventions for improving patient outcomes in type 2 diabetes and/or hypertension: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2020 Aug 4;8(8):e15779. [doi: 10.2196/15779] [Medline: 32459654]
37. Zhao M, Lu X, Zhang Q, et al. Effects of exergames on student physical education learning in the context of the artificial intelligence era: a meta-analysis. Sci Rep 2024 Mar 26;14(1):7115. [doi: 10.1038/s41598-024-57357-8] [Medline: 38531948]
38. 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 Jul;73(6):2651-2659. [doi: 10.1080/07448481.2024.2325929] [Medline: 38546660]
39. Kerfoot BP, Gagnon DR, McMahon GT, Orlander JD, Kurgansky KE, Conlin PR. A team-based online game improves blood glucose control in veterans with type 2 diabetes: a randomized controlled trial. Diabetes Care 2017 Sep;40(9):1218-1225. [doi: 10.2337/dc17-0310] [Medline: 28790131]
40. Purwandari A, Suparyatmo JB, Sugiarto S. Relationship between BMI, physical activity, zinc serum with fasting blood glucose in diabetes. 2019 Presented at: Proceedings of the 1st Seminar and Workshop on Research Design, for Education, Social Science, Arts, and Humanities, SEWORD FRESSH 2019; Apr 27-28, 2019 URL: http://eudl.eu/proceedings/SEWORD FRESSH/2019 [accessed 2026-5-29] [doi: 10.4108/eai.27-4-2019.2286799]
41. Sun Y, Xing J. The impact of gamification motivation on green consumption behavior—an empirical study based on ant forest. Sustainability 2023;15(1):512. [doi: 10.3390/su15010512]
42. König D, Kookhan S, Schaffner D, Deibert P, Berg A. A meal replacement regimen improves blood glucose levels in prediabetic healthy individuals with impaired fasting glucose. Nutrition 2014;30(11-12):1306-1309. [doi: 10.1016/j.nut.2014.03.014] [Medline: 25194963]
43. Nand K, Baghaei N, Casey J, Barmada B, Mehdipour F, Liang HN. Engaging children with educational content via gamification. Smart Learn Environ 2019 Dec;6(1). [doi: 10.1186/s40561-019-0085-2]
44. 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(12):e39019. [doi: 10.2196/39019] [Medline: 38133907]
45. Timpel P, Oswald S, Schwarz PEH, Harst L. Mapping the evidence on the effectiveness of telemedicine interventions in diabetes, dyslipidemia, and hypertension: an umbrella review of systematic reviews and meta-analyses. J Med Internet Res 2020 Mar 18;22(3):e16791. [doi: 10.2196/16791] [Medline: 32186516]

Abbreviations

AIC: Akaike information criterion

XSL•FO
RenderX

AR1: autoregressive correlation of order 1
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
DM: diabetes mellitus
FBS: fasting blood sugar
HbA_{1c}: glycated hemoglobin
LMM: linear mixed model
MOHME: Ministry of Health and Medical Education
OSF: Open Science Framework
TAM: Technology Acceptance Model
VR: virtual reality
XR: extended reality

Edited by I Steenstra, T Leung; submitted 28.Jul.2025; peer-reviewed by DG Martinez, P Pichayapinyo; revised version received 11.Apr.2026; accepted 30.Apr.2026; published 10.Jun.2026.

Please cite as:
Solat R, Pourhosein R, Bahrami Ehsan H, Pourabbasi A, Hashemi A, Hatami J
Improving Health-Related Symptoms and Behaviors in Children and Adolescents Diagnosed With Diabetes by Using a Virtual Reality–Gamified Self-Care Method: Randomized Controlled Trial
JMIR XR Spatial Comput 2026;3:e81402
URL: https://xr.jmir.org/2026/1/e81402
doi:10.2196/81402

© Ruhollah Solat, Reza Pourhosein, Hadi Bahrami Ehsan, Ata Pourabbasi, Amin Hashemi, Javad Hatami. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 10.Jun.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Mixed Reality–Based Slit Lamp for Ophthalmic Examination and Telemedicine: Technological Development and Validation Study

Rui Zhou[1], MD; Wei-Chiang Lin[1], PhD; Byron L Lam[2], MD; Rong Wen[2], PhD; Noble Amadi[1], MS; Shuliang Jiao[1], PhD

[1]Department of Biomedical Engineering, Florida International University, 10555 W Flagler St, EC-2610, Miami, FL, United States

[2]Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, FL, United States

Corresponding Author:
Shuliang Jiao, PhD
Department of Biomedical Engineering, Florida International University, 10555 W Flagler St, EC-2610, Miami, FL, United States

Abstract

Background: The slit-lamp biomicroscope is a fundamental diagnostic tool in ophthalmology for detailed examination of the eye. Current camera-equipped digital slit lamps were designed with a single optical channel, which results in the loss of depth information. Without that information, it can be challenging to visualize subtle anatomical variations in teleophthalmology applications and perform procedures guided by the digital view.

Objective: This study aimed to present a feasibility study of the mixed reality (MR)–based slit lamp (MR-SLP) capable of transmitting real-time stereoscopic views of the slit lamp to local and remote MR headsets, enabling stereoscopic teleophthalmology.

Methods: A prototype MR-SLP was built by integrating a calibrated stereoscopic camera pair on the left and right viewing channels of a conventional slit lamp and a real-time streaming network. The stereoscopic diagnostic images were transmitted to multiple MR headsets through the streaming network at 1080p and 30 frames per second (fps). The spatial resolution of the system was quantified using a US Air Force 1951 resolution target (Edmund Optics Inc). The 3D spatial accuracy and coordination were evaluated quantitatively by performing a tube-threading test. Five participants (mean age 42.2, SD 16.5 years) with normal visual function, best-corrected visual acuity of 20/20 or better, and a minimum stereoacuity of approximately 40 arc seconds participated in the tube-threading test. Teleophthalmology capability was assessed through real-time streaming across multiple remote sites at Florida International University and Bascom Palmer Eye Institute.

Results: The measured spatial resolution reached 102 line pairs/mm at 25× optical magnification. The tube-threading task was performed under 4 conditions. Task performance differed significantly between nonstereoscopic (2D) and direct eyepiece views (P=.03, Kruskal-Wallis test), but not between stereoscopic (3D) MR and direct views (P>.05, Kruskal-Wallis test). In the real-time remote streaming tests across multiple sites, the system achieved stable, low-latency transmission with an average round-trip time below 40 milliseconds. Participating ophthalmologists reported user experience and image quality comparable to traditional slit lamps.

Conclusions: The MR-SLP can provide real-time stereoscopic slit-lamp examination images and videos through a broadcasting network to local and remote locations. The spatial resolution and visuomotor performance are comparable to direct viewing through the eyepieces of a traditional slit lamp. This study demonstrated the feasibility of the MR-SLP for high-quality stereoscopic teleophthalmology.

(JMIR XR Spatial Comput 2026;3:e93513) doi:10.2196/93513

KEYWORDS

telemedicine; teleophthalmology; mixed reality; slit-lamp microscopy; depth perception; wearable; remote consultation

Introduction

The slit-lamp biomicroscope is a fundamental diagnostic tool in ophthalmology for detailed examination of both the anterior and posterior segments of the eye [1] with stereoscopic visualization [2]. Traditional slit lamps require the operator to view through the eyepieces to obtain a stereoscopic perspective. This necessity means the operator must stay close to the patient, which has become more concerning since the COVID-19 pandemic, given that close contact increases the risk of transmitting infectious diseases [3,4]. Modern slit lamps frequently come with a camera, whether dedicated or integrated into a smartphone, allowing them to capture images and videos of the eye during examinations. This feature is especially beneficial for data collection, telemedicine, and education, as captured images and videos can be viewed on a display, stored electronically, and shared remotely [5-7]. Current camera-equipped slit lamps often use a single-channel design, resulting in the loss of stereoscopic information and,

XSL•FO
RenderX

consequently, depth perception during image and video capture. Without depth perception, it can be challenging to visualize subtle anatomical variations and perform procedures guided by the slit-lamp view [8].

Mixed reality (MR) technology enables interactive experiences that blend both real and virtual environments by using a head-mounted display (HMD). A HMD is a wearable device designed to fit comfortably over the user's head, equipped with a stereo camera and individual screens for each eye. Additional sensors are incorporated to accurately track head and eye movements. The MR headset's see-through function allows users to view the real world with depth perception. By seamlessly blending digital content with the real world, MR facilitates intuitive interaction and manipulation of virtual objects within the user's actual environment. These advancements make MR a promising tool for applications such as 3D medical data visualization and telemedicine [9,10].

To address the limitations of conventional slit-lamp examinations in remote collaboration and stereoscopic information sharing, we developed the MR-based slit lamp (MR-SLP) that integrates MR technology with a traditional slit lamp. The MR-SLP enables local operators to interact naturally with patients using its see-through function while simultaneously visualizing a stereoscopic diagnostic view; it also allows real-time streaming of the view to remote physicians wearing headsets for immersive participation. This functionality supports collaborative decision-making, as remote specialists can actively guide local operators during examinations, and enables recording of stereoscopic diagnostic images and videos for documentation and education. In this study, we aim to develop and evaluate the MR-SLP system and assess its capability for real-time stereoscopic imaging, remote visualization, and collaborative ophthalmic examination and telemedicine.

Methods

Ethical Considerations

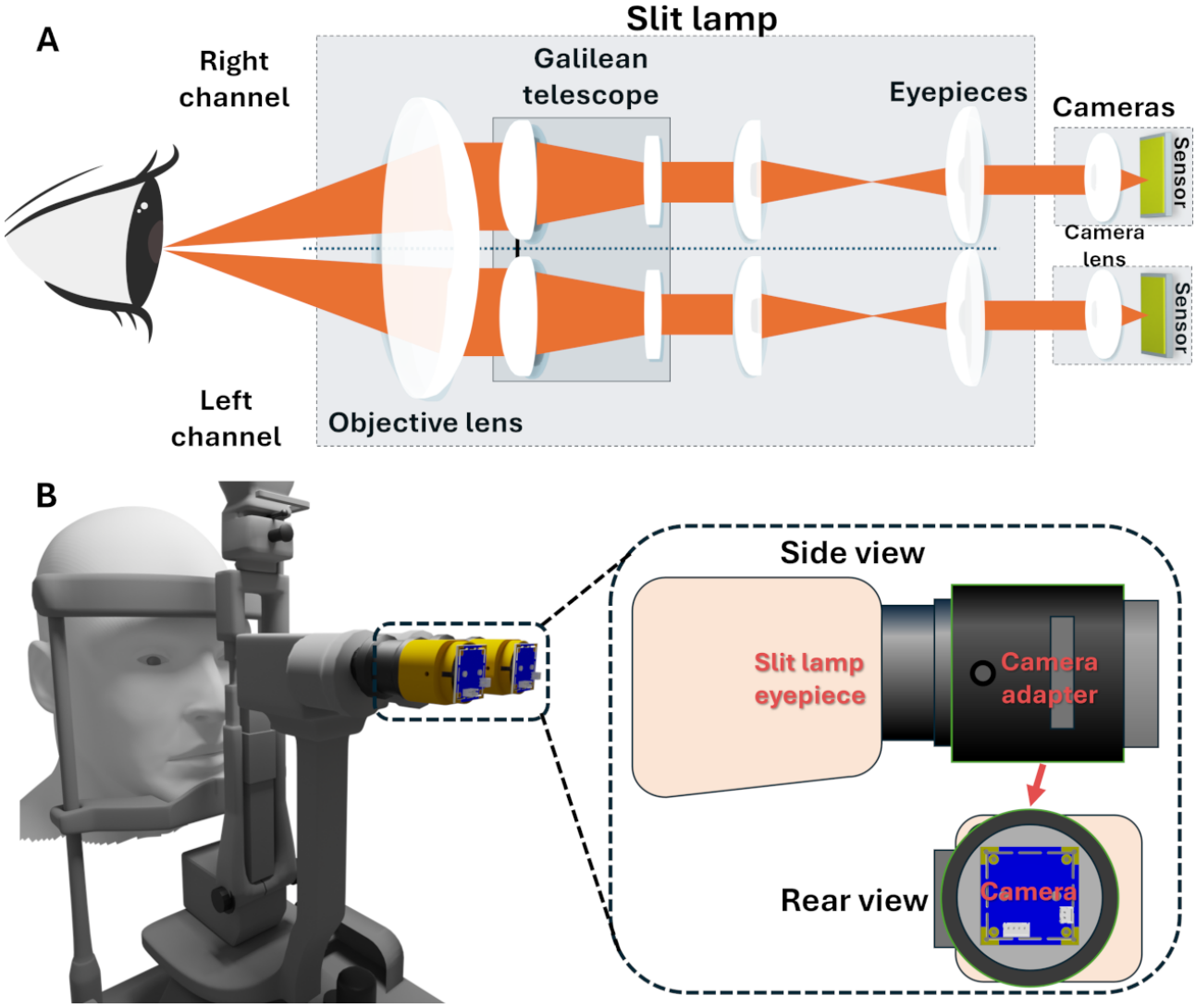
The study was conducted in accordance with the tenets of the Declaration of Helsinki and was approved by the institutional review board (IRB) of Florida International University (IRB-19 - 0112-CR02 and IRB-26 - 0004). Written informed consent was obtained from all participants prior to enrollment. All data were deidentified before analysis, and no personally identifiable information or protected health information was collected or stored by the MR-SLP system.

System Design

The MR-SLP system developed in this study was built on a commercial slit lamp (XCEL250, Recheit Inc). Two board-level USB video cameras (ELP-USB16MP01-L75, 1/2.8-inch sensor, maximum resolution: 4656×3496 pixels) formed a stereoscopic pair and were calibrated before mounted on the left and right eyepieces using custom-designed adapters. The adapters allow fine-tuning of the cameras' positions to ensure accurate alignment of their optical axes with the eyepieces. A schematic of the system is shown in Figure 1.

XSL•FO
RenderX

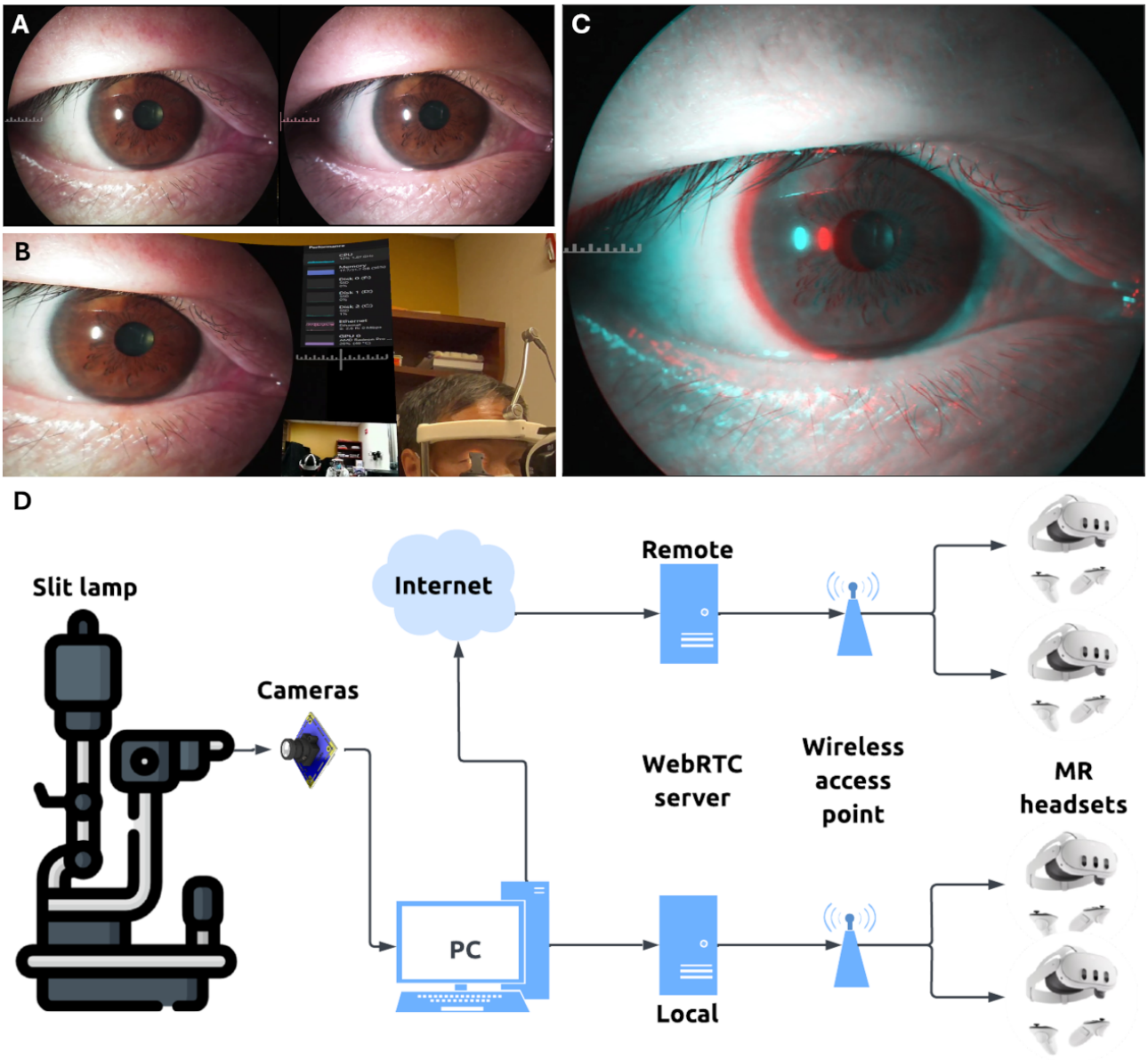
Figure 1. Design of the mixed reality–based slit-lamp (MR-SLP) system. (A) Schematic of the optical path of the MR-SLP. Light reflected by the examined area enters the front objective lens, then proceeds through the Galilean telescope assembly and the eyepieces before reaching the camera lens and being detected by imaging sensors. (B) Adaptation of a traditional slit lamp for MR-SLP functionality. Two cameras are mounted on the slit-lamp eyepieces using adapters, allowing for stereoscopic ocular imaging.



The videos captured by the 2 cameras were first streamed via USB to a local computer, where they were processed in real time using Open Broadcaster Software Studio (64-bit, version 31.0), an open-source video streaming application. For the video processing pipeline in Open Broadcaster Software Studio, we used the NVIDIA NVENC H.264 hardware encoder to achieve efficient, low-latency streaming. The encoder settings were specifically tuned for real-time performance. Key parameters included rate control set to constant bitrate, bitrate at 10 Mbps, keyframe interval set to 0 seconds for automatic management, and preset configured to “fastest” with “ultra low latency” tuning. The stream used the “baseline” profile to maximize compatibility and minimize processing overhead. The processed videos were formatted into side-by-side 3D mode (Figure 2A), a widely used format for conveying stereoscopic information in image processing [11]. Subsequently, the side-by-side 3D video was transmitted to local and remote Web Real-Time Communication (WebRTC) servers using the WebRTC HTTP ingestion protocol, with an output resolution of either 1080p or 720p. For local streaming, a free WebRTC server (OSSRS/SRS v6.0 from Docker Hub) was deployed on a PC, while remote streaming was facilitated by a commercial WebRTC service provided by dolby.io over the internet. The processed streams were broadcasted wirelessly to MR headsets using the WebRTC HTTP egress protocol. The headsets (Meta Quest 3; Meta Platforms Inc) then rendered the stream into a stereoscopic view with their built-in video player for immersive visualization (Figure 2B). An image in red-cyan stereoscopic format is provided (Figure 2C) as an example, which can be viewed using red-cyan anaglyph 3D glasses. A diagram of the broadcasting architecture is shown in Figure 2D.

XSL•FO
RenderX

Figure 2. Imaging procedures from camera to the head-mounted display (HMD)–rendered display. (A) Example images of a participant's eye captured with the mixed reality (MR)–based slit-lamp (MR-SLP) cameras in the side-by-side 3D (SBS) format. (B) The view of the local operator through their MR headset. The embedded window on the left displays the stereoscopic view (rendered from the SBS video) of the participant's eye. The participant and the environment are seen with the see-through function of the MR headset. (C) An image in red-cyan stereoscopic format demonstrates the 3D view as an example that can be observed using red-cyan anaglyph 3D glasses. (D) Broadcasting system architecture. The videos captured by the cameras on a slit lamp are processed by a local PC and transmitted wirelessly to local and remote MR headsets via Web Real-Time Communication (WebRTC) servers.



Spatial Resolution Evaluation

A US Air Force 1951 resolution target (Edmund Optics Inc) was used to evaluate the spatial resolution of the MR-SLP. The target was placed at the imaging plane of the MR-SLP and imaged at 25× optical magnification. Raw images of the resolution target were captured using the MR-SLP's USB cameras with the built-in camera app on Microsoft Windows 11, at a resolution of 4656 × 3496 pixels, which is the cameras' maximum resolution. Images displayed in the MR headset were captured through real-time video streaming, using the built-in snapshot tool of the Meta Quest 3 at a resolution of 1080p.

Intensity profiles were generated across the line pairs for selected group–element combinations from the acquired resolution target images using MATLAB (MathWorks Inc). The minimum resolvable line pairs were determined according to the Rayleigh criterion, wherein the central maximum of one pattern coincides with the first minimum of another. Additionally, a subjective evaluation was conducted in which participants compared the minimum readable line pairs observed directly through the eyepieces with those visible in the images captured via camera.

Tube-Threading Test for Evaluating Visuomotor Performance With the MR-SLP

To assess the visuomotor performance of the MR-SLP, a tube-threading test was designed based on the concept of a bead-threading task [12,13], a well-known method for estimating fine visuomotor performance that was critically affected by depth perception. This tube-threading test consisted of 2

XSL•FO
RenderX

sequential tasks. In task 1, the participant was asked to pick up each of the 10 tubes from a container using a tweezer, place the tube at the focal plane of the slit lamp to be seen clearly, and then return the tube to the container. In task 2, the participant was asked to pick up each tube, thread a wire into it under the visual guidance of the slit lamp, and return the tube to the container. The participants were required to complete 2 tasks 5 times under each designed condition shown in Table 1. High-precision tweezers (PL-30; Fisher Scientific) primarily for microscopy applications were used to grab the tubes. The tubes used in this test were made from micropipette tips. Their outer diameter ranged from 0.8 mm to 1.6 mm, and their inner diameter ranged from 0.4 mm to 1.2 mm. The diameter of the wire used in this test was 0.2 mm. The small size of these tubes and the wire, combined with 10× optical magnification of the slit lamp, required precise depth perception to complete the designed tasks.

Five individuals aged between 21 and 61 years (mean 42.2, SD 16.5 years) were recruited for this test. All participants provided written informed consent to participate in the study. Each participant had normal visual function, with a best-corrected visual acuity of 20/20 or better and a minimum stereoacuity of approximately 40 arc seconds, assessed using random-dot stereograms [14]. All participants did not have any history of ophthalmologic surgery, ocular motility disorders, known fine motor impairments, or other physical impairments that could interfere with the task performance. All participants completed a short acclimation session before the main test. This included (1) adjusting the eyepieces and MR headset to achieve a full field of view at test magnification, (2) finding a comfortable position so that their body and hands were stable and at ease while performing the tube-threading test, and (3) practicing the tube-threading task until they felt comfortable.

Table . Test conditions used in the tube-threading evaluation.

Condition ID (c)	Conditions	Description
1	MR-SLP[a], 2D, 60 fps[b]	Nonstereoscopic stream[c], output resolution 720p, frame rate 60 fps, with MR[d] headset.
2	MR-SLP, 3D, 30 fps	Stereoscopic stream, output resolution 720p, frame rate 30 fps, with MR headset.
3	MR-SLP, 3D, 60 fps	Stereoscopic stream, output resolution 720p, frame rate 60 fps, with MR headset.
4	Conventional slit lamp	Direct view through the eyepieces of the slit lamp, without MR headset.

[a]MR-SLP: mixed reality–based slit lamp.

[b]fps: frames per second.

[c]Nonstereoscopic stream was created by converting the video from a single mixed reality–based slit-lamp camera into the side-by-side 3D format.

[d]MR: mixed reality.

The time taken for each task under each condition was individually recorded, denoted as Ttask1(p,c,i) and Ttask2(p,c,i), where *p* is the participant number, *c* is the condition ID, and *i* is the repeat.

From a single repeat, the time to thread the tube Tthr(p,c,i)was calculated as

$$(1)\mathrm{Tthr}(p,c,i)=\mathrm{Ttask2}(p,c,i)-\mathrm{Ttask1}(p,c,i)$$

The mean of Ttask1(p,c,i) and Tthr(p,c,i) from the 5 repeats of a given participant under the same condition were calculated as

$$(2)\mathrm{T-task1}(p,c)=15\sum_{i=1}^{5}\mathrm{Ttask1}(p,c,i),\mathrm{T-thr}(p,c)=15\sum_{i=1}^{5}\mathrm{Tthr}(p,c,i)$$

and Ttask2(p,c,i) taken for each task under each condition was individually recorded. T-task1(p,c) and T-thr(p,c) for all 4 conditions were collectively analyzed using statistical methods to identify the effects of the test conditions on participants' performance.

Telemedicine Assessment

A telemedicine assessment was conducted to evaluate the streaming quality, remote usability, and applications of the MR-SLP. Five participants were assigned to different roles across 3 locations in Miami. The local site was at the main campus of Florida International University, where an ophthalmologist with around 10 years of clinical experience operated the MR-SLP and performed test examinations on a volunteer with normal visual function. Stereoscopic slit lamp video was captured and streamed in real time at an output resolution of 1080p. Simultaneously, a second ophthalmologist with more than 20 years of clinical experience watched the stereoscopic video from the MR-SLP using an MR headset at the first remote site, Bascom Palmer Eye Institute at the University of Miami Miller School of Medicine. He communicated in real-time with the MR-SLP operator at the local site, providing expert guidance and diagnostic opinions. In addition, a senior scientist at the first remote site also observed the examination procedure through an MR headset for technical assistance. Finally, a senior scientist at the second remote site, the Florida International University Engineering Center, observed the process using an MR headset. All local and remote participants verbally assessed the system's technical performance, including image quality and streaming stability. The latency, particularly the round-trip time, was provided by the remote WebRTC platform. Real-time audio communication between the local operator and remote participants was enabled using Zoom (Zoom Communications, Inc).

XSL•FO
RenderX

Data Analysis

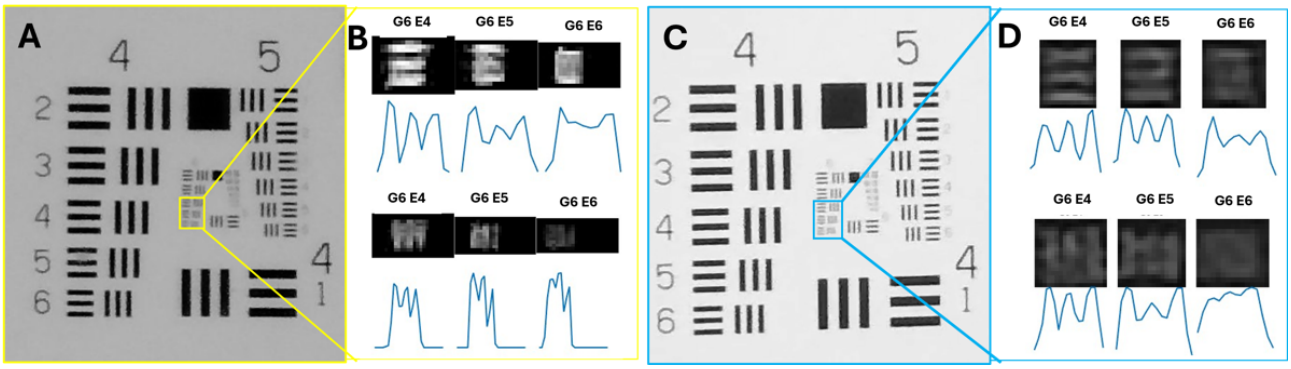
Statistical analyses were performed using R (version 4.5.1; R Foundation for Statistical Computing). The normality of the tube-threading task completion times was assessed using the Shapiro-Wilk test. Given the nonnormal distribution of the data, the nonparametric Kruskal-Wallis test was used to compare conditions, followed by post hoc analysis using the Dunn test with Bonferroni correction. Statistical significance was set at $P<.05$.

Results

Image Quality of the MR-SLP: Spatial Resolution

Our objective analysis revealed that both the MR-SLP cameras and MR headset were able to resolve element 5 in group 6 of the resolution target, corresponding to a spatial resolution of 102 line pairs/mm with a line width of approximately 9.8 µm (Figure 3). Additionally, we conducted a subjective evaluation with 3 participants who had a best-corrected visual acuity of 20/20 or better. These participants noted that, although they could discern element 5 in group 6 through the MR headset, they were able to identify up to element 6 in group 6 (114 line pairs/mm) when viewing directly through the slit-lamp eyepieces. All participants noted that the lines appeared slightly clearer when viewed through the traditional slit lamp compared with the MR headset display. This subjective difference may indicate that the fidelity of the current digital imaging pathway, which includes the camera and display, may not match that of direct optical viewing with the human eye.

Figure 3. Resolution target imaging. (A) Image of the line pairs (lps) from groups 4 to 6 for the resolution target captured directly by the mixed reality (MR)–based slit-lamp camera. (B) The upper panel shows the intensity profiles across the horizontal lps of elements 4 to 6 in group 6; the bottom panel displays the intensity profiles across the vertical lps. (C) Image of the lps from groups 4 to 6 of the resolution target captured from the MR headset. (D) The upper panel shows the intensity profiles across the horizontal lps of elements 4 to 6 in group 6; the bottom panel shows the intensity profiles across the vertical lps.



Tube-Threading Test With the MR-SLP

All participants successfully completed 2 sequential tasks (Figure 4) within a reasonable timeframe (<10 min). Their T-task1(p,c) and T-thr(p,c) records are summarized in Tables 2 and 3, respectively. According to the records, T-task1(p,c) was typically the shortest in the direct view condition (c=4). A Shapiro-Wilk test was conducted with T-task1(p,c) and T-thr(p,c) for each test condition. The test outcome indicated that the assumption of a normal distribution was not applicable in this dataset. Therefore, a nonparametric method, the Kruskal-Wallis test, was used to compare T-task1(p,c) and T-thr(p,c) under different conditions. Please find the whole dataset in Table S1 and Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 4. The tube-threading test and results. (A) The participant uses tweezers to grasp a tube, (B) positions the tube into the slit lamp's field of view until it is clearly visible, and (C) threads a wire into the tube under the slit lamp's view. (D) Box plot of Ttask1(p,c,i) across the 4 test conditions. (E) Box plot of Tthr(p,c,i) across the 4 test conditions. Each dark dot represents an individual time record, and the red diamond symbol represents the mean value. *P<.05 between 2 conditions; fps: frames per second. MR: mixed reality; ns: no statistically significant difference.

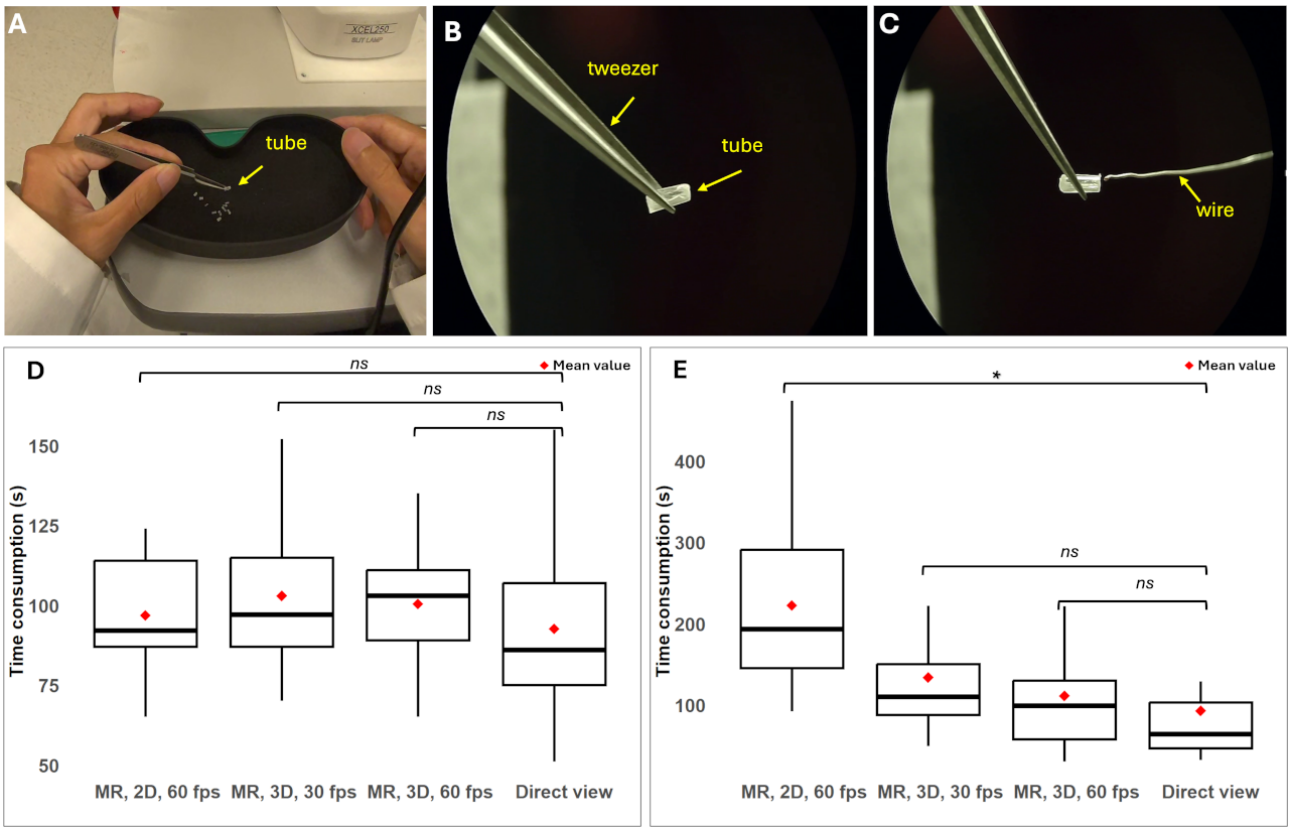


Table . T-task1(p,c) of all participants under all 4 test conditions.[a]

Participant (p)	MR[b], 2D, 60 fps[c,d], mean (SD)	MR, 3D, 30 fps[e], mean (SD)	MR, 3D, 60 fps[f], mean (SD)	Direct view[g], mean (SD)
1	77 (9)	92 (7)	84 (18)	63 (9)
2	107 (12)	126 (15)	111 (7)	116 (29)
3	118 (4)	120 (8)	124 (22)	107 (9)
4	91 (5)	93 (8)	91 (2)	75 (5)
5	91 (11)	85 (17)	91 (19)	103 (31)

[a]Time was recorded in seconds, and each number represents the mean (SD) from the 5 repeats.

[b]MR: mixed reality.

[c]fps: frames per second.

[d]Condition ID 1.

[e]Condition ID 2.

[f]Condition ID 3.

[g]Condition ID 4.

XSL•FO
RenderX

Table . T-thr(p,c) of all participants under all 4 test conditions.[a]

Participant (p)	MR[b], 2D, 60 fps[c,d], mean (SD)	MR, 3D, 30 fps[e], mean (SD)	MR, 3D, 60 fps[f], mean (SD)	Direct view[g], mean (SD)
1	241 (100)	138 (77)	153 (124)	68 (29)
2	317 (119)	133 (39)	110 (20)	76 (38)
3	131 (26)	92 (14)	75 (30)	57 (18)
4	175 (56)	93 (37)	69 (25)	48 (10)
5	249 (113)	213 (120)	149 (81)	215 (139)

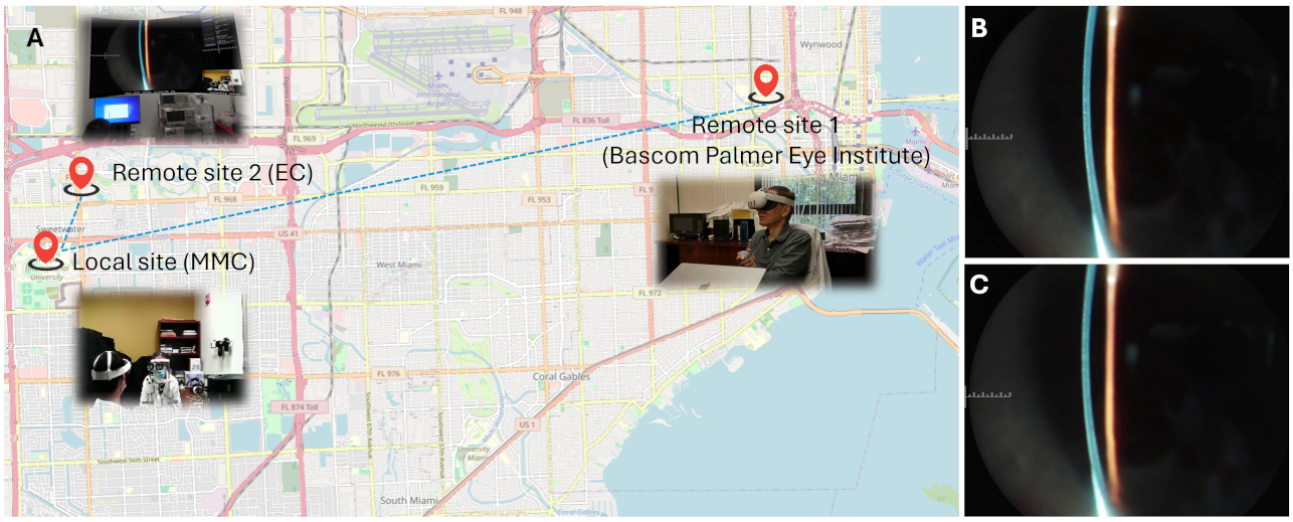
[a]Time was recorded in seconds, and each number represents the mean (SD) from the 5 repeats.
[b]MR: mixed reality.
[c]fps: frames per second.
[d]Condition ID 1.
[e]Condition ID 2.
[f]Condition ID 3.
[g]Condition ID 4.

As shown in Figure 4D, the distribution of Ttask1(p,c,i) for all 4 test conditions are similar. This observation was confirmed by the result of the Kruskal-Wallis test, indicating that there were no significant differences among T-task1(p,c) for all 4 conditions (n=20; *P*=.77). According to the trend depicted in Figure 4E, the mean value of T-thr(p,c) for each condition was the shortest in the direct view condition (c=4) and gradually increased from test condition 3 (MR 3D 60 frames per second [fps]) to test condition 2 (MR 3D 30 fps) and then to test condition 1 (MR 2D 60 fps). A statistically significant difference in T-thr(p,c) was observed among the 4 test conditions according to the Kruskal-Wallis test (n=20; *P*=.04). Subsequent post hoc analysis using Dunn test with Bonferroni correction revealed that only T-thr(p,1) and T-thr(p,4) are statistically significantly different, meaning that a statistically significant difference exists only between test condition 1 (MR 2D 60 FPS) and test condition 4 (direct view; n=5; *P*=.03). This analysis showed that the performance of the threading task had no significant difference between MR headset working on 3D mode and viewing directly through the eyepieces. The results also suggested that increasing the frame rate will reduce the time taken for the threading task.

Telemedicine Assessment

In the test, as depicted in Figure 5, the MR-SLP system's ability to simultaneously transmit high-resolution stereoscopic video to both local and remote MR headsets was successfully demonstrated. At the local site, the operator wearing the MR headset could simultaneously see the patient through the HMD's see-through function and view the stereoscopic examination images immersed in their real-world environment. Meanwhile, remote users connected via the internet could concentrate on the stereoscopic examination images, offering real-time expert guidance and diagnostic input, promoting effective collaboration over distances.

Figure 5. Test of mixed reality–based slit lamp (MR-SLP) for teleophthalmology. (A) The geographical relationship between the local site (Florida International University [FIU] main campus) and the remote sites (FIU Engineering Center [EC] and Bascom Palmer Eye Institute) involved in assessing image quality and user experience of the MR-SLP. (B) The images obtained directly from the camera of the MR-SLP; (C) The video frame displayed on the remote MR headset at the same time. No noticeable differences were observed between these 2 images. MMC: Modesto Maidique Campus.



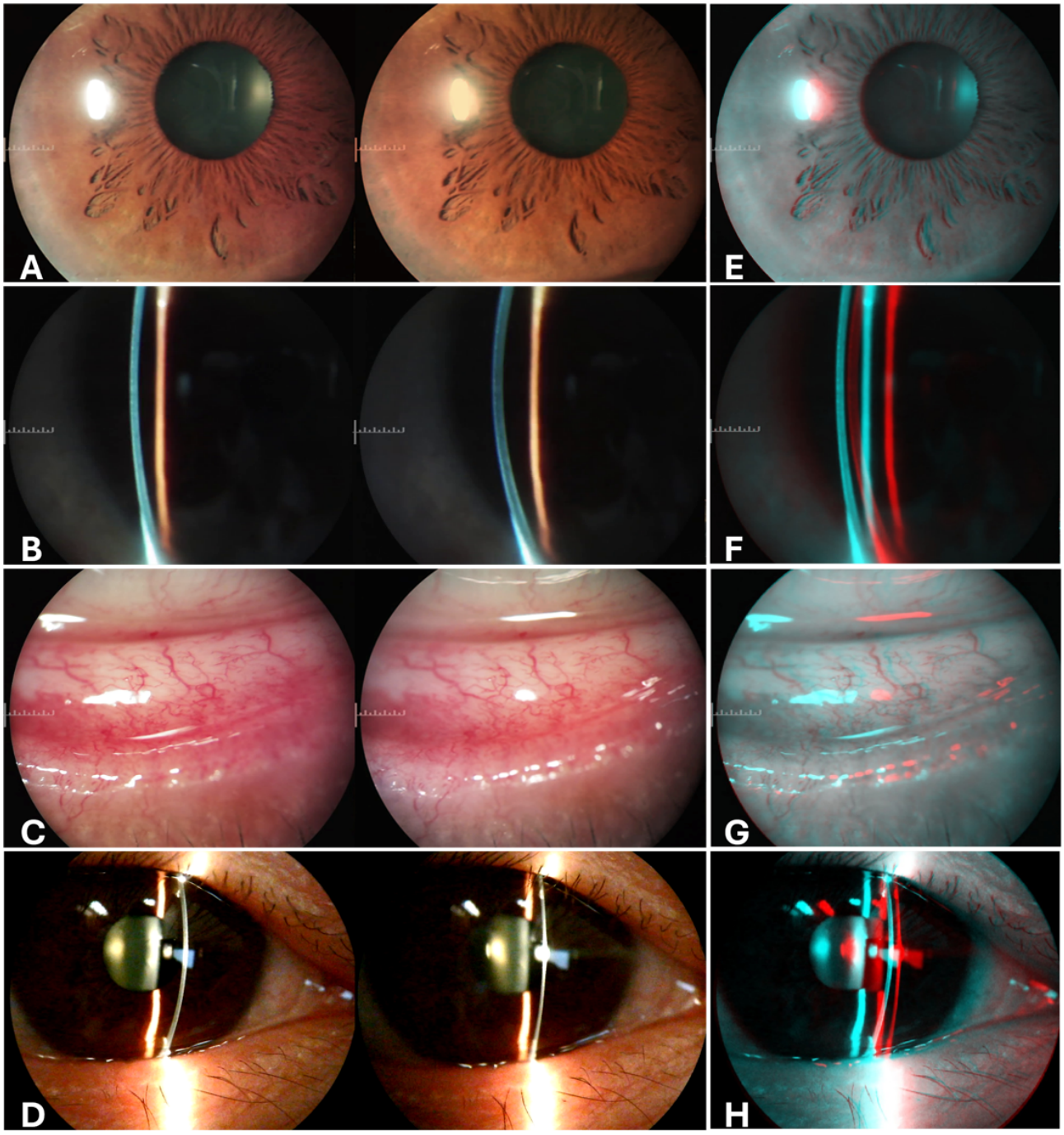
XSL•FO
RenderX

During examination of the volunteer's eye with the MR-SLP, anatomical details of the anterior segment (Figure 6) were accurately captured and displayed in real time on both local and remote MR headsets. According to the ophthalmologists participating in this trial, they reported that the image quality viewed in their MR headsets closely rivaled that observed through the slit-lamp eyepieces. The real-time video streams in this trial also maintained robust depth perception and exhibited low latency at remote sites, with an average round-trip time of less than 40 milliseconds and video jitter below 20 milliseconds. This allowed for a highly interactive and collaborative examination by the local and remote ophthalmologists. Both ophthalmologists involved in the trial stated that the MR-SLP system compared favorably to traditional slit lamp viewing and highlighted its potential advantages for clinical practice.

Fundus examination using the MR-SLP was also evaluated. Following the conventional procedure, the operator, wearing the MR headset, held an ophthalmic lens (90D; Volk Optical) between the volunteer's eye and the slit-lamp objective to visualize the retina. An example result is shown in Figure 7. Under a nonmydriatic condition, different regions of the retina were clearly visualized in the headsets by adjusting the slit lamp and the Volk lens. Streaming of the fundus images to the local and remote headsets was processed at the background without interfering with the operator's performance. In the test, the operator required only a brief additional period to adapt to the hand-eye coordination. When operating from alternative positions, such as an off-axis (lateral) position, maneuvering the slit lamp and Volk lens requires establishing a new hand-eye coordination and required additional practice. However, the operator could select the operating position for fundus examination without compromising overall performance.

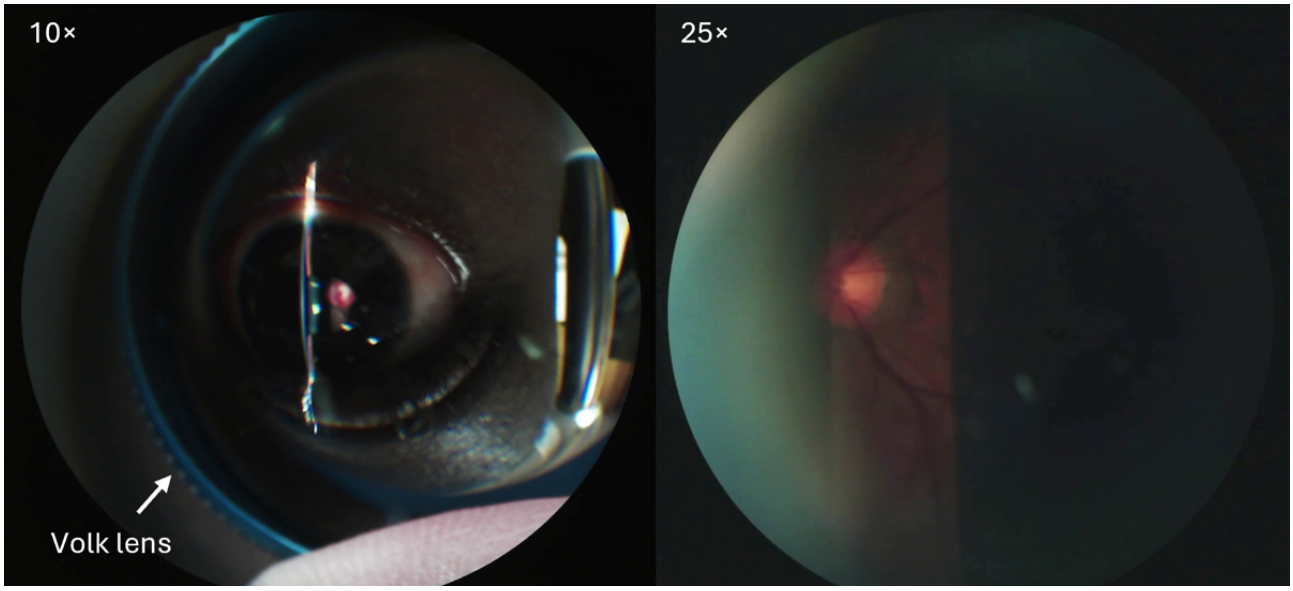
XSL•FO
RenderX

Figure 6. Images of the anterior segment structures captured by mixed reality–based slit lamp locally in side-by-side 3D format. (A) Iris, (B) corneal cross-section (blue stripe), (C) conjunctiva, and (D) lens. (E-H) Images in red-cyan stereoscopic format demonstrate the 3D view that can be observed using red-cyan anaglyph 3D glasses.



XSL•FO
RenderX

Figure 7. Fundus examination using the mixed reality–based slit lamp. Images were acquired through a handheld 90D Volk lens, showing retinal regions under a nonmydriatic condition at different optical magnifications.



Discussion

Principal Findings

We tested the performance and validated the feasibility of the MR-SLP for providing stereoscopic slit-lamp views of the eye, including the anterior segment and the fundus, in real time both locally and remotely. The MR-SLP may provide a new approach for slit-lamp examinations by eliminating the need for continuous close viewing through the eyepieces, which may offer more comfort for both the operator and the patient during examinations and more flexibility in operating the machine. Furthermore, the system enhances the efficiency of slit-lamp training by enabling multiple participants to view real-time, stereoscopic diagnostic images as if viewing them directly through the eyepieces. The remote ophthalmologist wearing the MR headset gains an immersive experience as if operating the slit lamp directly, facilitating stereoscopic teleophthalmology.

Factors Affecting Spatial Resolution

The MR headsets used in this study have a display resolution of approximately 2K per eye, while the cameras have a maximum resolution of 4K. The measured spatial resolution was similar (102 line pairs/mm) for both the images acquired by the MR-SLP cameras and the MR device, suggesting that at the tested magnification (25×) the system's effective resolution was limited by the slit-lamp optics. It is also possible that the MR device's rendering algorithms enhance the perceived clarity in the region of interest, compensating partially for the lower display pixel density [15]. However, the participants in the subjective evaluation test noted a difference in perceived clarity between the MR headset and direct eyepiece view. Since this observation was based on anecdotal reports from a small group of participants (n=3), future research should implement a larger psychophysical study to statistically quantify and validate any differences in perceived image quality.

Depth Perception, Effective Resolution, Frame Rate, and Latency

Accurate depth perception through the slit lamp is essential for precise clinical diagnosis and surgical procedures [16]. Our camera adapter and calibration algorithm allow fine-tuning of the camera position to precisely match the slit-lamp view with the digital camera view, ensuring consistent spatial mapping. The tube-threading test suggests that the MR-SLP in 3D mode may offer depth perception comparable to that of viewing through the eyepieces. Under test condition 1, participants reported that the main challenge was aligning the wire with the orifice of the tube without stereoscopic cues. A low frame rate also exacerbated the issue by creating mistiming between visual feedback and motion control, which increased the difficulty of the task. Additionally, while all participants completed a brief acclimation session, the study did not formally assess the learning curve for using the MR-SLP system. However, this preliminary test is valuable for the design of a more definitive, larger-scale validation study.

All the participants stated that the resolution target and tubes appeared clearer when viewed directly through the eyepieces. We believe this was caused by the effective resolution of the cameras used in this study during video streaming. The output frame rate was set at 60 fps only when the camera was operated at a resolution of 720p, which was limited by the camera hardware. This hardware bottleneck resulted in lower resolution of the streamed video than the slit lamp's optics and the MR headset display could provide. This issue can be resolved by using cameras that provide 1080p or higher video resolutions at a frame rate of 60 fps, which are abundant in today's market. The results of the tube-threading study, shown in Figure 4, also indicate that a higher frame rate led to better performance. This improvement may be attributed to the reduction of the motion blur at higher frame rates [17]. Future studies should explore how to best balance maintaining high frame rates with achieving higher resolutions, such as 1080p or 4K, for these dynamic procedural tasks. We believe that by introducing higher-quality

XSL•FO
RenderX

cameras and more advanced MR headsets, such as the Apple Vision Pro, the video quality of the MR-SLP at both local and remote sites will more closely match viewing through the eyepiece of a slit lamp.

System latency is a crucial factor in real-time applications, such as telemedicine, because delays can significantly impact the local operator's performance and diminish the quality of remote interactions. Latency can arise from network transmission, data processing, and display rendering [18]. While acceptable latency thresholds for demanding applications such as telesurgery are suggested to be below 100 to 200 milliseconds [19,20], platforms using WebRTC architecture can potentially achieve very low media transmission delays [21]. Remote tests conducted within the same city demonstrated a good performance with an average round-trip time of less than 40 milliseconds. Nevertheless, further evaluation through long-distance remote testing is needed.

MR Headset Challenges

In the current implementation of the MR-SLP system, Meta Quest 3 is used as the MR headset for both the local and remote locations. Quest 3 offers advantages for biomedicine applications, including high-resolution displays and advanced color pass-through that enable detailed 3D visualization of anatomical structures. The see-through function and accurate spatial mapping facilitate natural interaction and situational awareness, allowing users to seamlessly integrate virtual content with the real clinical environment. Additionally, Quest 3's onboard processing power efficiently converts side-by-side images into immersive 3D views, supporting real-time telemedicine and collaborative workflows. However, the headset's ergonomics and comfort remain suboptimal, particularly during prolonged use, such as continuous wear for more than 60 minutes. While the display resolution is high, further improvements—such as 4K or higher per eye—would be beneficial for even greater image clarity. We are also investigating alternatives to Quest 3, such as Apple Vision Pro, or augmented reality glasses. Our system's architecture is compatible with other extended reality platforms such as augmented reality for the local applications and virtual reality for remote physicians.

Limitations and Future Studies

While the technique holds promise for enhancing slit-lamp examinations, teleophthalmology, and medical training, several key limitations of our study need to be acknowledged, such as small sample size for statistical analysis and subjective evaluation based on anecdotal reports.

In this study, we did not evaluate the learning curve for different operators associated with adapting to the MR-SLP view, including potential differences in hand-eye coordination during complex maneuvers. For the evaluation of user experience with the MR-SLP, feedback from participating ophthalmologists was collected through verbal reports rather than using a validated assessment scale. Although this approach provided valuable initial insights for a feasibility study, it did not allow for quantitative analysis of user satisfaction or formal measurement of interrater agreement regarding the system. Future studies should use validated questionnaires to enable a more robust evaluation of the system's usability and acceptance in clinical settings. Additionally, the camera adapter and software implementation are cost-effective and require minimal modification, allowing for versatile integration with most slit lamps and potentially other binocular microscopes.

Conclusions

This study demonstrated the technical feasibility of the MR-SLP in capturing and transmitting real-time, stereoscopic slit-lamp examination videos to MR headsets. Our tests showed that the system can provide visuomotor performance comparable to direct viewing, and it provided low-latency streams for effective remote collaboration.

Acknowledgments

The authors attest that generative artificial intelligence was not used in the generation of any part of this manuscript.

Funding

This work was supported by National Institutes of Health grant R01EY031492.

Data Availability

The summarized data generated during this study are included in this published article and its multimedia appendix. The full raw datasets are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: SJ, RZ
Data curation: RZ
Formal analysis: RZ, WCL, SJ
Investigation: RZ, SJ, WCL, BLL, RW, NA
Resources: SJ
Writing–original draft: RZ, WCL, SJ
Writing–review and editing: RZ, WCL, SJ

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Time required for the tube-threading tasks and picture of mixed reality–based slit lamp operation.
[DOCX File, 204 KB - xr_v3i1e93513_app1.docx]

References

1. Kate A, Basu S. Slit lamp biomicroscopy. In: Das T, Satgunam P, editors. Ophthalmic Diagnostics: Technology, Techniques, and Clinical Applications: Springer Nature; 2024:167-181. [doi: 10.1007/978-981-97-0138-4_14]
2. Gellrich MM. History of the slit lamp. In: The Slit Lamp: Applications for Biomicroscopy and Videography: Springer Science & Business Media; 2013:189-210.
3. Fritz B, Paschko E, Young W, et al. Comprehensive compositional analysis of the slit lamp bacteriota. Front Cell Infect Microbiol 2021;11:745653. [doi: 10.3389/fcimb.2021.745653] [Medline: 34869057]
4. Shah SH, Garg AK, Patel S, Yim W, Jokerst JV, Chao DL. Assessment of respiratory droplet transmission during the ophthalmic slit-lamp exam: a particle tracking analysis. Am J Ophthalmol 2021 Feb;222:76-81. [doi: 10.1016/j.ajo.2020.08.046] [Medline: 32980331]
5. Dutt S, Nagarajan S, Vadivel SS, et al. Design and performance characterization of a novel, smartphone-based, portable digital slit lamp for anterior segment screening using telemedicine. Transl Vis Sci Technol 2021 Jul 1;10(8):29. [doi: 10.1167/tvst.10.8.29] [Medline: 34319384]
6. Huang Z, Yang J, Wang H, Pang CP, Chen H. Comparison of digital camera real-time display with conventional teaching tube for slit lamp microscopy teaching. Curr Eye Res 2022 Jan;47(1):161-164. [doi: 10.1080/02713683.2021.1952606] [Medline: 34224279]
7. Cao B, Vu CH, Keenan JD. Telemedicine for cornea and external disease: a scoping review of imaging devices. Ophthalmol Ther 2023 Oct;12(5):2281-2293. [doi: 10.1007/s40123-023-00764-3] [Medline: 37458978]
8. Schmidt TA. On slit-lamp microscopy. Doc Ophthalmol 1975 Nov 21;39(1):117-153. [doi: 10.1007/BF00578760] [Medline: 1201695]
9. Ong CW, Tan MC, Lam M, Koh VT. Applications of extended reality in ophthalmology: systematic review. J Med Internet Res 2021 Aug 19;23(8):e24152. [doi: 10.2196/24152] [Medline: 34420929]
10. Worlikar H, Coleman S, Kelly J, et al. Mixed reality platforms in telehealth delivery: scoping review. JMIR Biomed Eng 2023 Mar 24;8:e42709. [doi: 10.2196/42709] [Medline: 38875694]
11. Su GM, Lai YC, Kwasinski A, Wang H. 3D video communications: challenges and opportunities. Int J Commun Syst 2011 Oct;24(10):1261-1281. [doi: 10.1002/dac.1190]
12. O'Connor AR, Birch EE, Anderson S, Draper H. Relationship between binocular vision, visual acuity, and fine motor skills. Optom Vis Sci 2010 Dec;87(12):942-947. [doi: 10.1097/OPX.0b013e3181fd132e] [Medline: 21057348]
13. Piano ME, O'Connor AR. The effect of degrading binocular single vision on fine visuomotor skill task performance. Invest Ophthalmol Vis Sci 2013 Dec 17;54(13):8204-8213. [doi: 10.1167/iovs.12-10934] [Medline: 24222309]
14. Deepa BM, Valarmathi A, Benita S. Assessment of stereo acuity levels using random dot stereo acuity chart in college students. J Family Med Prim Care 2019 Dec;8(12):3850-3853. [doi: 10.4103/jfmpc.jfmpc_755_19] [Medline: 31879624]
15. Rao L, Argaman N, Zhuang J, et al. Display and optics architecture for Meta's AR/VR development. IEEE Open J Immersive Disp 2024;1:71-78. [doi: 10.1109/OJID.2024.3370888]
16. Bogdanova R, Boulanger P, Zheng B. Depth perception of surgeons in minimally invasive surgery. Surg Innov 2016 Oct;23(5):515-524. [doi: 10.1177/1553350616639141] [Medline: 27009686]
17. Li R, St George RJ, Wang X, et al. Moving towards intelligent telemedicine: computer vision measurement of human movement. Comput Biol Med 2022 Aug;147:105776. [doi: 10.1016/j.compbiomed.2022.105776] [Medline: 35780600]
18. Motiwala ZY, Desai A, Bisht R, Lathkar S, Misra S, Carbin DD. Telesurgery: current status and strategies for latency reduction. J Robot Surg 2025;19(1):153. [doi: 10.1007/s11701-025-02333-1]
19. Wirz R, Torres LG, Swaney PJ, et al. An experimental feasibility study on robotic endonasal telesurgery. Neurosurgery 2015 Apr;76(4):479-484. [doi: 10.1227/NEU.0000000000000623] [Medline: 25599203]
20. Xu S, Perez M, Yang K, Perrenot C, Felblinger J, Hubert J. Determination of the latency effects on surgical performance and the acceptable latency levels in telesurgery using the dV-Trainer(®) simulator. Surg Endosc 2014 Sep;28(9):2569-2576. [doi: 10.1007/s00464-014-3504-z] [Medline: 24671353]
21. Hamidouche W, Bariah L, Debbah M. Immersive media and massive twinning: advancing toward the Metaverse. IEEE Commun Mag 2024 Jul;62(7):20-32. [doi: 10.1109/MCOM.2300399]

Abbreviations

fps: frames per second

XSL•FO
RenderX

HMD: head-mounted display
IRB: institutional review board
MR: mixed reality
MR-SLP: mixed reality–based slit lamp
WebRTC: Web Real-Time Communication

Edited by I Steenstra; submitted 13.Feb.2026; peer-reviewed by A Saxena, L Liu; revised version received 23.Apr.2026; accepted 23.Apr.2026; published 14.May.2026.

Please cite as:
Zhou R, Lin WC, Lam BL, Wen R, Amadi N, Jiao S
Mixed Reality–Based Slit Lamp for Ophthalmic Examination and Telemedicine: Technological Development and Validation Study
JMIR XR Spatial Comput 2026;3:e93513
URL: https://xr.jmir.org/2026/1/e93513
doi:10.2196/93513

© Rui Zhou, Wei-Chiang Lin, Byron L Lam, Rong Wen, Noble Amadi, Shuliang Jiao. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 14.May.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Brief Virtual Reality and Mixed Reality Mindfulness Breathing Exercise for Emotional Well-Being and Cognitive Functions in University Students: Within-Subjects Experimental Design Study

Zoey K Y Eun, BSocSc; Charmaine Jiali Koh, BSocSc; Hwajin Yang, PhD; Adalia Y H Goh, MPhil; Meilan Hu, BSocSc; K T A Sandeeshwara Kasturiratna, MPhil; Andree Hartanto, PhD

School of Social Sciences, Singapore Management University, 10 Canning Rise, Singapore, Singapore

Corresponding Author:
Zoey K Y Eun, BSocSc
School of Social Sciences, Singapore Management University, 10 Canning Rise, Singapore, Singapore

Abstract

Background: Mindfulness has been shown to enhance emotional well-being and cognitive performance, yet much of this evidence stems from interventions requiring prolonged practice, making them time-consuming and less accessible. Recent studies suggest that brief mindfulness sessions may also yield positive outcomes, but the effectiveness of such interventions in virtual reality (VR) and mixed reality (MR) remains underexplored.

Objective: This study investigates the effects of brief mindfulness breathing exercises delivered through VR and MR on attentional and emotional restoration and self-control capacity.

Methods: Using a within-subjects experimental design, 102 undergraduate participants (n=83, 81.4% female; mean age 20.87, SD 1.89) completed a brief (approximately 15 min) VR and MR mindfulness breathing intervention delivered via a head-mounted display and a duration-matched mind-wandering audio control condition. Participants were undergraduates recruited via convenience sampling from psychology courses in a local university in Singapore. These conditions were separated by a 1-week washout period. Emotional well-being and self-control capacity were measured at baseline and post treatment, using self-report measures, whereas working memory capacity was measured at both time points, using operation span at baseline and rotation span post treatment.

Results: Repeated-measures ANOVAs (α=.05) indicated that VR and MR mindfulness breathing conditions significantly enhanced positive affect (P<.001, η_p^2=0.366), reduced negative affect (P<.001, η_p^2=0.279), and improved self-control capacity (P<.001, η_p^2=0.219), compared with the mind-wandering control condition. In contrast, no significant differences were observed for working memory, and Bayesian analyses provided moderate evidence in support of the null hypothesis for both the main effect of condition and the time×condition interaction (BF_{01}=7.42 and BF_{01}=5.55, respectively). Participants reported significantly greater absorption in the VR and MR conditions than in the control condition (Cohen's d=−1.61, 95% CI −1.91 to −1.32).

Conclusions: These findings suggest that a brief VR and MR mindfulness breathing exercise improves emotional well-being and self-control capacity relative to a mind-wandering control but does not yield short-term benefits for working memory. In contrast to existing studies that typically emphasize stress reduction or rely on multisession digital interventions, this study highlights that a single brief VR or MR session can enhance key emotional and self-regulatory outcomes. As such, these results underscore the potential of VR and MR mindfulness interventions as scalable and accessible tools for promoting mental well-being, while also pointing to the need for further research to optimize their cognitive impact.

(JMIR XR Spatial Comput 2026;3:e84239) doi:10.2196/84239

KEYWORDS

mindfulness; virtual reality; mixed reality; emotional well-being; working memory capacity

Introduction

Background

Mindfulness, defined as the enhanced awareness of moment-to-moment experiences, where one intentionally attends to their thoughts, emotions, and bodily sensations in a nonjudgmental and accepting manner [1,2], has consistently been shown to provide wide-ranging psychological benefits across both general [3,4] and clinical populations [5,6]. Evidence from numerous studies indicates that mindfulness-based interventions are associated with a range of positive outcomes related to well-being, including higher levels of life satisfaction [7-9], vitality [7,10], optimism [11,12], pleasant affect [7,13,14], improved emotional processing [15,16], and emotional

XSL•FO
RenderX

regulation [17-19]. Furthermore, these practices have been shown to effectively reduce depressive symptoms [7], stress [10,20,21], anxiety [22,23], and emotional reactivity [24-26]. These benefits arise not from suppressing or altering experiences, but from focusing on how individuals interpret their present-moment experiences [27,28]. When individuals perceive their thoughts and feelings as transient mental events that are impermanent in nature, they can reframe negative thoughts and decrease emotional reactivity, fostering a deeper sense of calm and well-being [29-31].

Beyond its mental health benefits, research has increasingly shown that mindfulness enhances cognitive performance [32-35] and is closely associated with improvements in executive functions, including inhibitory control, task-switching, and working memory [36-38]. The strong link between mindfulness and executive functions is theoretically well-grounded, as both rely on maintained attention and cognitive monitoring, indicating an overlap in their underlying cognitive mechanisms [39-41]. By fostering a relaxed mental state that promotes focused attention, mindfulness optimizes resource allocation and reduces mind-wandering, thereby improving cognitive performance [42,43].

In addition to cognitive benefits outlined above, mindfulness is closely linked to self-control. According to self-regulation theory [44], self-control operates as a limited resource that can be depleted under stressful and cognitively demanding conditions and replenished under restorative conditions. Cognitively demanding tasks require long-term executive control, such as maintaining attentional focus on goal-relevant information, suppressing distractions or impulses, and managing competing mental representations, and draw upon limited self-regulatory resources and temporarily reduce subsequent cognitive and emotional control capacity [45]. In contrast, restorative conditions are states or environments that facilitate the recovery of these depleted resources, including exposure to natural environments that promote attentional restoration, engagement in relaxation or mindfulness practices, or periods of rest that foster emotional rebalancing and cognitive recovery. When this resource is depleted, individuals may struggle with goal-directed behavior, impulse control, and regulation. A possible mechanism underlying this effect is that mindfulness enhances two fundamental self-control processes: (1) emotion regulation [17-19] and (2) attention regulation [33,46,47], both of which are crucial for maintaining self-regulatory capacity [44,48]. By enhancing these regulatory capacities, mindfulness provides a potential mechanism through which self-control resources can be replenished, thereby supporting long-term self-regulatory functioning.

Despite accumulating evidence on the positive outcomes of mindfulness on attentional and emotional restoration and self-control capacity, most findings are drawn from intensive programs, which typically span several weeks, requiring participants to engage in regular and often prolonged sessions under expert guidance [49]. Mindfulness-based stress reduction (MBSR) and mindfulness-based cognitive therapy are among the most pervasive and well-established interventions, with participants practicing mindfulness for up to 45 minutes daily and attending weekly group sessions that are 8 weeks long [2,50]. Specifically, 1 study demonstrated that MBSR is associated with reduced stress and enhanced psychological well-being [2], evidenced by improvements in emotional regulation, decreases in anxiety, and increases in life satisfaction among participants. Similar findings have been reported for other intensive mindfulness interventions. For example, a study found that participants who underwent a 1-month intensive mindfulness training reported significant reductions in anxiety and improvements in subjective well-being and self-compassion compared with a waitlist control group [51]. Additionally, another study found that novice meditators who participated in a 10-day intensive mindfulness training retreat demonstrated significant improvements in self-reported mindfulness, depressive symptoms, rumination, working memory, and maintained attention, relative to a comparison group who did not undergo any mindfulness training [32]. Generally, studies have highlighted the efficacy of intensive mindfulness interventions in improving well-being [52,53] and cognitive performance [54,55] compared with control conditions. Mindfulness fosters emotional regulation by increasing present-moment awareness and promoting acceptance of thoughts and feelings. According to the monitor and acceptance theory [56], mindfulness operates through two complementary mechanisms: (1) attention monitoring enhances awareness of moment-to-moment experiences, and (2) acceptance fosters a nonjudgmental and open stance toward these experiences. The interaction between these processes facilitates adaptive emotion regulation by allowing individuals to observe internal states without reacting impulsively.

Despite strong support for intensive mindfulness intervention, research on brief mindfulness has yielded mixed findings. On one hand, some studies report improvements in affect [57], reduced emotional reactivity, and enhanced attentional control [58,59]. Supporting this perspective, 1 study found that a single 15-minute mindfulness session significantly improved positive affect (PA) and reduced emotional reactivity in response to emotionally charged stimuli [60], highlighting the immediate benefits of mindfulness for emotional regulation. Similarly, another study showed that a single 15-minute brief mindfulness session could enhance attentional control by reducing mind-wandering during cognitive tasks [42], suggesting a notable impact on cognitive focus. While mindfulness has been associated with reduced emotional reactivity, it is important to note that reactivity itself is not inherently maladaptive. Emotional reactivity serves adaptive functions, which allow individuals to respond appropriately to meaningful positive or negative events. Excessive dampening of emotional responses, particularly to positive experiences, may reflect emotional blunting or reduced reward sensitivity, which can undermine well-being. Rather than eliminating reactivity, mindfulness promotes adaptive regulation, fostering flexibility to modulate emotions in a context-sensitive manner.

However, the outcomes of brief mindfulness practices are not always positive. Specifically, 1 study observed that a 15-minute mindfulness meditation did not significantly reduce negative affect (NA) relative to active control (guided progressive muscle relaxation training) and passive control (watching a TED Talks video) groups [61]. Likewise, another study found no

XSL•FO
RenderX

improvement in working memory capacity following a 15-minute mindfulness breathing exercise in 2 high-powered studies [62]. These mixed findings highlight both the promise and limitations of brief mindfulness and underscore the need for approaches that can deepen engagement without imposing the burden of long-term practice [63,64]. One potential explanation for these inconsistencies is that brief mindfulness exercises may lack the long-term engagement and depth provided by intensive practices, such as MBSR and mindfulness-based cognitive therapy, which are essential for achieving optimal outcomes [49,65]. To address this limitation, recent technological advancements in virtual reality (VR) and mixed reality (MR) may offer promising solutions to enhance brief mindfulness practices.

VR and MR technologies provide avenues for enriching mindfulness practices by offering immersive environments that foster greater focus and engagement [66-68]. For example, one study found that VR provides an immersive environment that may reduce typical barriers to mindfulness practice by increasing engagement and reducing distraction [66]. Another study reported that integrating mindfulness with immersive VR shows promise for improving mood, attention, and engagement [67]. Finally, another study demonstrated that immersive VR conditions can produce greater mindfulness and stress-reduction effects than conventional mindfulness formats in some samples [68]. VR, in particular, provides controlled multisensory environments that simulate natural settings and promote focus [69,70]. Additionally, VR provides multisensory experiences that allow users to interact with virtual natural environments, helping them to disconnect from everyday stressors and deepen their mindfulness practice [71]. Meanwhile, MR, which integrates real and virtual elements, can further enhance mindfulness training by providing real-time feedback and increasing interactivity, which may further boost engagement [72,73]. Recent studies have explored the potential benefits of VR-based mindfulness, showing associations with reduced negative emotions [69,74-76], increased positive emotional states [66,77], and cognitive improvements in attention and working memory [46,78]. Yet, findings have not been entirely consistent, with some studies reporting significant effects on either PA or NA but not both [66,79].

Despite the growing body of research on VR and MR mindfulness interventions, important gaps remain, particularly concerning their effectiveness in brief applications and whether they provide emotional and cognitive benefits comparable with those of more intensive mindfulness practices. Most evidence supporting the cognitive benefits of VR-based mindfulness comes from studies involving multiple sessions, leaving uncertainty about whether shorter interventions yield similar positive outcomes [46,78]. For example, 2 separate studies reported having 8 sessions of the VR mindfulness intervention [46,78]. Moreover, many of these studies face methodological limitations, such as small sample sizes [66] and inadequate or inconsistent control groups [80,81], which constrain the robustness and generalizability of their findings. Research on MR-based mindfulness remains even more limited, despite its potential for enhanced interactivity [73]. These gaps are increasingly salient given recent technological developments, such as the Meta Quest 3 and Apple Vision Pro, which support both VR and MR modes. Emerging mindfulness applications also blend these modalities by embedding virtual cues within real environments or transitioning into full immersion when deeper focus is needed [82]. Thus, addressing these gaps through further empirical investigations is crucial to establishing the efficacy of brief VR and MR mindfulness interventions in enhancing emotional well-being and cognitive functioning and to determining their potential for scalable implementation across diverse real-world settings.

This Study

This study examined whether a brief mindfulness breathing exercise delivered via VR and MR can enhance emotional restoration, attentional functioning, and self-control capacity relative to a mind-wandering control condition. Using a within-subjects experimental design, participants completed both VR and MR mindfulness breathing exercises as well as a mind-wandering audio condition where participants were instructed to let their mind wander freely with no mindfulness advice provided, with a 1-week washout period between conditions to minimize carryover effects. We hypothesized that a brief mindfulness breathing intervention in VR and MR would lead to significant improvements in attentional and emotional restoration and self-control capacity compared with the mind-wandering control condition. Emotional and self-regulatory outcomes were assessed using standardized self-report measures, while working memory capacity was assessed using 2 well-established complex span tasks—the operation span (OSpan) task and the rotation span (RotSpan) task. These tasks were chosen for their strong validity and sensitivity to within-person changes in attentional functioning [83]. Repeated-measures ANOVA was used to examine condition-based differences across emotional, cognitive, and self-regulatory domains.

In this study, we selected PA and NA as immediate, subjective indicators of emotional well-being to capture participants' affective states following mindfulness and VR or MR interventions. Although mindfulness has been shown to enhance several executive functions, including inhibitory control, task-switching, and working memory, we focused on working memory as it represents a central component of executive function and reflects the capacity to maintain goal-relevant information during mindful attention [36]. Previous research suggests that mindfulness training can improve working memory capacity by reducing mind-wandering and enhancing attentional stability, making it an appropriate measure for detecting short-term cognitive changes in a brief intervention [42]. We also conceptualized self-control within the limited resource framework of self-regulation, where momentary acts of regulation, attention control, or emotional management draw from a finite pool of control resources [84]. Because self-control sits at the intersection of executive control and emotional regulation [44], it serves as a meaningful link between our cognitive outcome (working memory) and affective outcomes (PA and NA). Hence, this study aimed to examine the effects of a brief (approximately 15 min) VR and MR mindfulness breathing exercise on attentional and emotional restoration and

XSL•FO
RenderX

self-control capacity, compared with a mind-wandering control condition.

Methods

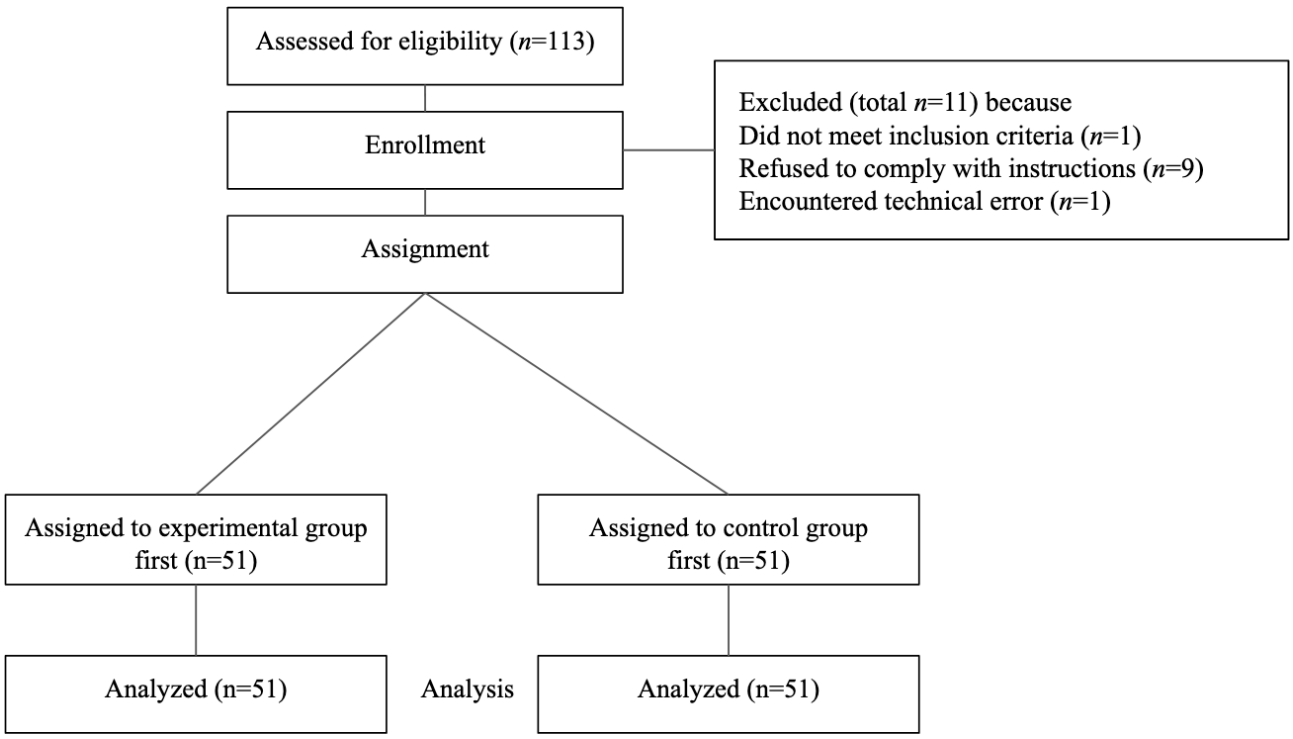
Transparency and Openness

This study's design and analysis plans were preregistered. The details of the preregistration are available on AsPredicted [85]. All preregistration documents and supplementary materials are publicly available on ResearchBox #3410 [86]. The study was reported in accordance with the APA Journal Article Reporting Standards (JARS) [87]. Frequentist analyses were conducted using IBM SPSS 29.0.1.0 [88], while Bayesian analyses were conducted using JASP version 0.17.3 [89] to examine evidence for null effects. The extent and pattern of missing data were examined across all primary outcome variables (PA, NA, self-control capacity, and working memory capacity) using SPSS Missing Value Analysis.

Participants

A total of 113 participants were recruited from a local university in Singapore. All participants were psychology undergraduates recruited using convenience sampling, and they received course credits for their participation. Each participant completed 2 counterbalanced sessions. Data from 11 participants were excluded from the analyses due to noncompliance with task instructions or insufficient immersion during either the experimental or control conditions, as indicated by experimenter observation. These participants appeared distracted or failed to engage consistently with the instructions. Specifically, 1 participant, who reported having dyslexia, was removed due to the potential impact of dyslexia on cognitive performance in the RotSpan task. Additionally, 1 participant encountered a technical error, and 9 participants failed to comply with study instructions. This resulted in a final analytic sample of 102 participants. Participant recruitment, eligibility screening, exclusions, assignment, and the final analyzed sample are summarized in a JARS-adapted flowchart in Figure 1. The demographic characteristics of participants are provided in Table 1.

Figure 1. Journal Article Reporting Standards–adapted participant flowchart depicting recruitment, eligibility screening, exclusions, assignment, and final analyzed sample in a within-subjects experiment.



XSL•FO
RenderX

Table . Demographic characteristics of the sample (N=102) in a within-subjects experiment comparing brief VR[a] and MR[b] mindfulness breathing and a mind-wandering control.

Characteristic	Value	Observed range
Sex (female), n (%)	83 (81.4)	—[c]
Ethnicity (Chinese), n (%)	78 (76.5)	—
Age (y), mean (SD)	20.87 (1.89)	18 - 30
Monthly household income[d], mean (SD)	3.70 (1.49)	1 - 6
Subjective socioeconomic status[e], mean (SD)	6.30 (1.42)	2 - 9

[a]VR: virtual reality.

[b]MR: mixed reality.

[c]Not applicable.

[d]Participants rated their monthly household income on a 6-point scale ranging from less than SGD 2000 to more than SGD 20,000; (1) less than SGD 2000 (US $1570.38); (2) SGD 2000-5999 (US $1570.38-US $4710.34); (3) SGD 6000-9999 (US $4711.13-US $7851.10); (4) SGD 10,000-14,999 (US $7851.88- US $11,777.04); (5) SGD 15,000-19,999 (US $11,777.82-US $15,702.98); (6) more than SGD 20,000 (US $15,703.76).

[e]Participants rated their subjective socioeconomic status using the MacArthur Scale of Subjective Social Status [90], a ladder which represented where people stood in society, and participants had to estimate where one stood on the ladder.

The sample size of this study was determined using an a priori analysis in G*Power 3.1.9.7 [91] to ensure that the study was designed to meet the minimum sample size needed to detect an effect with 80% statistical power. Based on a medium effect size of f=0.25, an α level of .05, and at least 80% power (1–β), the analysis indicated that a minimum of 34 participants would be needed. Additionally, to ensure sufficient statistical power, a post hoc sensitivity power analysis was conducted using the same G*Power settings. This analysis indicated that with our final sample size (n=102), the study was adequately powered to detect the time×condition interaction effects specified by the study design as small as f=0.14 (equivalent to η_p^2=0.019), which represents a small-to-medium effect size. No missing values were observed for any of the primary outcome variables.

Ethical Considerations

All procedures were approved by the university's Institutional Review Board (IRB-24 - 119-A087-M2(1124)) and complied with its ethical guidelines. Participants were informed about the study aims, procedures, potential risks, and their right to withdraw at any time without penalty and provided written informed consent before the first session. To protect privacy and retain confidentiality, participants were identified only by unique identification codes. No personally identifiable information was stored with the research data. All analyses were conducted using deidentified datasets, with data access restricted to the research team. Participants received course credit in exchange for completing both study sessions. No images included in the manuscript or supplementary materials allow for the identification of individual participants. Informed consent was obtained from the individual depicted in Figure 2 for the use of their image in this publication.

XSL•FO
RenderX

Figure 2. Participant wearing a Meta Quest 3 VR headset and hand controllers during the virtual reality and mixed reality mindfulness breathing condition in a within-subjects experiment.

Study Design

This study used a 2 (condition: VR+MR mindfulness breathing vs mind-wandering)×2 (time: baseline vs postintervention) within-subjects experimental design. The within-subjects approach was chosen to control for interpersonal variability, thereby reducing errors associated with individual differences and enhancing the ability to detect true differences between conditions, ultimately increasing statistical power. All participants experienced both the experimental and control conditions. In the experimental condition, participants engaged in a brief VR and MR mindfulness breathing intervention that lasted approximately 15 minutes. In the control condition, they engaged in a mind-wandering audio that lasted approximately 15 minutes. The order of conditions was counterbalanced, with half of the participants completing the experimental condition first and the other half completing the control condition first. To mitigate potential carry-over effects, sessions were separated by a 1-week interval, which served as a washout period to allow any practice or residual effects from the first session to subside [92,93]. Participants underwent random assignment through Qualtrics to establish the sequence in which they experienced the 2 different conditions.

Materials

Emotional Well-Being

State affect was measured using the 18-item Circumplex Model of Affect Scale [94], which evaluates emotional states along 2 independent dimensions—PA and NA. Participants rated their current emotional state on a 5-point Likert scale (1=*Not at all*, 5=*Extremely*) in response to the question, “Overall, how do you feel right now?” The PA scale includes 9 items, that is, energetic, enthusiastic, excited, happy, cheerful, pleasant, calm, content, and relaxed. The NA scale includes 9 items, that is, angry, hostile, irritable, nervous, anxious, tense, dejected, sad, and unhappy. Higher scores indicated greater agreement with the respective states. The PA scale demonstrated good internal consistency at baseline (α_{pre}=.90), after the VR or MR mindfulness condition (α_{post}=.93), and after the mind-wandering control condition (α_{post}=.91). The NA scale demonstrated good internal consistency at baseline (α_{pre}=.90), after the VR or MR mindfulness condition (α_{post}=.88), and after the mind-wandering control condition (α_{post}=.86).

Self-Control Capacity

Self-control capacity was measured using the 5-item Brief State Self-Control Capacity Scale [95]. Participants rated their agreement with each statement (eg, I feel drained now; I feel calm and rational now) on a 7-point Likert scale (1=*Very untrue of me*, 7=*Very true of me*), in response to the question, “Overall, how do you feel right now?” Higher scores indicated greater agreement with the respective states. The scale demonstrated good internal consistency at baseline (α_{pre}=.75), after the VR or MR mindfulness condition (α_{post}=.74), and after the mind-wandering control condition (α_{post}=.78).

XSL•FO
RenderX

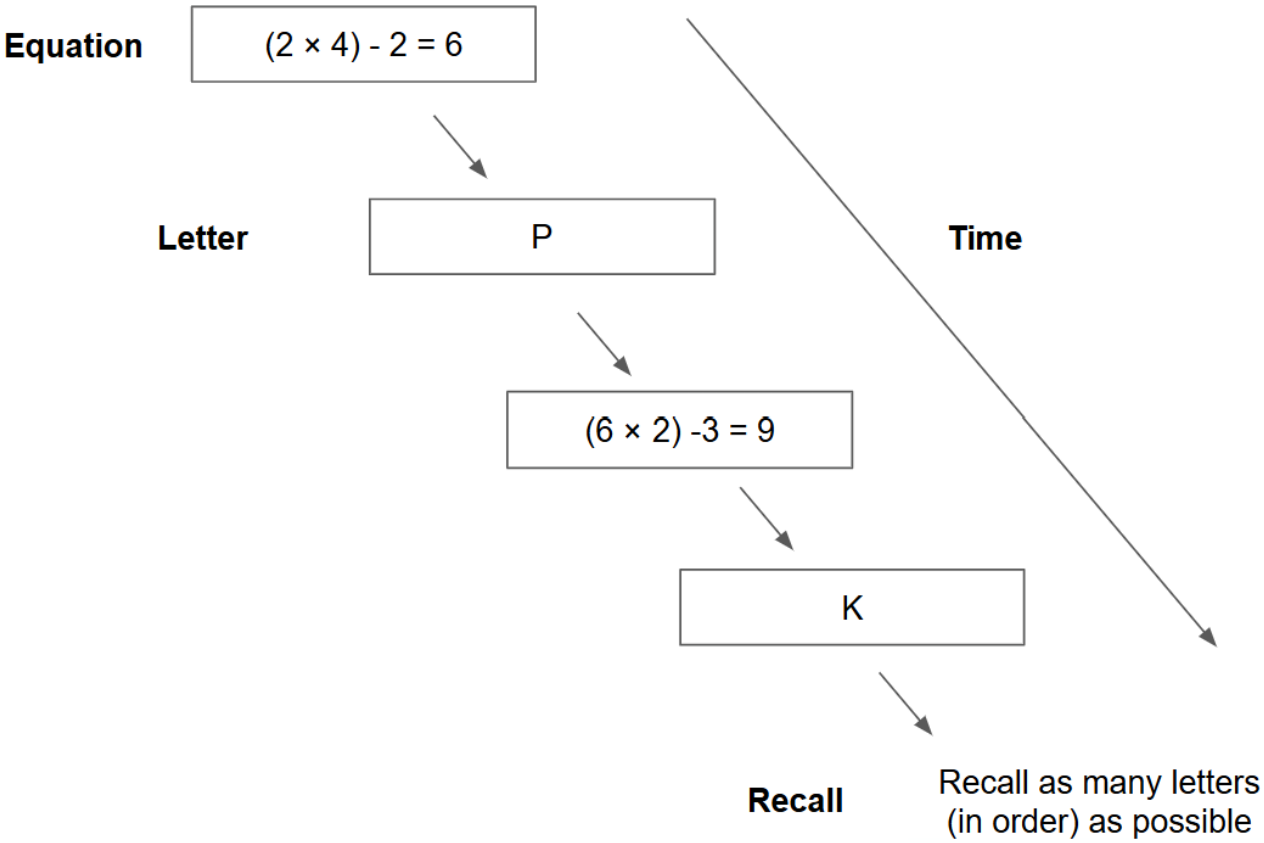
Working Memory Tasks

Working memory capacity was measured using 2 established shortened complex span tasks—OSpan task and the RotSpan task, administered via E-Prime 3.0 [96]. These tasks are accessible on the Attention & Working Memory Lab website [97]. These tasks were selected as performance on complex span tasks has been shown to be malleable and sensitive to experimental manipulation such as training instructions [83]. Although the OSpan (verbal) and RotSpan (spatial) tasks differ in stimulus modality, both tasks are validated measures of domain-general working memory capacity and share the same dual-task structure, which makes them comparable indicators of working memory capacity.

In both tasks, participants were required to memorize a series of items (eg, letters or arrows) while performing interspersed distractor tasks (eg, solving math equations or identifying rotated letters). The number of items to be memorized varied per trial, ranging from 2 to 7, with each sequence length appearing 3 times in a randomized order. To minimize the likelihood of rehearsing memorized items during the distractor task, participants were required to respond to the distractor tasks at a steady pace. Before beginning the main trials, participants completed a practice round for each task to ensure they fully understood the instructions.

Baseline working memory capacity was assessed using the OSpan task. In this task, participants solved mathematical equations while simultaneously memorizing a sequence of letters for later recall (Figure 3). The letters served as items to be remembered, and the math problems acted as distractors. During each trial, participants first solved a math equation and then viewed a letter. This math-letter sequence was repeated between 3 and 7 times, with the number of repetitions varying unpredictably across trials. At the end of each trial, participants were asked to recall the letters in the correct order. Performance was scored using the partial credit unit method, which is based on the number of correctly recalled letters divided by the total number of letters presented within each trial and then averaged across all trials [83]. Each participant completed 2 blocks of trials.

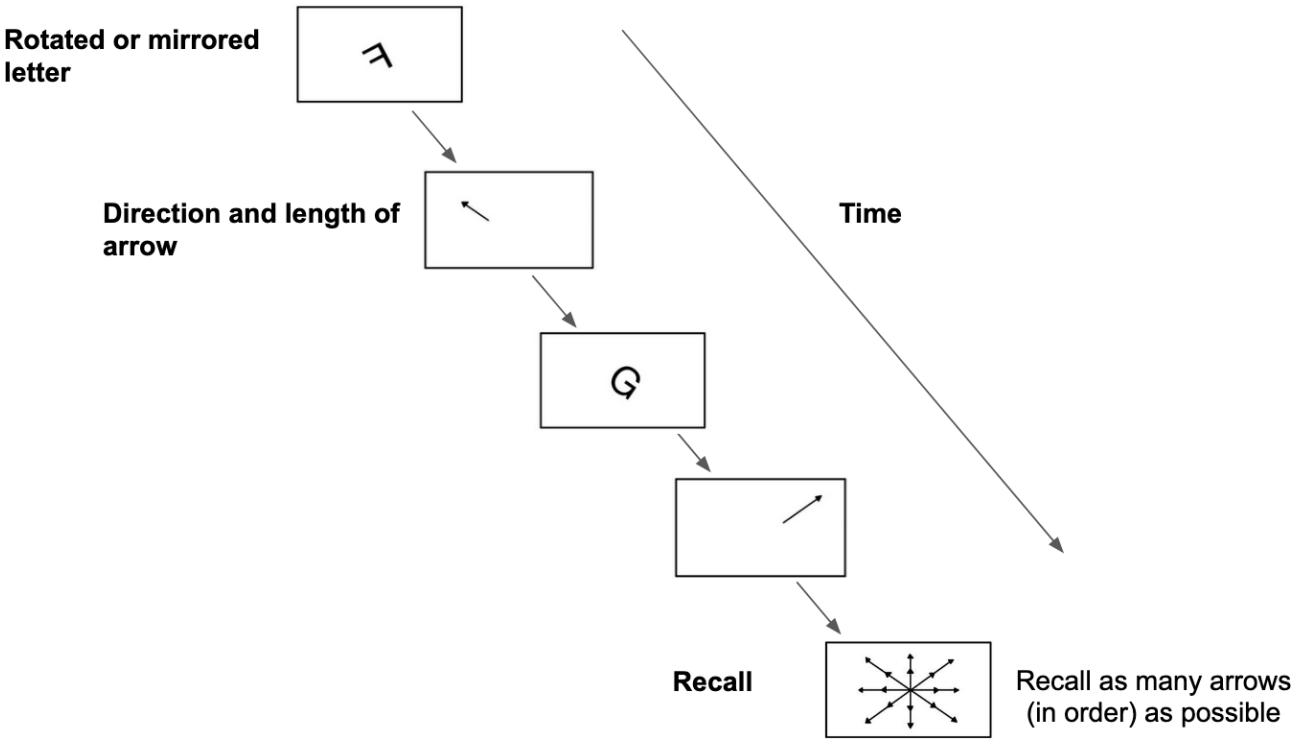
Figure 3. Operation span task flow to assess working memory capacity in an experimental within-subjects trial in undergraduate students in Singapore.



Posttreatment working memory capacity was measured using the RotSpan task, which is similar to the OSpan task but involves different items and operations. In this task, participants judged whether a rotated letter was correctly oriented or mirrored while also memorizing a sequence of arrows, each varying in direction and length (Figure 4). The arrows served as items to be remembered, and the rotated letters acted as distractors. During each trial, participants first judged the orientation of a rotated letter and then viewed an arrow. The rotation-arrow sequence was repeated 2-5 times, with the number of repetitions varying across trials. After each sequence, participants were required to recall the arrows in the correct order. Each participant completed 2 blocks of trials. The partial credit unit score was used to index performance [83].

XSL•FO
RenderX

Figure 4. Rotation span task flow to assess working memory capacity in an experimental within-subjects trial in undergraduate students in Singapore.



Brief VR and MR Mindfulness Intervention

In the VR and MR mindfulness breathing condition, participants engaged in a mindfulness breathing exercise that lasted approximately 15 minutes using the Meta Quest 3 VR headset (Figure 2). Before the mindfulness breathing exercise, the experimenter instructed participants to prepare and explained how to use the controllers for 2 specific functions, which were used to navigate the virtual environments. Participants were informed that their VR screen would be shared on an external monitor (but not recorded) and reminded to follow the guided audio throughout the exercise. The VR and MR mindfulness breathing condition was administered through the Headspace XR app. Headspace XR offers both virtual and MR experiences, featuring mood-enhancing games, personalized guided meditations, and exercises designed to help users improve their mind-body connection through movement and breathing techniques. For this study, participants were tasked to engage in 3 mindfulness breathing activities within Headspace XR.

The first activity, “Take a Beat Portal,” adopts a technique known as “box breathing,” which takes place in an MR setting (Figure 5). In this activity, participants would first find an object or feature in the room that is shaped in a square. Then, they were instructed by the audio to draw a box-shaped virtual overlay that they would be using for the breathing exercise. Once the shape is drawn, participants complete three sets of four controlled breathing cycles. The guided audio included instructions, such as “Inhaling…feeling your chest and stomach expanding [...] Holding…feeling the breath in your body [...] Exhale… feeling the release of any tension [...] Holding… before the next breath.” This entire activity lasts approximately 5 minutes.

XSL•FO
RenderX

Figure 5. Screenshots of the three mindfulness breathing activities used in the virtual reality and mixed reality conditions. The 3 panels depict the mindfulness breathing environments used in the virtual reality and mixed reality conditions. From left to right: Take a Beat Portal, Firefly Treehouse, and Boxy Treehouse.

The second activity, "Firefly Treehouse," is a breathing exercise accompanied by arm movements (Figure 5). This activity takes place in a virtual setting where participants are guided to take deep breaths, raising their arms while inhaling and lowering them during exhaling. The guided audio included directions, such as "Take in a deep breath, and as you inhale raise your hands up above your head [...] Hang there for a moment and as you exhale, bring your hands back to your side." This entire activity lasts approximately 3 minutes.

The third activity, "Boxy Treehouse," is a box breathing exercise similar to the first activity, with the only difference being that it is carried out in a VR setting (Figure 5). In this box breathing exercise, participants would similarly undergo 3 sets of 4 breathing exercises. The guided audio also included similar instructions as the first activity, such as "Inhaling…feeling your chest and stomach expanding [...] Holding…feeling the breath in your body [...] Exhale… feeling the release of any tension [...] Holding… before the next breath." Throughout all 3 activities, participants were instructed to direct their attention to the present moment, feeling their breath, and being aware of each of their breaths. This entire activity lasts approximately 5 minutes.

Mind-Wandering Control

In the mind-wandering condition, participants listened to a 15-minute audio track that allowed for free mind-wandering. Participants listened to the audio using headphones on the desktop. It included instructions, such as "Now we're going to do an exercise for 15 minutes [...] Now simply think about whatever comes to mind, let your mind wander freely without thinking about anything in particular [...] Let your mind roam as it normally would [...] Allow your thoughts to wander wherever they may go [...] Go ahead and follow whatever thoughts that come to mind [...] Continue letting your mind wander, allowing your thoughts to wander wherever they may go." Throughout the mind-wandering audio, there were no mindfulness instructions or interventions given.

Procedures

The within-subject experiment (VR+MR mindfulness breathing vs mind-wandering) consisted of 2 sessions held 1 week apart in the laboratory. Data collection spanned 13 weeks during the semester, and analysis commenced only after the final session was completed. Participants were randomly assigned to either the VR and MR mindfulness breathing experimental condition or the mind-wandering control condition during their first session and assigned to the other condition in the second session, a week later. Throughout the study, the experimenter monitored participants to maintain data quality.

At the start of each session, participants received a link to the Qualtrics survey and were directed to the informed consent page. Participants used a desktop computer with headphones and a VR headset to complete the study. After providing informed consent, each participant generated a unique personal ID to be used across both sessions. Participants began each session by completing baseline measures, including the Circumplex Model of Affect, SMS-5, and OSpan. Following the baseline assessments, participants proceeded to complete activities in their assigned conditions.

After undergoing their respective conditions, a manipulation check was conducted to assess whether the participants were absorbed during the VR and MR mindfulness breathing intervention [42,98,99]. Participants rated the extent to which they felt absorbed in the present moment, focused on their breathing, and the physical sensations of their breathing. The 3 questions were asked using a 7-point Likert scale (1=*Not at all absorbed*, 7=*Extremely absorbed*), where higher scores correspond to greater levels of absorption. A manipulation check consisting of the same questions was conducted for the mind-wandering condition. Finally, they completed posttreatment measures, including the Circumplex Model of Affect, SMS-5, and RotSpan. Participants filled in their demographics at the end of the first session. At the end of the second session, participants were debriefed about the purpose of the study (Figure 6).

XSL•FO
RenderX

Figure 6. Flowchart of the study procedure illustrating the counterbalanced within-subjects design comparing brief virtual reality and mixed reality mindfulness breathing and a mind-wandering control condition. Total sample size was n=102. MR: mixed reality; VR: virtual reality.

XSL•FO
RenderX

Results

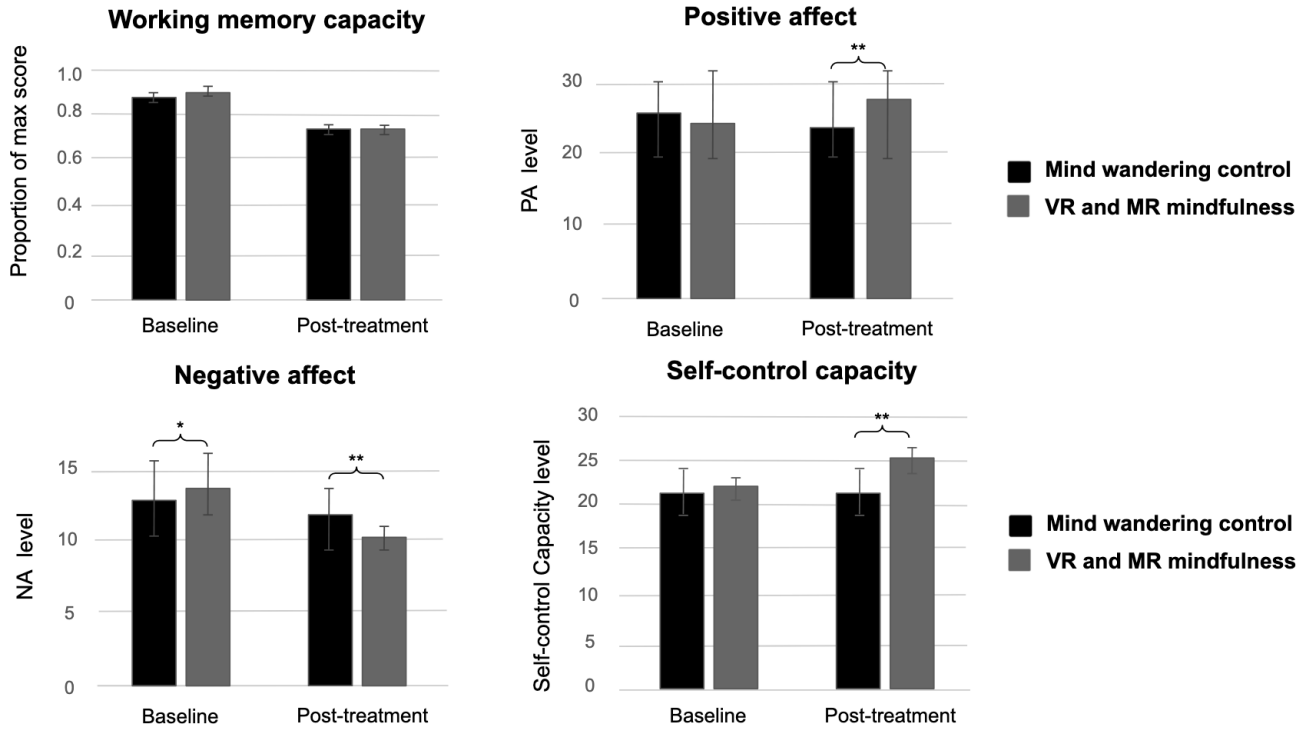
Manipulation Check

In line with our preregistration, we conducted a paired *t* test analysis, which revealed a significant difference in absorption between conditions, t_{101}=−16.29, Cohen *d*=−1.61, 95% CI −1.91 to −1.32; *P*<.001, indicating that participants reported greater absorption in the VR and MR mindfulness breathing condition (mean 15.46, SD 3.32) than in the mind-wandering control condition (mean 9.11, SD 4.07). Bayesian analysis provided extreme evidence in favor of the alternative hypothesis, BF_{10}>100 [100], indicating a substantial difference in absorption levels between conditions.

Working Memory Capacity

Following our preregistered analytic plan, differences in working memory scores between conditions via a 2 (condition: VR+MR mindfulness vs mind-wandering)×2 (time: baseline vs posttreatment) were tested using repeated-measures ANOVA. A significant main effect of time on working memory was observed ($F_{1,101}$=149.34, η_p^2=0.597; *P*<.001). However, there was no significant main effect of condition ($F_{1,101}$=0.343, η_p^2=0.003; *P*=.56). Similarly, the interaction effect between condition and time on working memory capacity was not significant ($F_{1,101}$=1.366, η_p^2=0.013; *P*=.25), and working memory scores differed similarly between baseline (mean 0.90, SD 0.13) and post treatment (mean 0.74, SD 0.18) in the VR and MR mindfulness breathing condition, and between baseline (mean 0.88, SD 0.15) and post treatment (mean 0.74, SD 0.17) in the mind-wandering control condition (Figure 7). This indicates that the difference in working memory performance between baseline and post treatment did not differ between the VR and MR mindfulness breathing condition and the mind-wandering control condition. Bayesian analysis further provided moderate evidence in support of the null hypothesis for both the main effect of condition and the time×condition interaction (BF_{01}=7.42 and BF_{01}=5.55, respectively), suggesting that VR and MR mindfulness breathing did not produce measurable improvements in working memory capacity compared to the control condition.

Figure 7. Working memory capacity (proportion of maximum score), positive affect, negative affect, and self-control capacity at baseline and post treatment across mind-wandering control and virtual reality and mixed reality mindfulness breathing conditions in a within-subjects experiment. Values indicate mean scores with error bars depicting SEs, which are consistent with the frequentist framework used for hypothesis testing. Total sample size was n=102. MR: mixed reality; NA: negative affect; VR: virtual reality.



PA

Consistent with our preregistered analytic plan, repeated-measures ANOVA was conducted to examine whether levels of PA were influenced by time and condition. First, no significant main effect of time on PA was observed after Bonferroni correction ($F_{1,101}$=5.367, η_p^2=0.050; *P*=.02). This indicates that PA did not change significantly across time, regardless of the condition. However, a significant main effect of the condition on PA was found, $F_{1,101}$=13.278, η_p^2=0.116; *P*<.001, indicating that the VR and MR mindfulness breathing condition showed significantly higher levels of PA (mean 26.47, SD 6.73) than the mind-wandering control condition (mean 24.20, SD 5.95). A Bonferroni correction was applied to account for multiple comparisons. With a corrected significance threshold of *P*<.0167 (ie, .05/3) for 3 comparisons, the main effect remained significant.

Most importantly, the interaction effect between condition and time on PA was significant, $F_{1,101}$=58.30, η_p^2=0.366; *P*<.001, suggesting that the changes in PA scores over time differed

XSL•FO
RenderX

significantly between the VR and MR mindfulness breathing condition and the mind-wandering control condition. Specifically, PA scores increased from baseline (mean 24.59, SD 6.74) to posttreatment (mean 28.35, SD 7.63) in the VR and MR mindfulness breathing condition, whereas scores in the mind-wandering control condition declined from baseline (mean 25.03, SD 6.24) to posttreatment (mean 23.36, SD 7.25), reflecting a statistically significant difference between conditions (Figure 7). This effect remained significant under the Bonferroni-corrected threshold (adjusted P<.0167). Bayesian analysis further provided extreme evidence supporting the alternative hypothesis (BF_{10}>1000, BF_{01}<0.001). The 95% credible intervals for the cell means indicated that PA decreased in the control condition from 25.03 (23.80-26.26) at baseline to 23.36 (21.94-24.79) post treatment, whereas it increased in the VR and MR mindfulness condition from 24.59 (23.26-25.91) to 28.35 (26.86-29.85).

NA

A similar analytical approach was used to examine changes in NA. A significant main effect of time on NA was observed, $F_{1,101}$=62.619, η_p^2=0.383; P<.001, suggesting that NA changed significantly from baseline to post treatment. While there was no main effect of the condition on NA, $F_{1,101}$=1.384, η_p^2=0.014; P=.24, the interaction effect between condition and time on NA was significant ($F_{1,101}$=39.087, η_p^2=0.279; P<.001). The results showed that the change in NA scores over time was significantly different between the VR and MR mindfulness breathing condition and the mind-wandering control condition. Specifically, NA scores declined from baseline (mean 14.15, SD 4.84) to post treatment (mean 10.58, SD 2.74) in the VR and MR mindfulness breathing condition, whereas the decline was less pronounced in the mind-wandering control condition from baseline (mean 13.15, SD 5.12) to post treatment (mean 12.36, SD 4.27). This effect remained statistically significant under the Bonferroni-corrected threshold (adjusted P<.0167). Bayesian analysis also provided extreme evidence in favor of the alternative hypothesis (BF_{10}>1000, BF_{01}<0.001). The 95% credible intervals indicated that NA decreased modestly in the control condition, from 13.15 (12.14-14.15) at baseline to 12.36 (11.52-13.20) post treatment, but showed a substantially larger decrease in the VR and MR mindfulness condition from 14.15 (13.20-15.10) to 10.58 (10.04-11.12).

Self-Control Capacity

Similar analyses were conducted to examine changes in self-control capacity. A significant main effect of time was observed, $F_{1,101}$=15.480, η_p^2=0.133; P<.001, indicating that self-control capacity changed significantly from baseline to post treatment. Moreover, a significant main effect of condition was found, $F_{1,101}$=24.910, η_p^2=0.198; P<.001, and remained statistically significant after the Bonferroni correction. Consistently, the interaction effect between condition and time on self-control capacity was also statistically significant ($F_{1,101}$=28.348, η_p^2=0.219; P<.001). The results showed that self-control capacity scores increased from baseline (mean 22.09, SD 4.60) to post treatment (mean 25.25, SD 4.42) in the VR and MR mindfulness breathing condition, whereas scores in the mind-wandering control condition decreased from baseline (mean 21.77, SD 5.14) to post treatment (mean 21.31, SD 5.46). Bayesian analysis further provided extreme evidence in favor of the alternative hypothesis (BF_{10}>1000, BF_{01}<0.001). Collectively, these findings suggest that exposure to VR and MR mindfulness breathing was associated with an improvement in self-control capacity.

Discussion

Despite the growing body of research on mindfulness, much of the existing literature focuses on intensive, multisession interventions [46,78], leaving a gap in our understanding of the effectiveness of brief mindfulness practices. Some evidence suggests that even a single session of mindfulness can produce temporary benefits, but findings remain inconsistent, highlighting the need for more engaging and immersive approaches [42,60-62]. This study examined whether immersive VR technologies could enhance the effectiveness of brief mindfulness practices, potentially offering a more engaging and effective alternative. Brief mindfulness interventions often face challenges, such as limited attentional depth, insufficient repetition to consolidate self-regulatory skills, and inconsistent environmental contexts, that hinder transfer to daily life [66]. VR and MR-based delivery may help address these limitations by providing immersive environments that minimize distraction, provide sustained attentional focus, and simulate realistic emotional situations for practicing acceptance [66-68].

Consistent with our hypothesis, both VR and MR mindfulness breathing improved emotional well-being by increasing PA and reducing NA. These findings align with previous research showing that mindfulness breathing enhances PA [7,13,14] while alleviating depressive symptoms [7,101], stress [10,20,21], and anxiety [22,23,101]. According to the monitor and acceptance theory [56], mindfulness promotes emotion regulation through two core components: (1) attention monitoring, which heightens present-moment awareness; and (2) acceptance, which enables individuals to observe emotions without reacting impulsively. The immersive nature of VR and MR may further strengthen these processes by increasing engagement, reducing external distractions, and fostering a heightened sense of presence during mindfulness practice [66,102]. By creating a more absorbing and controlled environment, VR and MR may deepen both attention monitoring and acceptance, leading to greater emotional benefits [51,69]. It is also important to note that mindfulness aims to cultivate adaptive flexibility [34], allowing individuals to experience and regulate both positive and negative emotions in a context-sensitive manner. Our findings suggest that even a brief 15-minute mindfulness breathing session can enhance emotional well-being, adding to evidence that short-duration mindfulness interventions can yield immediate psychological benefits. Overall, these findings underscore the potential of VR and MR as effective tools for mood enhancement by strengthening core mindfulness processes. However, it is also possible that part of the emotional improvements observed may partly reflect the novelty and sensory engagement of immersive VR itself rather than mindfulness-specific mechanisms. Previous research shows

XSL•FO
RenderX

that immersive VR can induce relaxation and PA through heightened presence and environmental realism, even without formal mindfulness instruction [67]. Thus, the affective benefits observed in this study may reflect a combination of novelty, attentional engagement, and mindfulness-related processes.

Beyond its benefits for emotional well-being, the study found that VR and MR mindfulness breathing also enhanced self-control capacity. Participants in the VR and MR conditions demonstrated greater self-control than those in the mind-wandering control condition, consistent with previous research showing that brief mindfulness practice can counteract self-control depletion [103,104]. This effect may stem from mindfulness enhancing 2 fundamental self-control processes: (1) emotion regulation [17-19] and (2) attention regulation [46,47,105], both of which are crucial for maintaining self-regulatory capacity [44,48]. The immersive qualities of VR and MR may further amplify these effects by deepening attentional absorption, minimizing external distractions, and alleviating cognitive strain. By reducing cognitive load, VR and MR mindfulness exercises may help preserve the mental resources necessary for self-control, thereby reinforcing self-regulatory functioning [44]. These findings contribute to the growing literature on mindfulness and self-control regulation and suggest that brief VR and MR mindfulness interventions could be useful for mitigating self-regulatory depletion [35,106]. Such interventions may be particularly suitable for individuals seeking to restore attentional and self-control resources during breaks in cognitively demanding environments, including workplaces and educational settings. By stabilizing attentional focus and alleviating cognitive strain, these interventions may enhance both productivity and well-being [52].

Contrary to our hypothesis, we found no significant difference in working memory between the VR and MR mindfulness and mind-wandering control conditions. This contrasts with previous research showing that mindfulness training can enhance working memory [32,33], suggesting potential limits to the immediate cognitive benefits of brief interventions in immersive VR and MR formats. These results highlight the need for caution when interpreting short-term cognitive outcomes following minimal exposure and are consistent with evidence that improvements in working memory typically emerge only after prolonged, repeated mindfulness training over multiple sessions [62,92]. Future research should therefore investigate whether extended VR and MR mindfulness programs yield cumulative cognitive gains as participants become more familiar and engaged with the practice.

Several factors may help to explain the limited cognitive benefits observed in our study. First, intervention duration and intensity are critical. Previous research indicates that mindfulness-based interventions typically require multisession practice to yield measurable cognitive benefits [33,42]. For example, studies in which participants engaged in 30 - 45 minutes of mindfulness training over several weeks have reported measurable improvements in working memory [33,42]. In contrast, the single 15-minute session in our study may not have provided sufficient opportunity for participants to fully engage with the mindfulness practice or experience cognitive benefits. Moreover, working memory capacity is relatively stable and tends to change gradually rather than in response to brief interventions [107,108]. Given that working memory relies on complex neural processes that require continuous reinforcement [109,110], longer training durations may be necessary to induce measurable improvements. This account aligns with evidence that the cognitive benefits of mindfulness depend on repeated and prolonged engagement, which strengthens underlying attentional and working memory mechanisms [4,32,54,55]. Future research should therefore explore whether extended VR and MR mindfulness interventions produce cumulative cognitive gains over time. It will be important to determine whether repeated practice enhances working memory or whether habituation results in reduced engagement. Clarifying these long-term dynamics is crucial for evaluating the feasibility of VR and MR-based mindfulness as a long-term cognitive intervention.

An alternative explanation for the lower working memory scores observed post treatment relative to baseline is the influence of uncontrolled stressors or fatigue during the study period. Academic workload, sleep disruption, or daily emotional stress may have temporarily impaired prefrontal functioning and reduced attentional capacity during testing [111-113]. Previous research indicates that acute stress can disrupt working memory by diverting cognitive resources from executive processes [114]. Repeated testing may also have contributed to cognitive fatigue, particularly given the demanding nature of the complex span tasks [109]. The comparable pattern observed across both conditions suggests that these differences reflect general influences rather than effects of the experimental manipulation itself.

Beyond intervention duration and intensity, the environmental context of mindfulness practice may also influence cognitive outcomes. While previous studies have often used nature-based settings (eg, forests, rivers, and beaches) for mindful practice, which have been known to promote relaxation and reduce cognitive load [66,115], our study used VR metaverse and MR environments. Although VR and MR provide immersive experiences, they may also introduce additional cognitive demands, such as the need to navigate virtual interfaces, which could detract from mindfulness engagement, particularly for individuals unfamiliar with the technology [82]. This may help explain why VR and MR did not yield measurable cognitive improvements; compared with natural environments, immersive virtual settings may not provide the restorative conditions necessary for cognitive recovery [92], potentially contributing to the null findings in our study.

While this study highlights the potential psychological benefits of VR and MR mindfulness interventions, several limitations should be considered. First, future research should include more diverse samples to enhance generalizability beyond young adults. Younger participants may differ from older adults or clinical populations in their adaptability to emerging technology and stress profiles [78]. Including individuals from a wider range of age groups, occupational backgrounds, and mental health conditions would provide a more comprehensive understanding of the effectiveness of VR and MR mindfulness interventions.

XSL•FO
RenderX

Second, the working memory assessment in this study may be subject to task impurity [116,117]. We used the OSpan task at baseline and the RotSpan task post treatment, which engage different cognitive processes (eg, verbal vs spatial processing) [83]. Relying on only 2 complex-span tasks may have provided an incomplete evaluation of working memory capacity, making the null findings potentially task-specific rather than reflective of actual cognitive changes. In addition, we did not include tasks from other paradigms, such as N-back or updating tasks, which assess complementary components of working memory (eg, tracking, updating, and replacing information) [108]. Future research should incorporate a broader set of complex span tasks to improve construct validity.

Third, the study did not account for individual differences in stress exposure, fatigue, or sleep quality across sessions. Variability in these uncontrolled factors may have contributed to fluctuations in working memory performance, underscoring the need for future studies to monitor or standardize participants' stress levels across testing periods [118]. Another limitation concerns participants' familiarity with VR technology. Because our sample primarily consisted of young adults, their higher levels of technological comfort may have facilitated immersion and engagement during the intervention. These effects may not generalize to older adults or clinical populations with limited VR experience. Future research should therefore assess VR familiarity as a potential moderator and examine whether individual differences in technological comfort influence intervention outcomes.

Finally, the absence of a VR-based control condition (ie, a nonmindfulness VR control condition) presents another limitation. We selected a mind-wandering task as the control condition to ensure that participants engaged in a cognitively active but nonmindful activity, thereby providing a clearer contrast with the mindfulness intervention [43]. Although this design differentiates mindfulness-specific effects from general cognitive engagement, it does not fully disentangle the potential contribution of VR immersion itself from those of mindfulness practice. Future research should include an appropriate VR control condition to clarify the distinct contributions of immersion and mindfulness.

In summary, this study examined the effects of VR and MR mindfulness on cognitive and emotional outcomes relative to a mind-wandering control condition. While no cognitive improvements were observed, the findings underscore the potential of VR and MR mindfulness in enhancing emotional well-being, consistent with research suggesting that immersive environments may facilitate psychological benefits [71,79]. In comparison to existing studies that have primarily focused on stress reduction or relied on multisession digital interventions [46,78], this study demonstrates that a single brief VR or MR session can enhance key emotional and self-regulatory outcomes. As such, these findings underscore the potential of VR and MR mindfulness applications as innovative and scalable tools for promoting mental well-being. As these technologies become increasingly incorporated into mental health interventions, future research should investigate their long-term efficacy and applicability across diverse populations and identify the specific features, such as sensory immersion, attentional focus, and guided engagement, that optimize their therapeutic potential. Furthermore, the declining cost and increasing accessibility of stand-alone VR headsets (eg, Meta Quest) increase the feasibility of implementing immersive mindfulness programs in organizational and educational settings [68]. Despite initial setup and training demands, these interventions can be cost-effective over time because the same digital content can be reused across participants with minimal supervision. Their scalability and portability make them suitable for brief restorative sessions that enhance attention and emotional well-being [92]. Future research should therefore examine implementation feasibility, including usability, cost-benefit considerations, and long-term adherence, to determine the broader applicability of VR- and MR-based mindfulness interventions. By leveraging the unique affordances of immersive technology, VR and MR mindfulness may provide innovative and accessible approaches to enhancing emotional regulation and well-being across clinical, educational, and workplace settings.

Acknowledgments

A Generative Artificial Intelligence tool (i.e., ChatGPT) was used only for refinement, correction, editing, or formatting of the manuscript to improve clarity of language.

Funding

This research was supported by grants awarded to Andree Hartanto by Singapore Management University through research grants from the Ministry of Education Academy Research Fund Tier 1 (25-SOSS-SMU-006) and Lee Kong Chian Fund for Research Excellence. The funder had no involvement in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

ZKYE: Methodology, Writing – original draft, Writing – review & editing.
CKJ: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing.
HY: Writing – original draft, Writing – review & editing.

XSL•FO
RenderX

AYHG: Conceptualization, Methodology.
MH: Conceptualization, Methodology.
KTASK: Conceptualization, Methodology.
AH: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing.

Conflicts of Interest

None declared.

References

1. Hayes AM, Feldman G. Clarifying the construct of mindfulness in the context of emotion regulation and the process of change in therapy. Clinical Psychology: Science and Practice 2004;11(3):255-262. [doi: 10.1093/clipsy.bph080]
2. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. Clinical Psychology: Science and Practice 2003;10(2):144-156. [doi: 10.1093/clipsy.bpg016]
3. Khoury B, Lecomte T, Fortin G, et al. Mindfulness-based therapy: a comprehensive meta-analysis. Clin Psychol Rev 2013 Aug;33(6):763-771. [doi: 10.1016/j.cpr.2013.05.005] [Medline: 23796855]
4. Zeidan F, Johnson SK, Diamond BJ, David Z, Goolkasian P. Mindfulness meditation improves cognition: evidence of brief mental training. Conscious Cogn 2010 Jun;19(2):597-605. [doi: 10.1016/j.concog.2010.03.014] [Medline: 20363650]
5. Goldberg SB, Tucker RP, Greene PA, et al. Mindfulness-based interventions for psychiatric disorders: a systematic review and meta-analysis. Clin Psychol Rev 2018 Feb;59:52-60. [doi: 10.1016/j.cpr.2017.10.011] [Medline: 29126747]
6. Strauss C, Cavanagh K, Oliver A, Pettman D, editor. Mindfulness-based interventions for people diagnosed with a current episode of an anxiety or depressive disorder: a meta-analysis of randomised controlled trials. PLoS ONE 2014;9(4):e96110. [doi: 10.1371/journal.pone.0096110] [Medline: 24763812]
7. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003 Apr;84(4):822-848. [doi: 10.1037/0022-3514.84.4.822] [Medline: 12703651]
8. Hollis-Walker L, Colosimo K. Mindfulness, self-compassion, and happiness in non-meditators: a theoretical and empirical examination. Pers Individ Dif 2011 Jan;50(2):222-227. [doi: 10.1016/j.paid.2010.09.033]
9. Schutte NS, Malouff JM. Emotional intelligence mediates the relationship between mindfulness and subjective well-being. Pers Individ Dif 2011 May;50(7):1116-1119. [doi: 10.1016/j.paid.2011.01.037]
10. Grossman P, Niemann L, Schmidt S, Walach H. Mindfulness-based stress reduction and health benefits. A meta-analysis. J Psychosom Res 2004 Jul;57(1):35-43. [doi: 10.1016/S0022-3999(03)00573-7] [Medline: 15256293]
11. Fredrickson BL, Cohn MA, Coffey KA, Pek J, Finkel SM. Open hearts build lives: positive emotions, induced through loving-kindness meditation, build consequential personal resources. J Pers Soc Psychol 2008 Nov;95(5):1045-1062. [doi: 10.1037/a0013262] [Medline: 18954193]
12. Kiken LG, Shook NJ. Looking up: mindfulness increases positive judgments and reduces negativity bias. Soc Psychol Personal Sci 2011 Jul;2(4):425-431. [doi: 10.1177/1948550610396585]
13. Garland EL, Fredrickson B, Kring AM, Johnson DP, Meyer PS, Penn DL. Upward Spirals of Positive Emotions Counter Downward Spirals of Negativity: Insights from the Broaden-and-Build Theory and Affective Neuroscience on the Treatment of Emotion Dysfunctions and Deficits in Psychopathology: The University of North Carolina at Chapel Hill University Libraries; 2010. [doi: 10.17615/8TFX-YY95]
14. Geschwind N, Peeters F, Drukker M, van Os J, Wichers M. Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression: a randomized controlled trial. J Consult Clin Psychol 2011 Oct;79(5):618-628. [doi: 10.1037/a0024595] [Medline: 21767001]
15. Cooper D, Yap K, Batalha L. Mindfulness-based interventions and their effects on emotional clarity: a systematic review and meta-analysis. J Affect Disord 2018 Aug 1;235:265-276. [doi: 10.1016/j.jad.2018.04.018] [Medline: 29660642]
16. Wu R, Liu LL, Zhu H, et al. Brief mindfulness meditation improves emotion processing. Front Neurosci 2019;13:1074. [doi: 10.3389/fnins.2019.01074] [Medline: 31649501]
17. Chambers R, Gullone E, Allen NB. Mindful emotion regulation: an integrative review. Clin Psychol Rev 2009 Aug;29(6):560-572. [doi: 10.1016/j.cpr.2009.06.005] [Medline: 19632752]
18. Hölzel BK, Lazar SW, Gard T, Schuman-Olivier Z, Vago DR, Ott U. How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. Perspect Psychol Sci 2011 Nov;6(6):537-559. [doi: 10.1177/1745691611419671] [Medline: 26168376]
19. Teper R, Segal ZV, Inzlicht M. Inside the mindful mind: how mindfulness enhances emotion regulation through improvements in executive control. Curr Dir Psychol Sci 2013 Dec;22(6):449-454. [doi: 10.1177/0963721413495869]
20. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J Altern Complement Med 2009 May;15(5):593-600. [doi: 10.1089/acm.2008.0495] [Medline: 19432513]
21. Sharma M, Rush SE. Mindfulness-based stress reduction as a stress management intervention for healthy individuals: a systematic review. J Evid Based Complementary Altern Med 2014 Oct;19(4):271-286. [doi: 10.1177/2156587214543143] [Medline: 25053754]

XSL•FO
RenderX

22. Evans S. Mindfulness-based cognitive therapy for generalized anxiety disorder. In: Eisendrath SJ, editor. Mindfulness-Based Cogn Ther: Springer International Publishing; 2016:145-154. [doi: 10.1007/978-3-319-29866-5_13]
23. Vøllestad J, Nielsen MB, Nielsen GH. Mindfulness- and acceptance-based interventions for anxiety disorders: a systematic review and meta-analysis. Br J Clin Psychol 2012 Sep;51(3):239-260. [doi: 10.1111/j.2044-8260.2011.02024.x] [Medline: 22803933]
24. Farb NAS, Anderson AK, Mayberg H, Bean J, McKeon D, Segal ZV. "Minding one's emotions: Mindfulness training alters the neural expression of sadness": Correction to Farb et al (2010). Emotion 2010 Apr;10(2):215-215. [doi: 10.1037/a0019263]
25. Goldin PR, Gross JJ. Effects of mindfulness-based stress reduction (MBSR) on emotion regulation in social anxiety disorder. Emotion 2010 Feb;10(1):83-91. [doi: 10.1037/a0018441] [Medline: 20141305]
26. Keng SL, Smoski MJ, Robins CJ. Effects of mindfulness on psychological health: a review of empirical studies. Clin Psychol Rev 2011 Aug;31(6):1041-1056. [doi: 10.1016/j.cpr.2011.04.006] [Medline: 21802619]
27. Bishop SR, Lau M, Shapiro S, et al. Mindfulness: a proposed operational definition. Clinical Psychology: Science and Practice 2004;11(3):230-241. [doi: 10.1093/clipsy.bph077]
28. Shapiro SL, Carlson LE, Astin JA, Freedman B. Mechanisms of mindfulness. J Clin Psychol 2006 Mar;62(3):373-386. [doi: 10.1002/jclp.20237] [Medline: 16385481]
29. Economides M, Martman J, Bell MJ, Sanderson B. Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone spp: a randomized controlled trial. Mindfulness (N Y) 2018;9(5):1584-1593. [doi: 10.1007/s12671-018-0905-4] [Medline: 30294390]
30. Feldman G, Greeson J, Senville J. Differential effects of mindful breathing, progressive muscle relaxation, and loving-kindness meditation on decentering and negative reactions to repetitive thoughts. Behav Res Ther 2010 Oct;48(10):1002-1011. [doi: 10.1016/j.brat.2010.06.006] [Medline: 20633873]
31. Hoge EA, Bui E, Goetter E, et al. Change in decentering mediates improvement in anxiety in mindfulness-based stress reduction for generalized anxiety disorder. Cogn Ther Res 2015 Apr;39(2):228-235. [doi: 10.1007/s10608-014-9646-4]
32. Chambers R, Lo BCY, Allen NB. The impact of intensive mindfulness training on attentional control, cognitive style, and affect. Cogn Ther Res 2008 Jun;32(3):303-322. [doi: 10.1007/s10608-007-9119-0]
33. Jha AP, Stanley EA, Kiyonaga A, Wong L, Gelfand L. Examining the protective effects of mindfulness training on working memory capacity and affective experience. Emotion 2010 Feb;10(1):54-64. [doi: 10.1037/a0018438] [Medline: 20141302]
34. Moore A, Malinowski P. Meditation, mindfulness and cognitive flexibility. Conscious Cogn 2009 Mar;18(1):176-186. [doi: 10.1016/j.concog.2008.12.008] [Medline: 19181542]
35. Tang YY, Ma Y, Wang J, et al. Short-term meditation training improves attention and self-regulation. Proc Natl Acad Sci U S A 2007 Oct 23;104(43):17152-17156. [doi: 10.1073/pnas.0707678104] [Medline: 17940025]
36. Black DS, Semple RJ, Pokhrel P, Grenard JL. Component processes of executive function-mindfulness, self-control, and working memory-and their relationships with mental and behavioral health. Mindfulness (N Y) 2011 Sep;2(3):179-185. [doi: 10.1007/s12671-011-0057-2] [Medline: 22013495]
37. Chan RCK, Shum D, Toulopoulou T, Chen EYH. Assessment of executive functions: review of instruments and identification of critical issues. Arch Clin Neuropsychol 2008 Mar;23(2):201-216. [doi: 10.1016/j.acn.2007.08.010] [Medline: 18096360]
38. Hartanto A, Yang H. The role of bilingual interactional contexts in predicting interindividual variability in executive functions: a latent variable analysis. J Exp Psychol Gen 2020 Apr;149(4):609-633. [doi: 10.1037/xge0000672] [Medline: 31486665]
39. Gallant SN. Mindfulness meditation practice and executive functioning: breaking down the benefit. Conscious Cogn 2016 Feb;40:116-130. [doi: 10.1016/j.concog.2016.01.005] [Medline: 26784917]
40. Holas P, Jankowski T. A cognitive perspective on mindfulness. Int J Psychol 2013;48(3):232-243. [doi: 10.1080/00207594.2012.658056] [Medline: 22506609]
41. Lutz A, Slagter HA, Dunne JD, Davidson RJ. Attention regulation and monitoring in meditation. Trends Cogn Sci 2008 Apr;12(4):163-169. [doi: 10.1016/j.tics.2008.01.005] [Medline: 18329323]
42. Mrazek MD, Franklin MS, Phillips DT, Baird B, Schooler JW. Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. Psychol Sci 2013 May;24(5):776-781. [doi: 10.1177/0956797612459659] [Medline: 23538911]
43. Vago DR, Zeidan F. The brain on silent: mind wandering, mindful awareness, and states of mental tranquility. Ann N Y Acad Sci 2016 Jun;1373(1):96-113. [doi: 10.1111/nyas.13171] [Medline: 27398642]
44. Baumeister RF, Vohs KD, Tice DM. The strength model of self-control. Curr Dir Psychol Sci 2007 Dec;16(6):351-355. [doi: 10.1111/j.1467-8721.2007.00534.x]
45. Nigg JT. Annual research review: on the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. J Child Psychol Psychiatry 2017 Apr;58(4):361-383. [doi: 10.1111/jcpp.12675] [Medline: 28035675]
46. Argüero-Fonseca A, Marchioro DM, López-Beltrán I. Effect of a mindfulness intervention with virtual reality in adolescents on attention and working memory. Journal of Positive Psychology & Wellbeing 2022;6(1):1810-1830 [FREE Full text]

XSL•FO
RenderX

47. Hodgins HS, Adair KC. Attentional processes and meditation. Conscious Cogn 2010 Dec;19(4):872-878. [doi: 10.1016/j.concog.2010.04.002] [Medline: 20430650]
48. Tice DM, Bratslavsky E. Giving in to feel good: the place of emotion regulation in the context of general self-control. Psychol Inq 2000 Jul;11(3):149-159. [doi: 10.1207/S15327965PLI1103_03]
49. Kabat-Zinn J. Full Catastrophe Living, Revised Edition: How to Cope with Stress, Pain and Illness Using Mindfulness Meditation: Hachette UK; 2013.
50. Segal Z, Williams M, Teasdale J. Mindfulness-Based Cognitive Therapy for Depression: Guilford press; 2012.
51. Orzech KM, Shapiro SL, Brown KW, McKay M. Intensive mindfulness training-related changes in cognitive and emotional experience. J Posit Psychol 2009 May 1;4(3):212-222. [doi: 10.1080/17439760902819394]
52. Goyal M, Singh S, Sibinga EMS, et al. Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. JAMA Intern Med 2014 Mar;174(3):357-368. [doi: 10.1001/jamainternmed.2013.13018] [Medline: 24395196]
53. Ostafin BD, Chawla N, Bowen S, Dillworth TM, Witkiewitz K, Marlatt GA. Intensive mindfulness training and the reduction of psychological distress: a preliminary study. Cogn Behav Pract 2006 Aug;13(3):191-197. [doi: 10.1016/j.cbpra.2005.12.001]
54. Charness G, Le Bihan Y, Villeval MC. Mindfulness training, cognitive performance and stress reduction. J Econ Behav Organ 2024 Jan;217:207-226. [doi: 10.1016/j.jebo.2023.10.027]
55. Schanche E, Vøllestad J, Binder PE, et al. Can clinical psychology students benefit from brief and intensive mindfulness training? Couns and Psychother Res 2020 Jun;20(2):311-324. [doi: 10.1002/capr.12273]
56. Lindsay EK, Creswell JD. Mechanisms of mindfulness training: monitor and acceptance theory (MAT). Clin Psychol Rev 2017 Feb;51:48-59. [doi: 10.1016/j.cpr.2016.10.011] [Medline: 27835764]
57. Schumer MC, Lindsay EK, Creswell JD. Brief mindfulness training for negative affectivity: a systematic review and meta-analysis. J Consult Clin Psychol 2018 Jul;86(7):569-583. [doi: 10.1037/ccp0000324] [Medline: 29939051]
58. Gilmartin H, Goyal A, Hamati MC, Mann J, Saint S, Chopra V. Brief mindfulness practices for healthcare providers - a systematic literature review. Am J Med 2017 Oct;130(10):1219. [doi: 10.1016/j.amjmed.2017.05.041] [Medline: 28687263]
59. Howarth A, Smith JG, Perkins-Porras L, Ussher M. Effects of brief mindfulness-based interventions on health-related outcomes: a systematic review. Mindfulness (N Y) 2019 Oct;10(10):1957-1968. [doi: 10.1007/s12671-019-01163-1]
60. Arch JJ, Craske MG. Mechanisms of mindfulness: emotion regulation following a focused breathing induction. Behav Res Ther 2006 Dec;44(12):1849-1858. [doi: 10.1016/j.brat.2005.12.007] [Medline: 16460668]
61. Somaraju LH, Temple EC, Bizo LA, Cocks B. Brief mindfulness meditation: can it make a real difference? Curr Psychol 2023 Mar;42(7):5530-5542. [doi: 10.1007/s12144-021-01897-z]
62. Quek FYX, Majeed NM, Kothari M, Lua VYQ, Ong HS, Hartanto A. Brief mindfulness breathing exercises and working memory capacity: findings from two experimental approaches. Brain Sci 2021 Feb 1;11(2):175. [doi: 10.3390/brainsci11020175] [Medline: 33535395]
63. Creswell JD. Mindfulness interventions. Annu Rev Psychol 2017 Jan 3;68(1):491-516. [doi: 10.1146/annurev-psych-042716-051139] [Medline: 27687118]
64. Zhang D, Lee EKP, Mak ECW, Ho CY, Wong SYS. Mindfulness-based interventions: an overall review. Br Med Bull 2021 Jun 10;138(1):41-57. [doi: 10.1093/bmb/ldab005] [Medline: 33884400]
65. Segal ZV, Ferguson AM. Mindfulness-Based Cognitive Therapy: Treatment Development from a Common Cognitive Therapy Core Sci Pract Cogn Ther Found Mech Appl: The Guilford Press; 2018:159-174.
66. Seabrook E, Kelly R, Foley F, et al. Understanding how virtual reality can support mindfulness practice: mixed methods study. J Med Internet Res 2020 Mar 18;22(3):e16106. [doi: 10.2196/16106] [Medline: 32186519]
67. Failla C, Marino F, Bernardelli L, et al. Mediating mindfulness-based interventions with virtual reality in non-clinical populations: the state-of-the-art. Healthcare (Basel) 2022 Jun 29;10(7):1220. [doi: 10.3390/healthcare10071220] [Medline: 35885747]
68. Ma J, Zhao D, Xu N, Yang J. The effectiveness of immersive virtual reality (VR) based mindfulness training on improvement mental-health in adults: a narrative systematic review. EXPLORE (NY) 2023;19(3):310-318. [doi: 10.1016/j.explore.2022.08.001] [Medline: 36002363]
69. Navarro-Haro MV, López-Del-Hoyo Y, Campos D, et al. Meditation experts try virtual reality mindfulness: a pilot study evaluation of the feasibility and acceptability of virtual reality to facilitate mindfulness practice in people attending a mindfulness conference. In: Ito E, editor. PLOS ONE 2017;12(11):e0187777. [doi: 10.1371/journal.pone.0187777] [Medline: 29166665]
70. Puente-Torre P, Delgado-Benito V, Rodríguez-Cano S, García-Delgado M. Virtual reality as an interactive tool for the implementation of mindfulness in university settings: a systematic review. MTI 2024 Oct 11;8(10):89. [doi: 10.3390/mti8100089]
71. Arpaia P, D'Errico G, De Paolis LT, Moccaldi N, Nuccetelli F. A narrative review of mindfulness-based interventions using virtual reality. Mindfulness (N Y) 2022 Mar;13(3):556-571. [doi: 10.1007/s12671-021-01783-6]
72. Navas-Medrano S, Soler-Dominguez JL, Pons P. Mixed reality for a collective and adaptive mental health metaverse. Front Psychiatry 2023;14:1272783. [doi: 10.3389/fpsyt.2023.1272783] [Medline: 38250268]

XSL•FO
RenderX

73. Riva G, Baños RM, Botella C, Mantovani F, Gaggioli A. Transforming experience: the potential of augmented reality and virtual reality for enhancing personal and clinical change. Front Psychiatry 2016;7:164. [doi: 10.3389/fpsyt.2016.00164] [Medline: 27746747]
74. Cawley A, Tejeiro R. Brief virtual reality mindfulness is more effective than audio mindfulness and colouring in reducing stress in university students. Mindfulness (N Y) 2024 Feb;15(2):272-281. [doi: 10.1007/s12671-024-02306-9]
75. Modrego-Alarcón M, López-Del-Hoyo Y, García-Campayo J, et al. Efficacy of a mindfulness-based programme with and without virtual reality support to reduce stress in university students: a randomized controlled trial. Behav Res Ther 2021 Jul;142:103866. [doi: 10.1016/j.brat.2021.103866] [Medline: 33957506]
76. Pallavicini F, Argenton L, Toniazzi N, Aceti L, Mantovani F. Virtual reality applications for stress management training in the military. Aerosp Med Hum Perform 2016 Dec 1;87(12):1021-1030. [doi: 10.3357/AMHP.4596.2016] [Medline: 28323588]
77. Waller M, Mistry D, Jetly R, Frewen P. Meditating in virtual reality 3: 360° video of perceptual presence of instructor. Mindfulness (N Y) 2021;12(6):1424-1437. [doi: 10.1007/s12671-021-01612-w] [Medline: 33777253]
78. Cinalioglu K, Lavín P, Bein M, et al. Effects of virtual reality guided meditation in older adults: the protocol of a pilot randomized controlled trial. Front Psychol 2023;14:1083219. [doi: 10.3389/fpsyg.2023.1083219] [Medline: 37575420]
79. Poetar CR, Bradley N, Voinescu A. Immersive virtual reality or computerised mindfulness meditation for improving mood? Preliminary efficacy from a pilot randomised trial. Front Psychol 2023;14:1157469. [doi: 10.3389/fpsyg.2023.1157469] [Medline: 37954176]
80. Blackmore R, Giles C, Tremain H, et al. Examining the use of virtual reality to support mindfulness skills practice in mood and anxiety disorders: mixed methods study. J Med Internet Res 2024 Dec 6;26:e45640. [doi: 10.2196/45640] [Medline: 39641990]
81. Ma J, Li H, Liang C, Li S, Liu Z, Qu C. A brief virtual reality-based mindfulness intervention can improve olfactory perception while reducing depression and anxiety symptoms in university students. Humanit Soc Sci Commun 2025 Mar 3;12(1):294. [doi: 10.1057/s41599-025-04584-7]
82. Kelly RM, Seabrook EM, Foley F, Thomas N, Nedeljkovic M, Wadley G. Design considerations for supporting mindfulness in virtual reality. Front Virtual Real 2022 Jan 21;2:672556. [doi: 10.3389/frvir.2021.672556]
83. Foster JL, Shipstead Z, Harrison TL, Hicks KL, Redick TS, Engle RW. Shortened complex span tasks can reliably measure working memory capacity. Mem Cognit 2015 Feb;43(2):226-236. [doi: 10.3758/s13421-014-0461-7] [Medline: 25217113]
84. Baumeister RF, Heatherton TF. Self-regulation failure: an overview. Psychol Inq 1996 Jan;7(1):1-15. [doi: 10.1207/s15327965pli0701_1]
85. Pre-registrations | “Technology use and cognitive functions”. AsPredicted. URL: https://aspredicted.org/4qp8r.pdf [accessed 2026-05-01]
86. ResearchBox #3410 - “VR mindfulness”. ResearchBox. URL: https://researchbox.org/3410 [accessed 2026-05-01]
87. 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] [Medline: 29345484]
88. IBM SPSS statistics. IBM. URL: https://www.ibm.com/products/spss-statistics [accessed 2026-05-01]
89. JASP Team. 2023. URL: https://jasp-stats.org/ [accessed 2026-05-01]
90. Adler NE, Epel ES, Castellazzo G, Ickovics JR. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy, white women. Health Psychol 2000 Nov;19(6):586-592. [doi: 10.1037/0278-6133.19.6.586]
91. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
92. Hartanto A, Kasturiratna K, Chua YJ, et al. Does a 15-minute brief mindfulness breathing exercise temporarily enhance inhibitory control and cognitive flexibility? A within-subject experimental approach. In: Freitas A, editor. Collabra 2023 Nov 29;9(1). [doi: 10.1525/collabra.90219]
93. Hartanto A, Kaur M, Kasturiratna K, Quek FYX, Majeed NM. A critical examination of the effectiveness of gratitude intervention on well-being outcomes: a within-person experimental daily diary approach. J Posit Psychol 2023 Nov 2;18(6):942-957. [doi: 10.1080/17439760.2022.2154704]
94. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980 Dec;39(6):1161-1178. [doi: 10.1037/h0077714] [Medline: 7431205]
95. Lindner C, Lindner MA, Retelsdorf J. Die 5-Item-Skala zur Messung der momentan verfügbaren Selbstkontrollkapazität (SMS-5) im Lern- und Leistungskontext. Diagnostica 2019 Oct;65(4):228-242. [doi: 10.1026/0012-1924/a000230]
96. Psychology Software Tools, Inc. URL: https://support.pstnet.com/hc/en-us [accessed 2026-05-01]
97. Task downloads. EngleLab task downloads, EngleLab. URL: https://englelab.gatech.edu/taskdownloads [accessed 2026-05-01]
98. Clinton V, Swenseth M, Carlson SE. Do mindful breathing exercises benefit reading comprehension? A brief report. J Cogn Enhanc 2018 Sep;2(3):305-310. [doi: 10.1007/s41465-018-0067-2]
99. Hafenbrack AC, Kinias Z, Barsade SG. Debiasing the mind through meditation: mindfulness and the sunk-cost bias. Psychol Sci 2014 Feb;25(2):369-376. [doi: 10.1177/0956797613503853] [Medline: 24317419]

XSL•FO
RenderX

100. Quintana DS, Williams DR. Bayesian alternatives for common null-hypothesis significance tests in psychiatry: a non-technical guide using JASP. BMC Psychiatry 2018 Jun 7;18(1):178. [doi: 10.1186/s12888-018-1761-4] [Medline: 29879931]
101. Hofmann SG, Sawyer AT, Witt AA, Oh D. The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review. J Consult Clin Psychol 2010 Apr;78(2):169-183. [doi: 10.1037/a0018555] [Medline: 20350028]
102. Döllinger N, Wienrich C, Latoschik ME. Challenges and opportunities of immersive technologies for mindfulness meditation: a systematic review. Front Virtual Real 2021;2. [doi: 10.3389/frvir.2021.644683]
103. Deng Y, Zhang B, Zheng X, Liu Y, Wang X, Zhou C. The role of mindfulness and self-control in the relationship between mind-wandering and metacognition. Pers Individ Dif 2019 Apr;141:51-56. [doi: 10.1016/j.paid.2018.12.020]
104. Friese M, Messner C, Schaffner Y. Mindfulness meditation counteracts self-control depletion. Conscious Cogn 2012 Jun;21(2):1016-1022. [doi: 10.1016/j.concog.2012.01.008] [Medline: 22309814]
105. Jha AP, Krompinger J, Baime MJ. Mindfulness training modifies subsystems of attention. Cogn Affect Behav Neurosci 2007 Jun;7(2):109-119. [doi: 10.3758/cabn.7.2.109] [Medline: 17672382]
106. Mitsea E, Drigas A, Skianis C. Digitally assisted mindfulness in training self-regulation skills for sustainable mental health: a systematic review. Behav Sci (Basel) 2023 Dec 10;13(12):1008. [doi: 10.3390/bs13121008] [Medline: 38131865]
107. Morrison AB, Chein JM. Does working memory training work? The promise and challenges of enhancing cognition by training working memory. Psychon Bull Rev 2011 Feb;18(1):46-60. [doi: 10.3758/s13423-010-0034-0] [Medline: 21327348]
108. Shipstead Z, Redick TS, Engle RW. Is working memory training effective? Psychol Bull 2012 Jul;138(4):628-654. [doi: 10.1037/a0027473] [Medline: 22409508]
109. Jaeggi SM, Buschkuehl M, Jonides J, Perrig WJ. Improving fluid intelligence with training on working memory. Proc Natl Acad Sci U S A 2008 May 13;105(19):6829-6833. [doi: 10.1073/pnas.0801268105] [Medline: 18443283]
110. Zhang Y, Fu J, Zhao X. Neural correlates of working memory training: an fMRI meta-analysis. Neuroimage 2024 Nov;301:120885. [doi: 10.1016/j.neuroimage.2024.120885]
111. Hyndych A, El-Abassi R, Mader EC. The role of sleep and the effects of sleep loss on cognitive, affective, and behavioral processes. Cureus 2025 May;17(5):e84232. [doi: 10.7759/cureus.84232] [Medline: 40525051]
112. Majeed NM, Kasturiratna K, Li MY, Chia JL, Lua VYQ, Hartanto A. COVID-19 stress and cognitive failures in daily life: a multilevel examination of within- and between-persons patterns. Journal of Pacific Rim Psychology 2023 Jan;17:18344909231208119. [doi: 10.1177/18344909231208119]
113. Raver CC, Blair C. Neuroscientific insights: attention, working memory, and inhibitory control. Future Child 2016;26(2):95-118. [doi: 10.1353/foc.2016.0014]
114. Shields GS, Sazma MA, Yonelinas AP. The effects of acute stress on core executive functions: a meta-analysis and comparison with cortisol. Neurosci Biobehav Rev 2016 Sep;68:651-668. [doi: 10.1016/j.neubiorev.2016.06.038] [Medline: 27371161]
115. Chandrasiri A, Collett J, Fassbender E, De Foe A. A virtual reality approach to mindfulness skills training. Virtual Real 2020 Mar;24(1):143-149. [doi: 10.1007/s10055-019-00380-2]
116. Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. Cogn Psychol 2000 Aug;41(1):49-100. [doi: 10.1006/cogp.1999.0734] [Medline: 10945922]
117. Hartanto A, Chua YJ, Quek FYX, Wong J, Ooi WM, Majeed NM. Problematic smartphone usage, objective smartphone engagement, and executive functions: a latent variable analysis. Atten Percept Psychophys 2023 Nov;85(8):2610-2625. [doi: 10.3758/s13414-023-02707-3] [Medline: 37188861]
118. Blasiman RN, Was CA. Why is working memory performance unstable? A review of 21 factors. Eur J Psychol 2018 Mar;14(1):188-231. [doi: 10.5964/ejop.v14i1.1472] [Medline: 29899806]

Abbreviations

JARS: Journal Article Reporting Standards
MBSR: mindfulness-based stress reduction
MR: mixed reality
NA: negative affect
OSpan: operation span
PA: positive affect
RotSpan: rotation span
VR: virtual reality

XSL•FO
RenderX

Edited by I Steenstra, S Brini; submitted 16.Sep.2025; peer-reviewed by S Ye, YH Wang; revised version received 24.Mar.2026; accepted 25.Mar.2026; published 11.May.2026.

Please cite as:
Eun ZKY, Koh CJ, Yang H, Goh AYH, Hu M, Kasturiratna KTAS, Hartanto A
Brief Virtual Reality and Mixed Reality Mindfulness Breathing Exercise for Emotional Well-Being and Cognitive Functions in University Students: Within-Subjects Experimental Design Study
JMIR XR Spatial Comput 2026;3:e84239
URL: https://xr.jmir.org/2026/1/e84239
doi:10.2196/84239

© Zoey K Y Eun, Charmaine Jiali Koh, Hwajin Yang, Adalia Y H Goh, Meilan Hu, K T A Sandeeshwara Kasturiratna, Andree Hartanto. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 11.May.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Immersive Virtual Reality–Supported Cognitive-Behavioral Therapy for Patients With Mild to Borderline Intellectual Disabilities and Substance Use Disorders: Two Exploratory Studies

Samantha Murray[1,2*], MSc; Simon Langener[3,4,5*], PhD; Hanneke Kip[1,2], PhD; Randy Klaassen[3], PhD; Saskia Marion Kelders[1], PhD; Dirk Heylen[3], Prof Dr; Joanne VanDerNagel[3,4,5], MD, PhD

[1]Behavioural Management and Social sciences, Psychology Health and Technology, University of Twente, Drienerlolaan 5, Enschede, The Netherlands
[2]Department of Research, Transfore, Deventer, The Netherlands
[3]Human Media Interaction, University of Twente, Enschede, The Netherlands
[4]Centre for Addiction and Intellectual Disability, Tactus Addiction Care, Enschede, The Netherlands
[5]Nijmegen Institute for Scientist-Practitioners in Addiction, Radboud University, Nijmegen, The Netherlands
[*]these authors contributed equally

Corresponding Author:
Samantha Murray, MSc
Behavioural Management and Social sciences, Psychology Health and Technology, University of Twente, Drienerlolaan 5, Enschede, The Netherlands

Abstract

Background: Substance use disorders (SUDs) are prevalent and characterized by high relapse rates. Individuals with mild to borderline intellectual disability (MBID) are more likely to develop SUDs and face barriers within treatment related to difficulties they experience with abstract thinking, verbal skills, and generalizing learned strategies to real-world contexts. Therefore, experiential, context-rich approaches are needed that reduce reliance on retrospective verbal reflection, support in-context identification of triggers, and allow the rehearsal of coping responses. Immersive virtual reality (IVR) may provide realistic, safe environments where patients with SUD and MBID can practice cognitive and behavioral skills with visual and practice-oriented materials.

Objective: This study aimed to generate design input for the development and clinical integration of IVR-supported therapy for individuals with MBID and SUD. Specifically, Study 1 explored alcohol-related triggers in patients with alcohol use disorder, whereas Study 2 examined the feasibility and acceptability of practicing nicotine-related coping strategies in patients with nicotine dependence (ND).

Methods: Two explorative studies were conducted at an inpatient clinic for patients with MBID and SUD in the Netherlands. Study 1 included 10 adults with alcohol use disorder and MBID who participated in interviews to determine relevant risk situations, triggers, and therapeutic goals for IVR-cognitive behavioral therapy (CBT). Study 2 included 10 adults with MBID and nicotine dependence who practiced coping strategies within an existing IVR featuring craving-inducing and craving-reduction scenarios. A multiple-method approach was used to gather input for IVR-CBT development and to explore feasibility and acceptability (user evaluation interviews, the Questionnaire of Smoking Urges, and Visual Analog Scale ratings).

Results: In study 1, we identified high-risk situations, including at-home routines (eg, sitting on the couch watching football), supermarkets (eg, confrontation with alcohol and advertisements), social gatherings (eg, invitations and peer pressure), and being outside or traveling (eg, public transport or passing alcohol-related places). Triggers clustered into multisensory cues (eg, seeing or smelling alcohol), social influences (peer pressure and interpersonal conflict), affective states (tension, distress, boredom, or euphoria), and personal habits (eg, rewarding oneself or associations with money). Participants expressed interest in using IVR to identify triggers, discuss affective states, and train refusal skills. In study 2, IVR elicited nicotine craving, which increased during cue exposure and decreased during tutorial and coping phases. The coping elements embedded in IVR included relaxation (eg, mindfulness or breathing exercises), distraction (eg, virtual pets and interactive games), and physical activity (eg, walking or sports).

Conclusions: IVR-CBT elements appear feasible and acceptable in inpatient MBID care with appropriate support. Findings provide patient-derived design insights for integrating trigger identification and coping rehearsal within IVR. Future work should use an iterative, user-centered design approach based on validated CBT-related techniques (eg, functional analysis or coping, or skills training) and compare IVR-CBT with CBT as usual to understand benefits and risks for patients and therapists.

(JMIR XR Spatial Comput 2026;3:e82601) doi:10.2196/82601

XSL•FO
RenderX

KEYWORDS

immersive virtual reality; cognitive behavioral therapy; substance use disorder; alcohol use disorder; nicotine dependence; mild intellectual disability; borderline intellectual functioning; craving; relapse prevention; user-centered design

Introduction

Background

Substance use disorders (SUDs) are highly prevalent and complex psychiatric conditions with a high disease burden [1]. Individuals with mild to borderline intellectual disabilities (MBID) are at an increased risk for developing SUDs [2,3]. MBID covers both persons with mild intellectual disability (MID) and borderline intellectual functioning (BIF) [4], and in the Dutch clinical context, it is defined by an IQ between 50 and 85, accompanied by limitations in adaptive functioning with onset during the developmental period. Individuals with MBID often experience cognitive limitations that affect comprehension, memory, and problem-solving abilities, as well as difficulties in understanding the consequences of substance use, which can increase their susceptibility to SUDs and reduce the efficacy of treatment as usual [2]. Additionally, the prevalence of SUDs among persons with MBID is exacerbated by social vulnerabilities, such as problems with social skills [5], as well as social isolation, trauma, and substance use as a form of self-medication against negative life events [6].

Despite the availability of various psychotherapeutic and pharmacological treatments, maintaining long-term recovery in individuals with MBID can be challenging. As reported by the National Institute on Drug Abuse (2020), relapse rates in the general population who were treated for SUDs vary between 40 and 60 percent. Although specific relapse data for individuals with MBID are limited, relapse rates in this group might be even higher, given the challenges in assessment and standard treatment they may encounter [7,8]. For instance, most individuals with MBID experience communication barriers as they struggle to verbalize their thoughts, emotions, behaviors, and everyday risk situations. Moreover, barriers in skill application are present, as they often experience difficulties with applying newly learned skills from treatment in real-world risky situations [8]. These barriers in communication and skill application can make it more difficult to engage in therapeutic processes such as cognitive behavioral therapy (CBT) that rely on reflective, language-based approaches, which require cognitive and adaptive skills. Although tailored CBT protocols for the dual diagnosis of MBID and SUDs are available in the Netherlands [8,9], they still require patients to remember and reflect on their own behavior within treatment when discussing high-risk situations. Furthermore, generalizing and applying the newly learned coping skills in the real world remains difficult [8]. Consequently, there is a need for experiential, context-rich treatment approaches that reduce the reliance on abstract verbal processing and support in-context rehearsal of coping responses.

From Virtual Reality Exposure Therapy to Immersive Virtual Reality–Based Cognitive Behavioral Therapy

Immersive virtual reality (IVR) offers a promising approach to address the cognitive barriers that people with MBID experience within standard treatments that are based on the principles of CBT. IVR can be described as a computer-generated simulation of a 3D setting by using special electronic equipment [10]. Typically, head-mounted displays (HMDs) are used to immerse the user into a virtual environment (VE). Previous research suggests that IVR might lower learning obstacles by making experiences and training abstract concepts and interrelationships more explicit, understandable, and trainable, thereby improving comprehension for people who struggle with abstract thinking, such as those with MBID [11-13]. This approach eliminates the requirement for disembedded thinking, which refers to the capacity to grasp concepts outside of immediate and physical contexts. This is a skill that individuals with MBID find challenging [14]. Moreover, IVR can facilitate active learning rather than passively receiving information, thus reducing the reliance on abstract reasoning and verbal instructions while fostering skill acquisition [14]. This implies that learning becomes an engaging experience that does not solely rely on a language-based experience. Instead, IVR facilitates skill development through experiential learning, allowing individuals to learn by doing in a safe, controlled setting and acquire new skills without solely relying on language-based instructions [15,16].

Various studies have shown that VEs can elicit realistic (physical) reactions and behaviors in patients with SUD [10,17,18], including tension and substance cravings [13]. Virtual reality exposure therapies (VRET) have been developed based on these “natural” reactions to stimuli in IVR, for example, within the domain of anxiety disorders [19]. Following this initial success, research has also explored possible uses of VRET for SUDs, drawn by its ecological validity in realistically eliciting cue reactivity, such as drug craving and psychophysiological reactions (eg, increased heart rate and sweating) to stimuli associated with substance use [20].

However, the effectiveness of VRET for SUDs remains questionable. While pilot studies using VRET for SUDs led to cue reactivity and participant satisfaction, as well as expectations of clinical benefit, a 2021 review indicates that this has not yet been demonstrated in research on clinical effectiveness [21]. This is supported by an earlier review from Segawa et al [22], who stated that although IVR is effective at eliciting craving, VRET treatment outcomes vary. One hypothesis is that exposure in the context of addiction does not lead to extinction but rather reactivates a pathologically disturbed reward system, unlike anxiety disorders. As a result, VRET might even increase the risk of relapse [23], which is consistent with the lack of effectiveness of regular exposure therapies for SUDs [24]. Taken together, the evidence at the time of this writing indicates that VRET is insufficient for lasting change and that it may be essential to integrate, for example, CBT techniques in IVR interventions to actively understand triggers and train coping behaviors. This approach aligns with a literature review by Taubin et al [25], who found that IVR protocols integrating mindfulness practice, cognitive reappraisal, or other

XSL•FO
RenderX

emotion-regulation tasks alongside IVR were more likely to achieve clinically meaningful effects on substance-use outcomes. Experimental data support this conclusion; in a randomized study, smokers who actively crushed virtual cigarettes, an IVR task that requires users to enact an immediate coping response, achieved significantly higher abstinence rates than smokers who performed a neutral control task [26]. These findings indicate that active rehearsal of coping strategies rather than exposure alone is essential for therapeutic benefits. Accordingly, an IVR protocol used in an active learning format that engages patients in practicing coping responses to trigger-related craving may positively influence internal biases and promote behavior change.

Building on these insights, embedding active learning tasks into an IVR can lead to interactive, user-centered interventions that are more compatible with the needs of individuals with MBID and SUDs. IVR-CBT for SUDs allows patients to safely confront triggers by exploring maladaptive thought patterns and gradually exposing them to drug-craving-provoking situations while practicing coping strategies in a secure, controlled VE. Moreover, IVR-CBT provides a supportive learning environment that allows users to experiment with alternative behaviors in a step-by-step manner, thereby enabling cognitive restructuring and learning of suitable behavioral responses to SUD-related symptoms. This aligns with the core objectives for SUDs in standard addiction care treatments that are based on CBT principles [27], including psychoeducation, skill development for coping strategies, developing more adaptive thinking patterns, strengthening problem-solving skills in a practical manner, and relapse prevention through self-control and reinforcement of positive behaviors [28]. However, research in interdisciplinary teams is needed to understand design requirements to develop IVR-CBT interventions, as well as implementation into treatment as usual.

This study reports on 2 explorative studies conducted with patients in a specialized addiction clinic for individuals with MBID and SUD. Study 1 aimed to identify patient-reported high-risk situations and craving triggers relevant to alcohol use during preparation for temporary clinical leave. Study 2 aimed to explore the feasibility and acceptability of engaging patients with MBID in an IVR protocol that induces nicotine craving while embedding coping components in virtual reality (VR; eg, distraction, relaxation, and physical activity). Feasibility was defined as completion of IVR phases and tolerability (eg, need for support/breaks, discontinuation), and acceptability was defined as perceived realism and perceived usefulness. Given the substance-specific focus of each study, findings are interpreted as exploratory and design-oriented, without assuming transferability of identified triggers or coping preferences across substances. Across both studies, our overarching goal was to generate initial design requirements and clinical integration considerations for future IVR-CBT modules for MBID and SUD populations.

Methods

Study 1: Identifying High-Risk Situations and Triggers for IVR-CBT

Study Design and Participants

We explored high-risk situations and triggers to prepare patients with MBID and alcohol use disorder (AUD) for temporary clinical leave during treatment as usual. In total, 10 adults with MBID, receiving inpatient AUD therapy, were included using convenience sampling. Participants were recruited by their treating therapist within the inpatient clinic when they met the study criteria and were considered able to participate. Diagnostics for MBID were conducted by psychologists in the addiction clinic or by other medical institutions in the care chain of patients (eg, disability care). Exclusion criteria included severe psychiatric disorders (eg, psychosis) or active substance use. Eight of ten patients had previous experience with IVR, mainly via participation in the previous studies by our research group. Data were collected from January to April in 2021 in a Dutch addiction clinic specialized in the treatment of patients with MBID and SUD.

Ethical Considerations

This research was approved by the Saxion University of Applied Sciences in Enschede, as well as the scientific board of Tactus Addiction Care (OZP 26-112020). Personal identifiable information was archived separately from research data at Tactus Addiction Care to protect our patients from data breaches. Moreover, we pseudonymized data and deleted related audio recordings after transcription. Before starting the interviews, participants signed an informed consent. Patients received no compensation for participation.

Materials and Procedures

We conducted semistructured interviews with 7 open questions to explore (1) how patients would prefer to use IVR for practicing the clinical leave (2 questions), (2) which alcohol-related situations (1 question) induce alcohol craving, and (3) which triggers (1 question) induce alcohol craving. The interview guide (Multimedia Appendix 1) was pilot tested with 2 patients prior to data collection. The researcher explained the study procedure by using examples (eg, photos and videos) from the “Go up in smoke” project [13] to familiarize patients unfamiliar with IVR with relevant concepts. The “Go up in smoke” project focuses on the induction of nicotine craving using triggers and the reduction of craving using coping skills in IVR. Subsequently, semistructured interviews were held in a quiet office in the addiction clinic and were audio recorded.

Data Analysis

The audio files were transcribed verbatim and analyzed based on Braun and Clarke’s reflexive thematic analysis approach [29,30]. For this, the researcher (Berlind van Ast) followed the 6-step protocol: for (1) data familiarization, the researchers listened to the recordings and read the transcripts. Subsequently, (2) relevant segments in the data were identified, coded, and collated. Then, (3) initial themes were generated using the coded data and (4) reviewed by revisiting themes with respect to the

XSL•FO
RenderX

aims. Finally, (5) themes were defined, and the (6) report was produced. For this, we used the qualitative data analysis program Atlas.ti (v.9; ATLAS.ti Scientific Software Development GmbH).

During this process, researchers (Berlind van Ast and Jolien Jongeling) reflected together on themes generated during the 6-step analysis (steps 4 and 5), as well as during meetings with supervisors. Both researchers hold a degree in advanced nursing studies and were working in the addiction clinics at the time of data collection. Therefore, both researchers possess previous knowledge about the target group, as well as skills to engage in research with the MBID population. However, previous experiences with addiction care and the context of this research focusing on the IVR-CBT paradigm may influence the confirmability and transferability of findings.

Study 2: (Don't) Go Up in Smoke, Exploring Coping Strategies to Reduce Nicotine Craving in IVR

Study Design and Participants

In study 2, we implemented and evaluated coping skills in IVR that were derived from self-control techniques (eg, distance, distraction, and declare) that are taught during the less booze or drugs [9] treatment for patients with comorbid MBID. For this, we explored typical risk-coping scenarios in IVR to reduce nicotine craving in a convenience sample of 10 non–nicotine-deprived adults with MBID and nicotine dependence (ND) (Fagerström ≥5) undergoing inpatient treatment in a Dutch addiction clinic. Participants were recruited by their treating therapist when they met the study criteria and were considered able to participate. Diagnostics for MBID were conducted by psychologists in the addiction clinic or by other medical institutions in the care chain of patients (eg, disability care). Exclusion criteria included a history of migraine, epilepsy, motion sickness, severe psychiatric disorder (eg, psychosis), and use of nicotine replacement therapy to protect participants from IVR-induced symptoms (eg, cybersickness and hallucination) induced by the IVR device or, in the case of nicotine replacement therapies, to avoid introducing bias on the data collected. Only one of the 10 patients had prior experience with IVR. Data were collected from January to April 2021 in a Dutch addiction clinic specialized in the treatment of patients with MBID and SUD.

Ethical Considerations

Ethical approval as nonmedical research was given by the medical ethics board of the MST hospital in Enschede (K19-34). Moreover, this work was approved by the scientific board of Tactus Addiction Care. Personal identifiable information was archived separately from research data at Tactus Addiction Care to protect our vulnerable patients from data breaches. Moreover, we pseudonymized data and deleted related audio recordings after transcription.

All participants were welcomed by the researcher and thoroughly informed about the research procedure. After informed consent was obtained, the audio and screen recordings were started. The participants received no compensation for participation.

Measures

ND was assessed with the Dutch version of the Fagerström Test for nicotine dependence (FTND) [31]. The FTND is a 6-item questionnaire assessing ND severity with scores ranging from 0 to 10. Scores of 0 - 2 indicate a light dependence, 3 - 5 a moderate dependence, 6 - 7 a severe dependence, and 8 - 10 a very severe dependence.

Nicotine craving was assessed with the Questionnaire of Smoking Urges Brief (QSU-Brief) and by using a Visual Analog Scale (VAS). The QSU-Brief is a 10-item questionnaire assessing the urge to smoke [32]. Scores on each item range from 1 ("Strongly disagree") to 7 ("Strongly agree"). The total score is obtained by calculating the mean of the 10 items. The VAS is a single-item scale ranging from 0 to 10, where 0 is interpreted as "no craving" and 10 as "severe craving."

A semistructured interview (Multimedia Appendix 2) with 14 open questions was conducted to explore (1) VE to induce and reduce nicotine craving (3 questions), (2) explored coping strategies to reduce nicotine craving (5 questions) and alternatives in IVR to reduce nicotine craving (5 questions), and (3) potential applications of IVR during treatment (1 question).

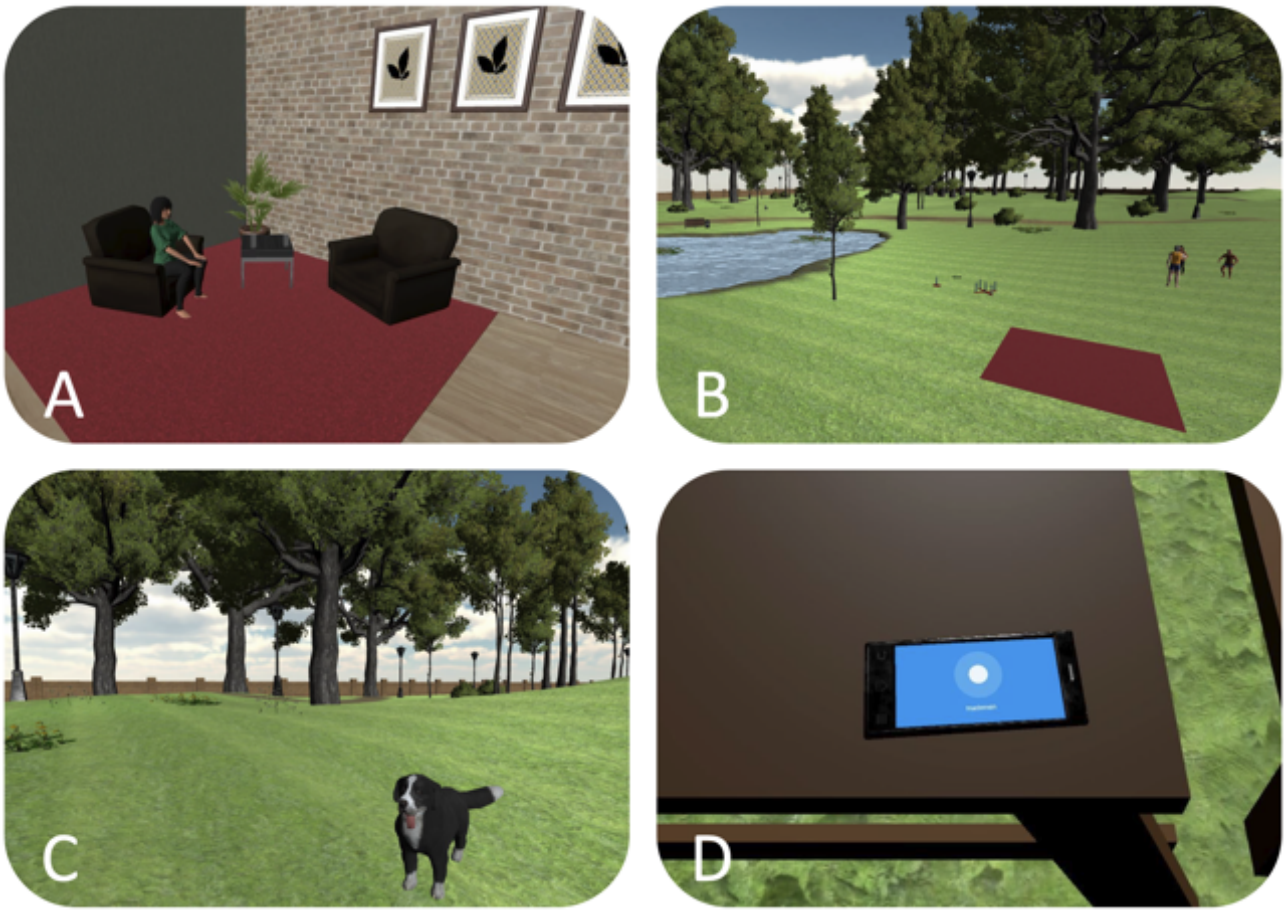
Hardware and VE

We used an HTC VIVE Pro Eye HMD (HTC Corp), 1440×1600 pixels per eye (2880×1600 combined), a 90 Hz refresh rate, and an 110-degree field of view, base stations, controllers, and a compatible laptop to display the IVR to the patient.

The VE was developed in Unity3D (v.2019.2.3f1; Unity Technologies) using the SteamVR Software Development Kit (Valve Corporation) and was partially adapted from previous work by our department [13], containing two main scenarios: (1) craving induction and (2) reduction (see Figures 1A-1D). The craving induction scenario comprised 3 main areas [13], including a crossroad with a bus stop, at home with a garden, and a restaurant with a terrace and smoking goods vendor. In contrast, the craving reduction scenario comprised 2 environments, including a therapy room with 2 virtual humans guiding 2 different mindfulness exercises that were prerecorded with a psychomotor therapist, as well as a virtual cat in the front yard, a park with a lake, a walkway, sporting agents, a blanket with calming music, a ring toss game, a badminton set, a dog that can be taken for a walk, and benches with a virtual smartphone to conduct a visual breathing exercise. Audio cues respecting the restaurant, interactive humans, cigarettes, blowing wind, running water, and calming music were used to deepen immersion. Participants were able to use teleport locomotion and an approximately 2×2 m room-scale area for natural locomotion.

XSL•FO
RenderX

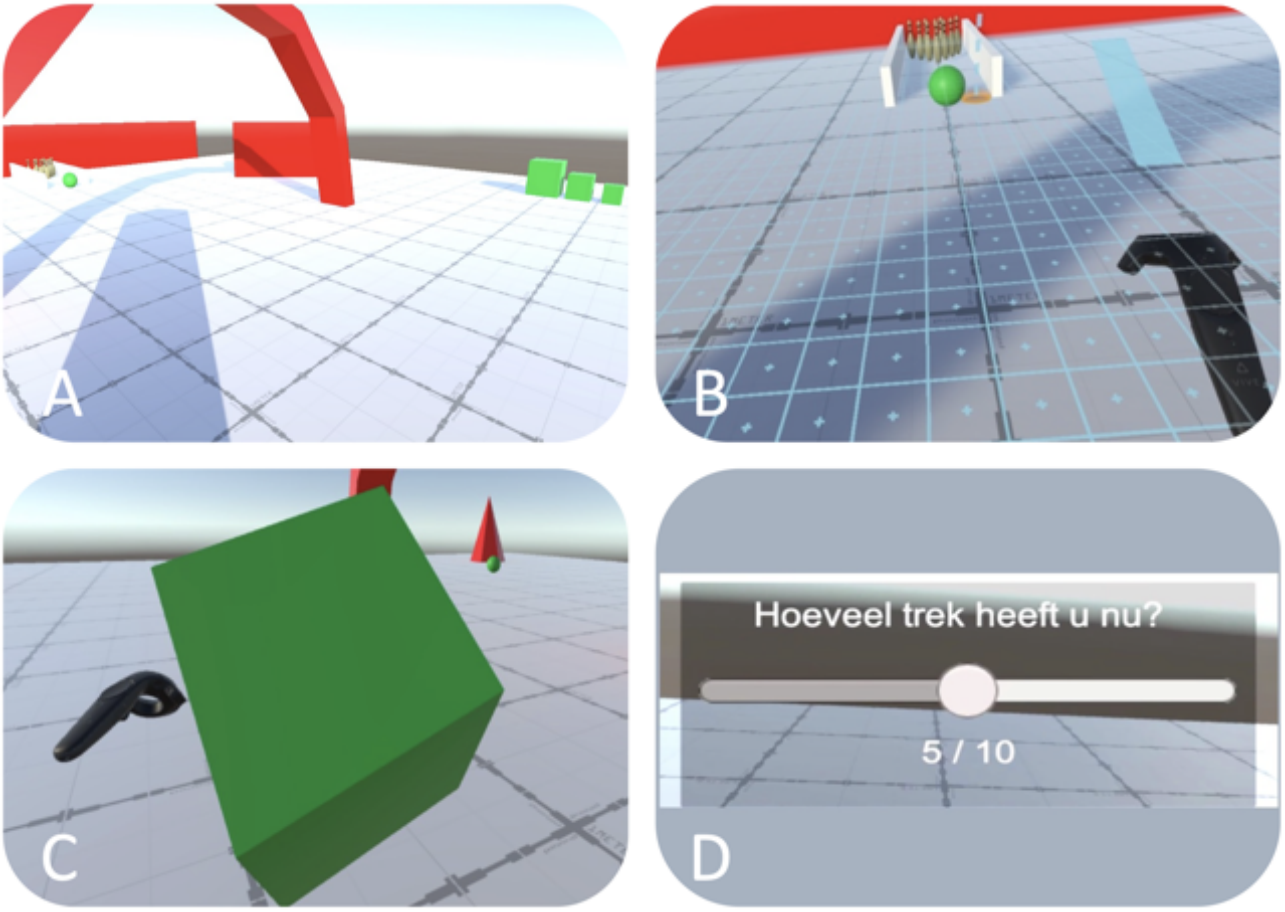
Figure 1. The virtual environment (VE) for craving reduction. (A) Mindfulness via a virtual therapist, (B) Park with sporting virtual humans, relaxation areas, and games, (C) Walk in the park with a dog, and (D) Breathing exercises on a virtual smartphone.



To train patients in using the controllers and ensure habitation within the IVR, a neutral tutorial was used with specific interactions (ie, grabbing objects or teleporting) by using a plane with a walking track, 3 interactable cubes, and a bowling alley (see Figure 2A-2D).

XSL•FO
RenderX

Figure 2. The tutorial environment to train interactions in immersive virtual reality (IVR), including a (A) "Walking track" to train teleporting, (B) "Bowling alley" to grab and throw a bowling ball, (C) three interactable cubes of different sizes to be stacked, and (D) Visual Analog Scale (VAS) for craving assessments.



Procedure

The research was carried out in a quiet room in an addiction clinic to enhance an ecologically valid surrounding that replicated the usual treatment setting. The procedure (maximum 90 min) was divided into five phases: (1) preparation (≈15 min), (2) tutorial (≈10 min), (3) trigger-induced craving (≈15 min), (4) coping skills and craving reduction exploration (≈15 min), as well as (5) postquestionnaires (≈30 min). During the preparation, demographics and baseline characteristics (ie, FTND and QSU-Brief) were assessed. If necessary, patients were allowed small breaks during phases 1 - 5 without leaving the room (eg, to smoke). At the beginning of the sessions, participants engaged in a short tutorial session in a neutral VE to become familiar with the HMD and controls and rate their craving using the VAS. Upon completion, participants removed the HMD and were verbally asked the QSU-Brief again before entering the cue reactivity environment. Participants were asked to wear the HMD again and conduct a guided exploration (see [13] of 5 specific situations aiming to induce cravings, including bus stop, restaurant terrace, a vendor with smoking goods, an at-home environment with a garden, and a man sitting on a bench smoking). After completing the craving induction stage, patients were instructed to interact with various coping skills, including distraction (ie, breathing and mindfulness exercises, sport, relaxation, and animals), distance (ie, walk in park or go outside), different thinking, different acting (ie, say "no" when a cigarette is offered, drink a soft drink, or use a nicotine patch), declaring (ie, use a virtual phone to call for help), which were derived from existing therapy protocols for people with dual diagnosis [9]. During both phases, patients were asked to rate their cravings after each interaction using the in-game VAS, followed by removing the HMD and completing the QSU-Brief. Finally, the semistructured interview was conducted verbally with participants. During the use of the HMD, patients' therapists were close by to provide help if needed.

Data Analysis

The audio files were transcribed verbatim and analyzed based on Braun and Clarke's reflexive thematic analysis approach [29,30]. For (1) data familiarization, the researcher (Jolien Jongeling) listened to the recording and read the transcript. Then, (2) relevant segments were identified, coded, and collated. The (3) initial themes were generated using the coded data and (4) reviewed by revisiting themes with respect to the dataset and goal. Finally, (5) the themes were defined, and the (6) report was produced. For this, we used the qualitative data analysis program Atlas.ti (v. 9). Data from the QSU-Brief and VAS were analyzed descriptively using RStudio (v. 1.3.1093; PBC).

During this process, researchers (ie, Jolien Jongeling and Berlind van Ast) reflected together on themes generated during the 6-step analysis (steps 4 and 5), as well as during meetings with supervisors (ie, SL and JV). Both researchers pursued a master's degree in advanced nursing studies and worked in the addiction

XSL•FO
RenderX

clinic described in both studies. Therefore, both researchers possess previous knowledge about our target group, as well as skills to engage in research with the MBID population. However, previous professional experiences and the context of this research with a focus on the IVR-CBT paradigm may influence the confirmability and transferability of findings to other settings.

Results

Study 1. Identifying High-Risk Situations and Triggers for IVR-CBT

Sample Description

Participants (n=10) had a mean age of 40.5 (9.7) years, and most identified as female (n=6). The sample included patients with BIF (n=7) and MID (n=3). Notably, most participants (n=7) had prior IVR experience from the "Go up in smoke" project.

High-Risk Situations and Triggers That Elicit Alcohol Craving

The participants identified high-risk situations (see Table 1), including at home, supermarkets, social gatherings, and being outside.

In the home environment, participants mostly described the living room to elicit alcohol habits, though these situations are usually individual, as others describe the bathroom or kitchen. Also, the conditions vary from clean to messy (eg, with bottles all around). Regarding the supermarket, participants described associations and confrontation with alcohol cues (eg, advertisement), potentially triggering affective states ("Then my craving goes up and I start sweating and then I just get difficult thoughts. That I have to drink. That everything gets better when you drink. Yes, very nervous, stressful." [P02]). Social gatherings include meetings with families and peers that entail invitations to drink, including peer pressure and conflict situations ("It is when I am with my father, that he triggers me with his negative energy, my living room, and the supermarket." [P1]). Finally, being outside passing beer gardens, supermarkets, or being in public transport (ie, bus and train) is probably causing distress or invitations from others when walking by.

Table . High-risk situations and triggers in patients with mild to borderline intellectual disability (MBID) and alcohol use disorder (AUD).

Theme	Example
High-risk situations	
At home	"As soon as I enter, I sit on the couch, grab a glass and turn on the football. Then I just sit and consume it on my own." (P08)
Supermarket	"You enter the store and then you see, and you are confronted and then you feel that urge for alcohol. And even if I were to walk by, I would still see them, and they would look at me too. They look like: 'don't forget me.' " (P08)
Social gatherings	"Often you will meet someone you know, and they often know that you like a glass of wine and a beer. And then it's: 'hey, come and sit down, nice and cosy.'" (P07)
Being outside	"Put [Dutch city] in there. There they have to do it themselves. Either by bus or by train. You could combine that with those virtual reality glasses, with the person who either has to go by bus or by train and see how the person reacts." (P01)
Triggers	
Multisensory experiences	"I also deliberately don't go through that hallway anymore, but you see the offer when you come in these days and it's actually terrible, like they do it on the TV." (P05)
Social influences	"For example, if someone would like to belong to a group, then that group encourages: 'Hey, join us!' You should put something like that in it. I would experience that as a trigger." (P07)
Affective states	"When I'm high in tension, I tend to go to the store very quickly. Then I'll go get some alcohol to push that away again. To feel relaxed again." (P01)
Personal habits	"I associate money with alcohol because then these are my groceries." (P08)

Triggers (see Table 1) to use alcohol in the abovementioned high-risk settings can be classified into multisensory experiences, social influences, affective states, and personal habits. The participants described seeing, smelling, and hearing alcohol-related cues in daily situations, such as people drinking, (alcohol) advertisements, bottles, supermarkets, soccer on TV (with sound), and (Dutch) music during social gatherings. Social influences comprise the described invitations, pressure, and conflict situations with (relevant) others to elicit strong distress/cravings. Generally, participants described a spectrum

XSL•FO
RenderX

of affective states, such as distress, boredom, and euphoria as causes of alcohol craving. Finally, personal habits, such as rewarding oneself, nice weather, watching TV, having money in one's wallet, or completing work, were also described, often intertwined with the aforementioned triggers (ie, multisensory cues, social influences, and affective states).

Applications of IVR for CBT in Clinical Care

The preferred use of IVR during clinical care in patients with MBID and AUD can be divided into assessment and treatment (see Table 2).

Table . Preferred use of immersive virtual reality (IVR) for cognitive behavioral therapy (CBT) in patients with mild to borderline intellectual disability (MBID) and alcohol use disorder (AUD).

Theme	Example
Assessment	
Identify high-risk situations and triggers	"I will feel safe. Yes, because suppose you really get those craving moments or a panic attack, then you have a backup to rely on." (P01)
Coping skills training	
Discuss feelings during high-risk situations with therapists	"Because you are still in a safe environment, you can better discuss what you are feeling at that moment and how you can best deal with it." (P07)
Training of distraction techniques for coping with alcohol craving	"Practicing with such glasses to look for distraction, that is very important. (.) For example, that I get a trigger to take a walk. So, I didn't do that before, I just remained seated and then went for a walk to the fridge." (P06)
Learn to say "no" when offered alcohol	"Seeing other people drinking seems like a very good thing to practice [with IVR], you will soon encounter that at birthdays, and I just want to be able to do it without it. And if they offer me one then I want to be stronger to say 'no.'" (P05)

For assessment, participants reported the need to identify personal risk situations and triggers using IVR. This includes measuring craving in the given context (by using a VAS) to gain insights into one's own vulnerabilities. Regarding treatment, participants reported the wish to discuss feelings during risk situations with therapists, to train distraction techniques for coping with cue reactivity, and learn to say "no" when offered alcohol, for example, by choosing alternatives. In doing so, participants said they would like to train and rehearse repeatedly before leaving the clinic ("If you have always said 'yes' [to substance use] it is very strange for you to say 'no.' And I think the more you practice that, the easier it gets." [P07]). Most participants indicated a positive intention to use IVR for practicing clinical leave ("Yes, that would be nice, because then you are not immediately placed in the real situation, and you can learn to deal with your feelings and with your craving." [P02]), though some found it redundant or anxiety-evoking ("It was too scary for me. Everything overwhelms you, so to speak." [P10]).

Study 2: (Don't) Go Up in Smoke–Exploring Coping Strategies to Reduce Nicotine Craving in IVR

Sample Description

Participants (n=10) had a mean age of 39 (9.0) years, of which half (n=5) identified as female. The sample included patients with BIF (n=7) and MID (n=3). The FTND was moderate (n=2), severe (n=4), and very severe (n=4), respectively. Patients smoked on average 18.7 (4.2) per day, and most (n=9) had no prior IVR experience.

IVR Craving Induction and Procedural Remarks

All participants reported the IVR to be craving-inducing. Especially, the man offering a cigarette and the bus stop were mentioned to cause severe cravings ("Yes and that man on that bench, who then offers you a cigarette and at that moment you just notice something goes through you, oh yes delicious." [S08]). Though some participants reported prior to using the IVR worries about their abilities to navigate and interact, the researchers observed mostly positive remarks after immersion in IVR. However, 2 patients described anxiety issues related to using IVR ("I do find it a little scary" [S01]), resulting in a single dropout.

Coping Strategies to Reduce Nicotine Craving in IVR

For the implemented coping skills (see Table 3), participants named distraction through games and animals; relaxation (ie, mindfulness and breathing exercises); and physical activity (ie, sport going and for a walk in the park [with a virtual dog]) as coping strategies to reduce nicotine craving in IVR. Moreover, all participants engaged in mindfulness/breathing exercises delivered by virtual humans or smartphones. The strategies were described as appealing, though some described control and concentration troubles ("Then I have to be very quiet. No people around me, because then it won't work." [S10]). Furthermore, engaging in physical activity was pleasant, for instance, through sports and games (eg, bowling, throwing rings, and stacking cubes), as most participants mentioned the neutral tutorial to reduce craving ("I immediately stopped thinking about smoking" [S06]). For animals, people enjoyed petting and walking a virtual dog and playing with a virtual cat, as some missed their pets during their clinical stay. Noteworthy, being busy with IVR was described as distracting by many participants. Coping strategies to add to IVR include yoga, additional sports (eg, football and basketball), meditation, talking to a therapist or other patients, and creative tasks.

XSL•FO
RenderX

Table . Coping strategies to reduce nicotine craving in immersive virtual reality (IVR).

Theme	Example
Distraction	
Interactive games (eg, bowling, ring toss, and stacking cubes)	"But that's because I focused on the actions I had to do, that I focused very much on that actually, so I might actually suppress the urge […]." (S05)
Virtual pets	"But also with my cat, which walks with me to the store, that sort of thing, yes crazy. Yes, I miss that animal so much and then they say it's just a cat. Yeah, it's easy for you to say. I see him as my child." (S10).
Sitting close to water (on blanket hearing music)	"That you were distracted, I think, especially that last one beside the water and then with that dog." (S02)
Talking to therapist or other patients[a]	"[…] that someone tells you that the craving will go away and so on, that you will think about it, that you maybe can let it go a bit better." (S08)
Creative tasks	"The open air, the bench to sit down quietly and eh meditate a bit or if necessary, bring a coloring book, because I still like to color very much." (S05)
Relaxation	
Mindfulness	"It did diminish a little with the relaxation exercises, when I sat in that chair like that for a while." (S08)
Breathing exercises	"[…] of breathing in and out, that's actually a kind of relaxation exercise." (S05)
Physical activity	
Going for a walk in the park (with a virtual dog)	"It was more feeling good again, that dog came running to me and eh you miss that very much here" (S08)
Sports (eg, squats, yoga, basketball, football)	"[…] because it is anyway necessary for my fitness, but also for distraction and yes sport is just better on all fronts." (S04)

[a]Not implemented in our immersive virtual reality coping strategies exploration for craving reduction.

Subjective Craving During Baseline, Tutorial, Craving Induction, and Craving Reduction

Table 4 describes data obtained from the QSU-Brief and VAS across the different measures at baseline (T0), tutorial (T1), craving induction (T2), and craving reduction (T3). The results show tendencies for craving reduction during the tutorial (T1) and coping skills exploration (T3).

Table . Difference in craving on Questionnaire on Smoking Urges Brief (QSU-Brief) and Visual Analogue Scale (VAS) at T0-T3.

Measures	T0	T1	T2	T3
QSU-Brief[a]				
Mean (SD)	3.41 (1.1)	3.04 (1.4)	4.20 (1.5)	3.66 (1.2)
Median (IQR)	3.50 (3.0-3.8)	3.15 (1.6-4.3)	4.35 (3.5-5.2)	3.55 (3.0-4.3)
VAS[b] in IVR[c]				
Mean (SD)	4.22 (2.2)	2.86 (2.8)	5.34 (1.8)	2.50 (2.1)
Median (IQR)	4.00 (4.0-5.0)	2.00 (1.0-4.0)	5.57 (4.6-6.6)	2.00 (1.5-2.8)

[a]QSU-Brief: Questionnaire on Smoking Urges Brief.

[b]VAS: Visual Analogue Scale.

[c]IVR: immersive virtual reality.

Discussion

Principal Findings

This study reported 2 exploratory, design-oriented studies that were conducted in an inpatient MBID setting, examining high-risk situations, triggers, and coping strategies to inform the development and clinical integration of IVR-supported elements for substance-use treatment.

In the first study, qualitative interviews were used to identify alcohol-related high-risk situations, triggers, and patient-defined treatment goals in preparation for temporary clinical leave. Participants identified a range of high-risk situations for AUD, including being at home, in supermarkets, at social gatherings,

XSL•FO
RenderX

in public spaces, or while traveling. Alcohol craving in these contexts was associated with multiple types of triggers, including multisensory cues, social influences, affective states, and personal habits. Participants described these alcohol-related high-risk situations as challenging and expressed interest in using IVR to identify personal triggers in context, discuss emotional responses while being immersed, and repeatedly rehearse refusal skills or alternative actions for temporary clinical leave.

In the second study, an existing IVR cue reactivity environment was used to explore the feasibility and acceptability of practicing coping strategies for nicotine craving, alongside descriptive craving assessments. All participants reported that the IVR elicited nicotine craving, and most were able to complete the procedure. This suggests that the use of IVR was feasible in an inpatient MBID setting, when therapist support was available. Participants perceived several coping strategies as helpful, which were related to distraction (eg, games and virtual pets), relaxation (eg, mindfulness and breathing exercises), and physical activity (eg, sport and walking the dog). Descriptively, nicotine craving tended to increase during cue exposure and decrease during the tutorial and coping phases. Notably, participants differed in how they experienced the specific IVR scenarios, and some required additional support due to anxiety or unfamiliarity with IVR. For example, one participant described a virtual riverside environment as triggering for nicotine craving due to personal associations with alcohol use ("I live close to the IJssel [Dutch river] myself and that was a very triggering place for me last summer, because automatically people sit there to drink alcohol" [S04]), whereas other participants experienced similar scenarios as craving reducing. This illustrates that the subjective meaning of IVR scenarios can vary substantially between individuals and that certain environments may evoke unintended craving responses depending on personal history.

Taken together, these findings suggest that IVR may provide a structured and experiential format for identifying high-risk situations and rehearsing coping responses in context for individuals with MBID. This is particularly relevant given the barriers in MBID treatment, including difficulties with abstract reflection and transfer of coping skills to everyday environments [2,8].

Interpretation and Comparison With Prior Work

Participants in Study 1 identified different high-risk alcohol situations such as supermarkets, home settings, social gatherings, and public spaces. These findings are consistent with previous research on triggers in AUD, which has highlighted the role of environmental cues such as bars and stores in eliciting cravings [21,33]. However, participants in the present study frequently described these environments in combination with social influences (eg, peer pressure and interpersonal conflict) and affective states (eg, tension, negative affect, and distress) as triggers for alcohol craving. For example, in supermarkets, alcohol exposure was described as stressful and confrontational due to advertisements and visibility of alcohol products. Social gatherings were linked to invitations and pressure to drink. In the home setting, alcohol use seemed to be described as embedded in a habitual sequence of actions, such as coming home, sitting on the couch, watching football, and then drinking.

These findings indicate that alcohol craving in this MBID sample was experienced as embedded in daily routines and interpersonal situations, rather than triggered by exposure to environmental stimuli alone. This aligns with broader conceptualizations of craving as context-dependent while being influenced by affective and social processes [33]. In this study, participants described complex situations in which contextual cues, emotions, and social interactions were closely intertwined. For individuals with MBID, who often find it difficult to analyze and describe their everyday risk situations [2,8], addressing such interconnected triggers in a traditional CBT session can be challenging. CBT typically relies on retrospective discussion and functional analysis of high-risk situations. This requires patients to reconstruct the situation, identify the links between context, emotion, and behavior and consider alternative responses. Within this context, IVR may offer a complementary way to examine and rehearse high-risk situations in a more concrete and structured way.

While most IVR cue reactivity research operationalizes craving through predefined substance-use environments, such as bars or public venues [10,17,18], the present study identified high-risk situations and triggers for alcohol craving in patients with MBID through a bottom-up approach derived directly via interviews with patients. The situations described by participants were often routine and home-based and embedded in everyday activities rather than limited to prototypical drinking locations. Using these patient-derived situations as a starting point when developing IVR scenarios may increase the practical relevance of IVR within MBID inpatient treatment.

In the second study, participants explored several coping strategies to reduce nicotine craving in IVR. All patients reported that the IVR environment induced nicotine craving, confirming its ecological validity, consistent with earlier findings [13]. Interestingly, nicotine craving decreased during the tutorial phase, which contrasts with previous results that showed increased nicotine craving during immersion [13]. One possible explanation for this difference is that the gamified tutorial required sustained attention and functioned as a distraction, thereby reducing nicotine craving. This observation suggests that the structure and content of the tutorial phase may influence craving responses and should therefore be considered carefully when designing IVR protocols.

Participants perceived coping strategies that were embedded within IVR as helpful, particularly those that involved concrete actions such as interacting with virtual animals or engaging in games. This suggests that including active coping components within IVR may be important, rather than relying on cue exposure alone [22,26]. However, participants differed in how they experienced certain environments, which highlights the need for careful scenario selection and therapist guidance. The riverside example illustrates that environments intended to be calming may evoke craving if they are associated with prior substance use experiences. This underscores the importance of assessing personal associations and monitoring tolerability when integrating IVR into clinical care.

XSL•FO
RenderX

Implications for the Development of an IVR-CBT for Patients With MBID

The findings from both studies provide practical starting points for the development of IVR-supported CBT elements within MBID SUD treatment. The present studies identified high-risk situations and coping strategies that patients themselves considered as relevant within inpatient MBID treatment and demonstrated that practicing coping skills within IVR was feasible and acceptable when therapist support was available. Participants also described treatment-related goals, such as discussing feelings with therapists during IVR sessions, practicing refusal skills, and learning to manage craving-related distress. These goals align with CBT mechanisms such as functional analysis and coping skills training [34], as well as with MBID-adapted CBT protocols that are used within Dutch inpatient treatment settings at the time of this writing [9].

Although clinical effectiveness was not evaluated, the descriptive reductions in craving during the coping phases suggest that CBT-based coping strategies, including self-control techniques tailored to the needs of patients with MBID, can be operationalized within IVR environments. The possibility to simulate socially and emotionally complex situations in a structured and controlled setting may allow for repeated rehearsal, which can be difficult to stimulate in traditional therapy but can be practiced safely and repeatedly in IVR. Prior studies have shown that IVR is effective at establishing cue reactivity by simulating real-life scenarios related to substance use. However, many applications focused primarily on exposure to SUD-related cues, without embedding active coping skill training or other therapeutic elements that are essential for behavioral change [35]. Integrating coping rehearsal directly within IVR may help to address this limitation by combining immersive exposure with the immediate practice of coping strategies.

Based on the findings from both studies, several design and implementation considerations for developing IVR-CBT interventions (ie, trigger-coping scenarios) can be identified. First, IVR environments may be useful for supporting two fundamental CBT objectives: (1) identification of high-risk situations and triggers (CBT: stimulus control) and (2) coping with trigger-induced craving (CBT: stimulus response prevention) [36]. IVR may function as a complementary modality within existing CBT protocols by enabling in-context rehearsal of skills that are otherwise discussed retrospectively. In doing so, a structured IVR-CBT framework may be developed that systematically links CBT objectives to certain IVR training modules with risk-coping scenarios, grounded in both CBT theory and immersive learning design. Stimulus control could be addressed through specific IVR trigger-assessment modules, in which patients, for instance, identify, label, and rate high-risk situations or triggers via simplified interfaces and real-time feedback tools via either the IVR or the therapist. Once patients can recognize and assess their high-risk situations and triggers, the next step would involve training them to actively manage their responses, thereby moving from stimulus control to stimulus-response prevention. The transition from stimulus control to stimulus-response prevention can be supported by an IVR-CBT framework with a staged approach that aligns technological affordances with a pedagogical learning design, as recommended by [37], by sequencing skill acquisition through concrete, repetitive, and context-specific activities.

Moreover, given the variability in individual trigger experiences observed in this study, IVR systems may benefit from allowing flexibility in scenario selection and coping options. Personalization through therapist-patient goal setting [37] and the use of customized scenarios that are based on each patient's individual triggers and coping needs [38] might enhance clinical relevance. For instance, patients presenting with predominantly social triggers can, for instance, engage in peer-pressure simulations involving assertive refusal training, while patients with predominantly environmental triggers (eg, being in a bar) may work within location-based scenarios that focus more on stimulus control and rehearsal of coping strategies (eg, going for a walk). However, it is still uncertain whether individualized tailoring improves outcomes when compared to standardized IVR modules, as this still requires empirical evaluation.

Furthermore, IVR content may benefit from alignment with individualized high-risk situations and trigger profiles systematically linked to personalized coping strategies, while keeping a balance between personalized IVRs and standardized treatments [39]. It is important that researchers distinguish therapeutic mechanisms from educational ones, as skill learning in IVR may occur even in the absence of immediate relief effects, which are typically associated with real-world coping. Therefore, future studies should not only evaluate whether patients can learn coping strategies in IVR but also how these strategies generalize to behavior outside of IVR. Finally, the development of IVR-CBT protocols may benefit from interdisciplinary collaboration to address clinical, technological, and implementation considerations, including questions related to intensity and dosage. Future research could contribute to the development and empirical evaluation of a CBT-informed framework for IVR interventions that links identified triggers to coping strategies across structured learning stages. In addition, further studies should examine personalization capabilities, implementation barriers, and acceptability of IVR-CBT across broader MBID populations and diverse treatment settings.

Limitations

Several limitations should be considered. First, both studies included small convenience samples recruited from a single clinic; therefore, limiting the extent to which the results can be generalized to other settings or MBID populations. Our convenience sampling might have over-represented individuals with MBID who were comparatively stable, motivated, and comfortable with technology since participation in the studies was voluntary and patients with severe psychiatric instability were excluded, resulting in selection bias. As a result, our findings may not fully apply to the perspectives and needs of individuals with higher levels of psychiatric complexity or lower technological literacy. Second, the 2 studies focused on different substances and different therapeutic targets. While AUD and ND may share some features, they may also differ in contextual triggers, behavioral patterns, and treatment mechanisms; therefore, the findings should be interpreted within each specific substance context. While overlapping processes such as cue

XSL•FO
RenderX

reactivity and coping rehearsal may be relevant across SUDs, the specific triggers and coping elements that were identified in this work are substance-specific and design-oriented. In addition, the studies were exploratory and design-oriented, thereby not assessing clinical effectiveness, long-term outcomes, or the generalization of learned coping strategies beyond the IVR context. Third, the data were collected verbally by nurses (in training to become specialists) to resemble a common treatment setting. However, this might have increased social desirability bias (eg, see contrary craving effects during the tutorial phase), though paper-based assessments appear too complex for our group. Fourth, the interactions within the IVR were limited and of short duration due to the explorative nature of our study. This might have restricted the patient's opportunity to fully practice coping responses extensively and may have limited the extent to which craving reduction could be observed. Fifth, it is important to note that we demonstrated and explored coping skills by using a pre-existing IVR system. This may have caused confirmation bias because participants' given responses may have been influenced by the available specific examples and functionalities that were presented, which may have limited the identification of additional coping strategies. Instead, our explorative research aimed to generate or understand how IVR can be integrated into current CBT treatment in patients with MBID and SUD. Finally, participants might have participated in prior studies by our group, which could influence the data collected.

Conclusions

IVR-supported CBT elements appear feasible and acceptable in an inpatient MBID setting when appropriate support is available. The present studies provide concrete, patient-derived design insights into alcohol-related high-risk situations and nicotine-related coping elements that can be embedded in IVR. The findings of these studies suggest that IVR may help to reduce barriers that are associated with abstract verbal reflection by enabling in-context identification of triggers and repeated rehearsal of coping responses within a controlled environment. However, IVR environments may activate personally learned substance-use associations that differ between individuals. Therefore, careful scenario selection, real-time craving monitoring, and therapist guidance seem to be necessary to prevent unintended increases in craving.

Acknowledgments

We would like to thank our participants from Tactus addiction care. Moreover, we want to thank Berlind van Ast (BvA) and Jolien Jongeling (JJ) for their work in collecting data and qualitative analyses as part of their master thesis projects. During revision of the manuscript, the authors used ChatGPT (free version, OpenAI) in a limited way to refine the wording of individual sentences and to improve grammar, clarity, and readability. The tool was not used to generate research ideas or hypotheses, plan the study, search for or evaluate the literature, analyze data, interpret findings, generate references, write conclusions, or produce the scientific content of the manuscript. Its role was limited to language assistance for text that was already written by the authors, and suggested edits were reviewed and retained only when considered accurate. The authors remain fully responsible for the content, accuracy, and conclusions of the manuscript.

Funding

This project was funded by Tactus Addiction Care and by the Pioneers in Health Care Innovation Fund in 2018.

Data Availability

The datasets generated and analyzed during the current study are not publicly available because they contain sensitive qualitative and clinical data from a vulnerable patient population. Data may be available from the corresponding author on reasonable request, subject to institutional and ethical approval.

Authors' Contributions

Conceptualization: SM, SL
Investigation: SL
Formal Analysis: SL
Writing – Original Draft: SM, SL
Writing – Review & Editing: SM, SL, HK, SK, JVN, DH, RK
Supervision: DH, RK, JVN
All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview on alcohol-related high-risk situations and triggers for immersive virtual reality-cognitive behavioral therapy (IVR-CBT).

XSL•FO
RenderX

[DOCX File, 17 KB - xr_v3i1e82601_app1.docx]

Multimedia Appendix 2
Semistructured interview on coping strategies to reduce nicotine craving in immersive virtual reality (IVR).
[DOCX File, 16 KB - xr_v3i1e82601_app2.docx]

References

1. Rehm J, Shield KD. Global burden of disease and the impact of mental and addictive disorders. Curr Psychiatry Rep 2019 Feb 7;21(2):1-7. [doi: 10.1007/s11920-019-0997-0] [Medline: 30729322]
2. Didden R, VanDerNagel J, Delforterie M, van Duijvenbode N. Substance use disorders in people with intellectual disability. Curr Opin Psychiatry 2020 Mar;33(2):124-129. [doi: 10.1097/YCO.0000000000000569] [Medline: 31743126]
3. van Duijvenbode N, VanDerNagel JEL. A systematic review of substance use (disorder) in Individuals with mild to borderline intellectual disability. Eur Addict Res 2019;25(6):263-282. [doi: 10.1159/000501679] [Medline: 31330514]
4. Seelen-de Lang BL, Smits HJH, Penterman BJM, Noorthoorn EO, Nieuwenhuis JG, Nijman HLI. Screening for intellectual disabilities and borderline intelligence in Dutch outpatients with severe mental illness. J Appl Res Intellect Disabil 2019 Sep;32(5):1096-1102. [doi: 10.1111/jar.12599] [Medline: 31033102]
5. Cosden M. Risk and resilience for substance abuse among adolescents and adults with LD. J Learn Disabil 2001;34(4):352-358. [doi: 10.1177/002221940103400410] [Medline: 15503579]
6. Taggart L, McLaughlin D, Quinn B, McFarlane C. Listening to people with intellectual disabilities who misuse alcohol and drugs. Health Soc Care Community 2007 Jul;15(4):360-368. [doi: 10.1111/j.1365-2524.2007.00691.x] [Medline: 17578397]
7. Slayter EM. Understanding and overcoming barriers to substance abuse treatment access for people with mental retardation. J Soc Work Disabil Rehabil 2008;7(2):63-80. [doi: 10.1080/15367100802009780] [Medline: 19042302]
8. VanDerNagel J, Kiewik M, Didden R. Handboek LVB En Verslaving:'iedereen Gebruikt Toch: Uitgeverij Boom; 2017.
9. Method less alcohol or drugs. Kennisplein Gehandicaptensector. 2016. URL: https://www.kennispleingehandicaptensector.nl/tips-tools/tools/methode-minder-drank-of-drugs [accessed 2026-04-21]
10. Emmelkamp PMG, Meyerbröker K. Virtual reality therapy in mental health. Annu Rev Clin Psychol 2021 May 7;17:495-519. [doi: 10.1146/annurev-clinpsy-081219-115923] [Medline: 33606946]
11. de Oliveira Malaquias FF, Malaquias RF. The role of virtual reality in the learning process of individuals with intellectual disabilities. Technol Disabil 2016;28(4):133-138. [doi: 10.3233/TAD-160454]
12. Keizer A, van Elburg A, Helms R, Dijkerman HC. A virtual reality full body illusion improves body image disturbance in anorexia nervosa. PLoS ONE 2016;11(10):e0163921. [doi: 10.1371/journal.pone.0163921] [Medline: 27711234]
13. Langener S, VanDerNagel J, Klaassen R, Van der Valk P, Heylen D. "Go up in smoke": feasibility and initial acceptance of a virtual environment to measure tobacco craving in vulnerable individuals. Presented at: 2021 IEEE 9th International Conference on Serious Games and Applications for Health(SeGAH); Aug 4-6, 2021. [doi: 10.1109/SEGAH52098.2021.9551854]
14. Standen PJ, Brown DJ. Virtual reality and its role in removing the barriers that turn cognitive impairments into intellectual disability. Virtual Real 2006 Nov 17;10(3-4):241-252. [doi: 10.1007/s10055-006-0042-6]
15. Kip H, Kelders SM, Weerink K, et al. Identifying the added value of virtual reality for treatment in forensic mental health: a scenario-based, qualitative approach. Front Psychol 2019;10:406. [doi: 10.3389/fpsyg.2019.00406] [Medline: 30873093]
16. Geraets CNW, van der Stouwe ECD, Pot-Kolder R, Veling W. Advances in immersive virtual reality interventions for mental disorders: a new reality? Curr Opin Psychol 2021 Oct;41:40-45. [doi: 10.1016/j.copsyc.2021.02.004] [Medline: 33714892]
17. Riva G, Wiederhold BK, Mantovani F. Neuroscience of virtual reality: from virtual exposure to embodied medicine. Cyberpsychol Behav Soc Netw 2019 Jan;22(1):82-96. [doi: 10.1089/cyber.2017.29099.gri] [Medline: 30183347]
18. Valmaggia LR, Latif L, Kempton MJ, Rus-Calafell M. Virtual reality in the psychological treatment for mental health problems: an systematic review of recent evidence. Psychiatry Res 2016 Feb 28;236:189-195. [doi: 10.1016/j.psychres.2016.01.015] [Medline: 26795129]
19. Dellazizzo L, Potvin S, Dou BY, et al. Association between the use of cannabis and physical violence in youths: a meta-analytical investigation. Am J Psychiatry 2020 Jul 1;177(7):619-626. [doi: 10.1176/appi.ajp.2020.19101008] [Medline: 32456503]
20. Ferreri F, Bourla A, Mouchabac S, Karila L. e-Addictology: an overview of new technologies for assessing and intervening in addictive behaviors. Front Psychiatry 2018;9:51. [doi: 10.3389/fpsyt.2018.00051] [Medline: 29545756]
21. Langener S, Van Der Nagel J, van Manen J, et al. Clinical relevance of immersive virtual reality in the assessment and treatment of addictive disorders: a systematic review and future perspective. J Clin Med 2021 Aug 18;10(16):3658. [doi: 10.3390/jcm10163658] [Medline: 34441953]
22. Segawa T, Baudry T, Bourla A, et al. Virtual reality (VR) in assessment and treatment of addictive disorders: a systematic review. Front Neurosci 2019;13:1409. [doi: 10.3389/fnins.2019.01409] [Medline: 31998066]

XSL•FO
RenderX

23. Marissen MAE, Franken IHA, Blanken P, van den Brink W, Hendriks VM. Cue exposure therapy for the treatment of opiate addiction: results of a randomized controlled clinical trial. Psychother Psychosom 2007;76(2):97-105. [doi: 10.1159/000097968] [Medline: 17230050]
24. Conklin CA, Tiffany ST. Applying extinction research and theory to cue-exposure addiction treatments. Addiction 2002 Feb;97(2):155-167. [doi: 10.1046/j.1360-0443.2002.00014.x] [Medline: 11860387]
25. Taubin D, Berger A, Greenwald D, et al. A systematic review of virtual reality therapies for substance use disorders: impact on secondary treatment outcomes. Am J Addict 2023 Jan;32(1):13-23. [doi: 10.1111/ajad.13342] [Medline: 36128667]
26. Girard B, Turcotte V, Bouchard S, Girard B. Crushing virtual cigarettes reduces tobacco addiction and treatment discontinuation. Cyberpsychol Behav 2009 Oct;12(5):477-483. [doi: 10.1089/cpb.2009.0118] [Medline: 19817561]
27. Keijsers GPJ, et al. Protocollaire Behandelingen Voor Volwassenen Met Psychische Klachten 1: Boom; 2017. [doi: 10.1007/978-90-313-7117-4_1]
28. Marlatt GA, Donovan DM. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors: Guilford; 2005.
29. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
30. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
31. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The fagerström test for nicotine dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
32. Littel M, Franken I, Muris P. Psychometric properties of the brief Questionnaire on Smoking Urges (QSU-Brief) in a Dutch smoker population. Neth J Psychol 2011:1-20 [FREE Full text]
33. Sayette MA. The role of craving in substance use disorders: theoretical and methodological issues. Annu Rev Clin Psychol 2016;12(1):407-433. [doi: 10.1146/annurev-clinpsy-021815-093351] [Medline: 26565121]
34. Carroll KM, Kiluk BD. Cognitive behavioral interventions for alcohol and drug use disorders: through the stage model and back again. Psychol Addict Behav 2017 Dec;31(8):847-861. [doi: 10.1037/adb0000311] [Medline: 28857574]
35. Langener S, Murray S, Berkhoff M, Bouman Y, Didden R, VanDerNagel JE. Ontwikkeling van ondersteuning met virtual reality in de verslavingsbehandeling van cliënten met een LVB: een cognitief-gedragstherapeutisch perspectief. Tijdschrift Gedragstherapie 2024;57(4):447-459. [doi: 10.5553/TG/016774542024057004009]
36. Schippers G, Smeerdijk M, Merkx M. Cognitieve gedragstherapie bij problematisch middelengebruik en gokken: inleiding, toelichting en verantwoording. In: M G, Schippers M, Smeerdijk MJM, editors. Handboek Cognitieve Gedragstherapie Bij Middelengebruik En Gokken: Resultaten Scoren; 2014.
37. Fowler C. Virtual reality and learning: where is the pedagogy? Brit J Educational Tech 2015 Mar;46(2):412-422 [FREE Full text] [doi: 10.1111/bjet.12135]
38. Kip H, Kelders SM. Putting the value in VR: how to systematically and iteratively develop a value-based VR application with a complex target group. Presented at: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland p. 1-12. [doi: 10.1145/3290605.3300365]
39. Kouijzer M, Kip H, Bouman YHA, Kelders SM. Implementation of virtual reality in healthcare: a scoping review on the implementation process of virtual reality in various healthcare settings. Implement Sci Commun 2023 Jun 16;4(1):67. [doi: 10.1186/s43058-023-00442-2] [Medline: 37328858]

Abbreviations

AUD: alcohol use disorder
BIF: borderline intellectual functioning
CBT: cognitive behavioral therapy
FTND: Fagerström Test for Nicotine Dependence
HMD: head-mounted display
IVR: immersive virtual reality
MBID : mild to borderline intellectual disability
MID: mild intellectual disability
ND: nicotine dependence
QSU-Brief: Questionnaire of Smoking Urges-Brief
SUD: substance use disorder
VAS: Visual Analogue Scale
VE: virtual environment
VR: virtual reality
VRET: virtual reality exposure therapy

XSL•FO
RenderX

Edited by T Leung; submitted 18.Aug.2025; peer-reviewed by P Schroeder; revised version received 13.Mar.2026; accepted 13.Mar.2026; published 04.May.2026.

Please cite as:
Murray S, Langener S, Kip H, Klaassen R, Kelders SM, Heylen D, VanDerNagel J
Immersive Virtual Reality–Supported Cognitive-Behavioral Therapy for Patients With Mild to Borderline Intellectual Disabilities and Substance Use Disorders: Two Exploratory Studies
JMIR XR Spatial Comput 2026;3:e82601
URL: https://xr.jmir.org/2026/1/e82601
doi:10.2196/82601

© Samantha Murray, Simon Langener, Hanneke Kip, Randy Klaassen, Saskia Marion Kelders, Dirk Heylen, Joanne VanDerNagel. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 4.May.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Virtual Reality in Health Professions Education: Qualitative Descriptive Study of Educators' Perspectives

Tak Wing Yu[1*], BSc, MSc; Michael Rowe[2*], BSc, MSc, PhD; Jose Frantz[3*], BSc, MSc, PhD

[1]Department of Physiotherapy, University of the Western Cape, Community and Health Science Building, Bellville Campus, 14 Blanckenberg Street, Bellville, Cape Town, Western Cape, South Africa

[2]Digital Innovation in Health & Social Care, School of Health & Care Sciences, University of Lincoln, Lincoln, United Kingdom

[3]University of the Western Cape, Physiotherapy, University of the Western Cape, Bellville, South Africa

[*]all authors contributed equally

Corresponding Author:
Tak Wing Yu, BSc, MSc
Department of Physiotherapy, University of the Western Cape, Community and Health Science Building, Bellville Campus, 14 Blanckenberg Street, Bellville, Cape Town, Western Cape, South Africa

Abstract

Background: As virtual reality (VR) technology has become more accessible, the potential of this technology has been increasingly investigated in higher education institutions to improve teaching and learning. While VR can provide immersive experiences to support visualization and active learning, its adoption in health professions education is often limited by high costs, technical complexity, and a lack of pedagogical fit. Furthermore, the educator, who is critical to curriculum design, is underrepresented in the discourse on VR integration, especially in resource-constrained environments in South Africa.

Objective: The purpose of this study was to investigate the perceptions of health professions educators regarding the potential value of immersive VR in the educational process, the implementation barriers, and the requirements for successful curriculum integration at a South African university.

Methods: A qualitative exploratory descriptive design was used. Eighteen educators (N=18) from a wide variety of health disciplines (including physiotherapy, occupational therapy, nursing, and dentistry) were recruited from the Faculty of Community and Health Sciences. Participants interacted with 3 immersive VR applications (The Body VR, Sharecare You VR, and Wraith VR), which were chosen to represent varying levels of clinical complexity. Semistructured interviews were used to explore their perceptions. Data were analyzed using reflexive thematic analysis in accordance with the framework of Braun and Clarke. Sample size adequacy was determined using information power.

Results: Five overall themes were constructed from the data: (1) the experience of VR, where educators enjoyed the engagement but initially felt that VR was a "gaming" novelty; (2) teaching and learning preferences, where VR was considered a potential tool for authentic and individualized learning; (3) challenges of VR, where issues of visual overload, cybersickness, and logistical barriers were reported; (4) clinical competency and patient safety, where VR was valued as a safe space to manage errors; and (5) curriculum integration, where there was a preference for scaffolding the use of VR from foundational anatomy in the junior years to procedural simulation.

Conclusions: Educators have a favorable view of VR as a powerful supplemental tool that can help boost student engagement and bridge the gap between theory and clinical practice. However, successful implementation involves more than purchasing hardware; it requires a strategic "scaffolded integration" approach that ensures VR applications align with specific curricular outcomes. It is the role of institutions to not only provide technical assistance but also pedagogical training and support for infrastructure to transform a technological gimmick into an ongoing educational resource.

(JMIR XR Spatial Comput 2026;3:e52925) doi:10.2196/52925

KEYWORDS

virtual reality; VR; teaching and learning; educators; teachers; health professions education; medical education; technology education; technology-enhanced learning; TEL; education; experiences; attitudes; perspectives; immersive

XSL•FO
RenderX

Introduction

Background

Virtual reality (VR) technology is no longer restricted to the gaming industry; it has entered the classroom as a pedagogical tool in health professions education (HPE) [1]. This step has provided learners with an immersive environment that can facilitate visualization of complex anatomical structures, which cannot be easily created physically without causing patient harm [2]. Extensive evidence exists that VR is useful in facilitating learning, spatial awareness, and motivation [3-6]. However, as the hype around the technology wears off, the question has moved past whether it will work to how we keep it going. This is particularly vital in resource-constrained settings, where implementing new technology can be a challenge amid other priorities within the education system [7].

Even though VR was initially promoted as a low-cost alternative to high-fidelity simulation-based learning (SBL) techniques that require expensive mannequins, consumables, and separate rooms, it does not replace any training tool. The finances are delicate [8-11]. VR involves a high initial cost for equipment, software, and technical support [12]. According to Ravichandran and Mahapatra [12], the initial expenses are high, but VR can be scaled in the long run; once the app is purchased, many thousands of students can use it at a very minimal incremental cost compared to the current expenses of cadaveric dissection or wet laboratories.

Although VR has been cited as having advantages over other pedagogical techniques, including SBL, it also presents realistic challenges in HPE. Two recent studies conducted by Nemani [13] and Soltani and Rostami [14] emphasize that successful integration is often hindered not only by hardware limitations and constraints but also by institutional and organizational constraints, including people (human resources). Furthermore, the exploration of “educator readiness” is required. Educators are the key gatekeepers of the curriculum. If they view VR as a “gimmick” or lack the pedagogical competency to integrate the technology, there is a risk of marginalizing it as a supplementary activity rather than a core learning modality.

This challenge is more pronounced in developing countries such as South Africa, where, due to the digital divide and infrastructural constraints, the use of immersive technologies can be a daunting prospect [15]. Mondal and Mondal [16] noted that to prevent the waste of technology, a detailed alignment with the local curriculum is required. However, educators’ voices on the use of VR are underrepresented in the global literature and are predominantly biased toward high-income nations [17]. Understanding how educators in South Africa view the pedagogical value, ease of use, and logistical overhead of VR is imperative for developing implementation strategies that are culturally and institutionally sustainable.

Therefore, the purpose of this study was to explore the perceptions of health professions educators at a South African university regarding the immersion of VR. Furthermore, the identification of potential pedagogical challenges (such as implementation) and possible solutions and strategies was explored.

Purpose of the Study

Health professions educators are among the key stakeholders in incorporating VR into the curriculum. Therefore, exploring educators’ perspectives on the use of VR is fundamental to the sustainable integration of VR in HPE. Thus, this study seeks to understand educators’ perceptions of VR in HPE, with special reference to perceived educational value, implementation difficulties, and issues in curriculum integration.

Methods

Study Design and Setting

This study used a qualitative, exploratory, and descriptive design within an interpretivist paradigm. This approach was chosen to bring out participants’ lived experiences as health professions educators as they interacted with novel VR technologies, thereby allowing for a rich, context-dependent understanding of their pedagogical perceptions [18,19]. The study was conducted in the Faculty of Community and Health Sciences at one of the research-intensive universities in the Western Cape, South Africa. It was an interdisciplinary setting that encompassed physiotherapy, occupational therapy, nursing, psychology, and dentistry, among other educational areas within the health sciences.

Participant Selection and Recruitment

Academic employees were selected through purposive sampling to ensure that they were actively engaged in teaching at the Faculty of Community and Health Sciences. Advertisements for recruitment were sent via email to the faculty and distributed at department meetings. To secure a wide range of opinions, maximum variation methods were used to select participants from various disciplines with different degrees of teaching experience (junior lecturers, lecturers, senior lecturers, and associate professors) and different levels of technology experience (junior or experienced). Participant recruitment continued until a certain level of information power was achieved [3]. The continuous recruitment ensured adequate specificity and depth to achieve the study’s goals and objectives, and 18 participants were recruited.

VR Intervention and Procedure

All participants were taken through a structured VR workshop prior to the interviews to establish a baseline of experiential knowledge. The lessons were conducted in a special simulation laboratory with HTC VIVE Cosmos VR goggles. The participants were requested to experience 3 particular VR applications, which were selected to illustrate different pedagogical affordances:

1. The Body VR: a narrative storytelling journey that takes the user within the bloodstream, via an immersive approach, throughout the body. The application is designed to illustrate abstract physiological concepts (key characteristics and components: passive and observational).

XSL•FO
RenderX

2. Sharecare You VR: a highly interactive anatomical atlas allowing for the dissection and manipulation of organs (key characteristics and components: active and exploratory).
3. Wraith VR: a procedural simulation requiring manual dexterity and sequence recall (key characteristics and components: active and procedural).

Each session lasted about 20 to 30 minutes, and participants were given time to familiarize themselves with the hardware and interface before reflecting on the experience.

More details on the VR application are provided in Multimedia Appendix 1.

Data Collection

Data were collected via semistructured individual interviews immediately after VR exposure (Multimedia Appendix 2). The semistructured interview guide was developed using the CAMIL (Cognitive Affective Model of Immersive Learning) framework [20]. Makransky and Petersen [20] framework provides a grounded basis for identifying key constructs relevant to the learning context, including interest, motivation, embodiment, and self-regulation. Although the thematic analysis was conducted using the reflexive thematic analysis (RTA) inductive approach, without using CAMIL as an a priori coding framework, it allowed themes and subthemes to emerge from the interviews rather than from existing theory and literature. The primary focus of the interviews revolved around the following:

1. Pedagogical affordances: What are the key components you think are needed for a successful VR integration in the educational system?
2. Implementation barriers: What is hindering the ease of using VR?
3. Curricular alignment: In which module will you implement them (describe how the module is able to link to VR)?

Interviews were audio-recorded with participants' consent. The primary researcher wrote field notes and reflexive memos throughout the data collection process to document nonverbal cues and early analytical thoughts.

Data Analysis

Thematic Analysis

The data analyzed utilized the RTA approach, as outlined by Braun and Clarke [21]. This was an iterative 6-phase process that involved the following:

1. Familiarization: repetitive listening to audio recordings and reading transcripts in order to immerse oneself in the data.
2. Coding: sequential production of systematic semantic codes and latent codes for the entire dataset.
3. Generating themes: pulling together codes into possible themes that represent shared patterns of meaning.
4. Reviewing themes: reviewing themes against the coded extracts and the overall dataset to ensure that they create a coherent narrative.
5. Defining and naming themes: developing the details of each theme to make them clear and distinct.
6. Writing up: writing the final report.

The Role of Artificial Intelligence in Data Analysis

After the primary researcher completed the initial analysis, the artificial intelligence (AI) model was provided with the human-developed codebook and the anonymous semistructured interview transcript for analysis. The initial prompt involved giving the AI model a role; the research aim, objective, and research question were provided (eg, "You're a Professor at a University working in the Physiotherapy Department"). The AI model was instructed to identify excerpts that differed from or codes that contradicted the researchers' primary findings. The results from the AI model were treated as unverified notes, and gaps were identified; they were cross-referenced by the research team against the raw audio and transcript to ensure contextual accuracy and truthfulness. The final decision on the results was determined by the researchers.

Trustworthiness and Rigor

Rigor was established by using the criteria of credibility, transferability, dependability, and confirmability [22].

- Credibility: ensured by spending time with the data and researcher reflexivity. The researchers scheduled quarterly meetings with an institutional researcher to discuss, debrief, and provide feedback and questions about the study's analysis process.
- Transferability: enhanced through the provision of detailed descriptions of the study setting, participants, and the VR intervention to enable the reader to judge whether they might apply to other contexts.
- Dependability and confirmability: an audit trail was maintained for key methodological decisions, including the development of a codebook, subthemes, and themes.

Reflexive Statement

The lead researcher is a physiotherapy lecturer working in the same faculty as all participants, with an understanding of the educational environment and potential biases. A desire to deploy VR in health sciences education, driven by a disciplinary background as a physiotherapist, was recognized at the outset as likely to affect coding and interpretation processes. This was moderated through reflexive memos during data collection and analysis processes to expose and challenge underlying biases. Regular supervisors' meetings helped to externalize thinking. Participants were encouraged to know that both critical and positive responses were sought and valued. During the interviews, participants were reminded of the lead researcher's role. A semistructured interview was used to allow participants to elaborate on their experiences and perceptions rather than simply responding to an investigative prompt.

Ethical Considerations

Ethics approval was granted by the Humanities and Social Sciences Research Ethics Committee of the University of the Western Cape (reference HS22/6/55). An information sheet and a consent form were sent prior to the interviews, and the participants signed the consent form before starting the interview process. All identifiable information was removed from the transcript before data analysis. To protect data privacy during AI processing, only deidentified textual data were input into

XSL•FO
RenderX

ChatGPT. Participants in this study did not receive any stipend or compensation.

Results

Demographics and Background

The participants' ages ranged from 32 to 54 years, with a mean age of 42 (SD 9.44) years. Occupational therapy, physiotherapy, dentistry, nursing, psychology, social work, sport, recreation and exercise science, biokinetics, and the Interprofessional Education Unit were represented. Details of the participants are available in Multimedia Appendix 3.

Themes, Subthemes, and Codes

The analytical process involved the creation of 105 initial codes, inductively developed from the interview data and field notes (Multimedia Appendix 4). These codes represented fragmented ideas, experiences, and perceptions expressed by participants and remained data-driven during initial familiarization and the initial round of coding. Through an iterative and reflexive process, the codes were compared, sorted, and refined to highlight patterns of meaning. Similar codes were then organized into 16 subthemes that captured common pedagogical, experiential, and contextual issues from participants' experiences. These subthemes were grouped into 5 high-level themes that captured higher-level patterns of meaning related to educators' views on VR in HPE. Themes were developed based on analytic consistency and interpretative meaning rather than on counts, as is typically the case in RTA. Consultation between the authors on codes, subthemes, and themes was ongoing throughout the data analysis to ensure the representativeness and fit of the codes and themes. This also enabled them to be defined and redefined through data analysis.

The Educator's Experience in the Use of VR

Before this study, most participants had only heard of or had a surface-level understanding of VR terminology. Some participants had VR experiences related to watching media or videos on their phones. Participant 14, a dentist, said, "Nothing was done before, not even a VR video." This sentiment was echoed by participant 3, a physiotherapist, who noted the transition from gaming to education: "Gaming, I've mostly seen it in gaming and gradually… [they are] trying to bring it into the teaching environment as well." For many of these professionals, the research project provided their first physical encounter with fully immersive virtual worlds. It is important to note that the interview questions addressed VR in general, but the participants' responses were heavily influenced by the high-fidelity, fully immersive experiences provided during the study.

The emotional reaction to this was largely positive, even though there was no prior experience, and it was defined by an emotional wow factor. The most frequently used words by the participants to describe their participation were immersion, excitement, and fun. Indicatively, a psychologist, participant 5, reported feeling deeply emotional about the experience of cellular visualization: "I felt like I was inside the cell. I was struck. I adjusted the wow factor to that. Wow. What is it? I have never experienced anything of the kind; therefore, in my case, it was the first time that I was exposed to the virtual environment." This feeling of belonging to the setting is a characteristic feature of high-presence VR and represents a major change from traditional classroom observation. Participant 3 (physio) pointed out the direct influence of these affective factors in the learning process: "It feels more fun, more fun. It is more immersive, and as I listen to the models, I can see their labels and look around to see where these models or structures fit." This participant also mentioned that this would be more interactive for students; it is more fun, suggesting that the fun of the VR experience might be an effective motivator for student engagement.

Teaching and Learning Preferences

The participants found VR to be a powerful alternative modality that can add value to education through its immersive, interactive nature. Instead of being a passive visual representation, the technology was viewed as a stimulus for various learning orientations, such as exploratory, active, and individualized learning.

Immersive Learning

The ability to immerse oneself in the content deeply, the feeling of being physically present, was considered to be one of the main motivators of retention. Participant 7 (sport science) claimed that the value of VR lies in its ability to aid memory through active participation: "The one value is that it helps you to remember. As you are part of what you are learning." This experiential "being part of" the content facilitates deep encoding and retrieval, as the student is not just observing a process but living it. Participant 9 (sport science) emphasized the importance of spatial perspective in developing clinical insight: "By seeing what happened from a difficult angle, it can help you understand [the condition] after a client has had surgery." This ability to manipulate viewpoints and observe clinical conditions from angles inaccessible in conventional environments provides a deeper level of understanding.

Active and Individualized Learning

Educators stressed the transition from passive receiving to active agency. Participant 7 (ERES) noted the importance of student autonomy: "The fact that you can choose your options, you can go back and choose which part of the body you can do the tests, see where the breakages are, how things are changing and flowing. That makes you feel like you're part of the lesson that's happening, and not just someone who's receiving the old ways of teaching." This sense of participation shifts the student's role from a consumer of information to an active investigator. Participant 15 (dentist) reflected on their own learning style and how VR might have aided their studies: "The type of student I was, and I am currently seeking help with my studying. If I can visualize it, it stays in my memory, as opposed to me going to a lecture, listening and then going home." This highlights the technology's ability to cater to those who struggle with traditional lecture-based formats. Participant 8 (ERES) observed that "It gives the opportunity for the student who learns slightly slower. I'm thinking of individualizing learning. It is almost custom-made for each student. It's like apps on your phone. I could decide how fast we are moving." This customization

XSL•FO
RenderX

addresses the diverse needs of learners and enables a more inclusive educational experience, much as specialized programs can help diverse groups express themselves creatively and overcome repressed emotions.

Authentic Learning

Lastly, the participants discussed the possibility of simulating near-real clinical situations. Participant 14 (dentist) stated, “Ultimately, we would like to strive for authentic learning experiences, and VR could offer that based on the setup scenario.” This can be achieved through VR, which serves as a safe, high-fidelity simulation of a clinical setting or a virtual human body.

Clinical Competencies and Patient Safety

One recurring issue among educators was the risk of student errors in the real-world clinical environment. The participants often placed VR as a necessary, nonthreatening intermediate through which it is possible to experience safe failure. Participant 7 (sport science) found value in the nonthreatening nature of the technology: “I could make mistakes. But I felt it was nonthreatening, which is definitely a good thing.” Participant 6 (social work) pointed to the moral necessity: “Even if the students make a mistake...they realize they learn from making mistakes. However, in the real environment, it is not possible, as you realize the patient might die.” This concept is transformative for clinical training, as it allows students to learn from mistakes that would be unacceptable in a real-world setting.

Students with this safety affordance were directly associated with preclinical preparation and the number of hours of clinical experience. Participant 2 (physio) stated, “VR will improve the student’s confidence. More so, to promote and enhance their experiences during clinical placements. Because students would develop competence prior to working with real patients.” This is supported by the fact that VR can accelerate skill acquisition and enhance overall clinical outcomes. Educators also regarded VR as a means of overcoming institutional barriers, including limited access to hospitals. Participant 8 (sport science) observed that VR could be used when students “cannot go into hospital settings,” as it can “bring the clinical environment to the students” and expose them to complex equipment: “You’re placing them in a [virtual] hospital room where they can see all the valves or machines we use in physio; whatever it is, they’ll be able to see it.” Participant 12 (physio) even proposed a formal recognition of virtual training: “So you could allocate clinical hours as well, not only when they are on block to prepare but also to give them that clinical experience and confidence.” Participant 10 (sport science) suggested that “the students’ confidence and also the patient [functional] outcome increases, maybe it will even increase students’ [clinical] competence.” This highlights a shift toward accepting virtual simulation as a legitimate part of professional clinical hours.

VR and Curriculum Alignment

Participants stressed that VR’s effectiveness depends on its scaffolded integration into the curriculum rather than its use as a generic add-on. Discipline-specific alignment, in which content aligns with the anticipated outcomes of a particular module, was very popular. For example, the participants proposed that anatomical applications are most appropriate in junior years (first and second years), whereas procedural simulations (eg, Wraith VR) should be used with senior students (third and fourth years) or postgraduate students. Participant 10 (sports science) observed that VR might familiarize junior students before they are expected to perform complex tasks, including the first surgical extraction.

Participants’ opinions on the level of integration were inconclusive. Some, such as participants 6 (social work) and 8 (ERE), said they would like VR to become a primary teaching tool, possibly replacing traditional lectures. Others, such as participant 3 (physio), noted that “Theory and textbooks, they do have their place. And I think VR can supplement.” However, participant 10 (sport science) suggested an “equal” balance: “I think it will be equal. I would use that equal...I would send them the theoretical information before the time. Then, the main tool is to put it into practice, because that is my work, applying it.” This variety of opinions suggests that the optimal level of integration may vary across disciplines and specific learning objectives. Irrespective of the level of use, teachers such as participant 8 (ERES) stated that “the technology should be linked to an outcome module,” and that, without understanding the pedagogical fit, the implementation process will be unsuccessful.

Challenges of VR Implementation

The enthusiasm notwithstanding, a number of challenges at different levels were identified, the main ones being cognitive load, cybersickness, technical, logistical, and financial obstacles.

The first problem identified was the risk of visual and mental overload. While powerful visualizations are a strength of VR, they can also be overwhelming for those not accustomed to the stimulus. Participant 11 (ERES) cautioned, “It might be too overwhelming with the visual and hearing all information, all at once, for initial exposure might be a little bit too much if that’s their first exposure to content.” Participant 6 (social work) noted that this stimulus is not suitable for everyone: “This would be beneficial for those who need this kind of [visual] stimulus, and then you have the other side of the spectrum where it would be overwhelming for those who do not.”

There were also complaints of physical discomfort and usability issues. Physical discomfort, specifically “cybersickness,” was reported. Participant 4 (psychologist) described a loss of balance: “I lost a bit of balance. And my knees suddenly became weak.” Participant 2 (physio) experienced visual issues: “At times, it got a bit blurry and then cleared up. And there were lags.” These physiological responses are common in early-stage VR adoption and highlight the need for sessions of appropriate duration and high-quality hardware to minimize vestibular conflict.

Technical obstacles, including the difficulty of controller buttons, served as a major distraction from the learning objectives. Participant 4 (psychologist) noted that “The buttons on the controller were difficult to use; it was complicated.” Participant 14 (dentist) found that the interface interfered with the learning objectives: “The biggest distraction for me is that if I pressed the button, I was either too high or too low on the

XSL•FO
RenderX

wrong side." Participant 2 (physio) also observed that "managing the material, my hands with the controls took a bit of practice." These findings suggest that the user interface must be intuitive to prevent the technology itself from becoming a barrier to learning.

Lastly, institutional obstacles such as cost and scalability were raised. Participant 6 (social work) questioned the university's ability to invest: "We're talking about thousands of rands. I don't know if the university has money to invest." The high cost of headsets and software licenses makes scalability difficult. Participant 4 (psychologist) pointed out that "We need to have more than one, or five for the number of students that we have, for it to be valuable."

Class size is a major logistical hurdle. Participants 3 (physio) and 15 (dentist) asked "A class of 75 - 80 students, how long will it take them to rotate one machine?" This rotation time can be a significant drain on teaching resources and limit the time each student can spend in the virtual environment. For VR to be truly effective in a large-class setting, multiple devices and efficient scheduling are required. The challenge of managing limited resources while providing a high-quality experience is a common theme in fields undergoing professional development.

Discussion

Principal Findings

Our study asked health professions educators for their perceptions after using immersive VR to teach certain courses or modules. The study found that educators perceive VR as an educational tool that motivates students, helps them understand complex concepts, and encourages active learning. Other studies have also found that VR is more motivating and can help students understand spaces and engage in active learning [3-6].

Participants in this study indicated that VR will be useful only when aligned with the course and pedagogy. Participants suggested that VR should be employed as part of a lesson aligned with the Course Intended Learning Outcome, rather than simply for "using" technology in teaching. This approach reinforces the importance of alignment and provides the educator and students with a clear aim and learning goal. As with any technology implementation, the initial attractiveness and novelty will maintain students' attention; however, for long-term educational value, sustainability and support are required [23]. These supports are provided by different levels of the institution. For example, a lab technician (manpower) and storage space (infrastructure) are required [24].

Participants, as educators, highlighted that VR is useful for learning from mistakes without risk to the patient, thus creating a safe environment for students to explore and be unsuccessful without causing actual harm to the patient. A safe learning environment is important for students, as it helps them construct knowledge from the provided content [25]. Besides a safe learning environment, it is important for students to learn from their mistakes [26]. VR can provide immediate feedback to help students review their mistakes and learn from them. Additionally, VR offers high replayability, enabling students to return to the virtual environment and redo the learning activity until the desired outcome is achieved.

Comparison With Prior Work

The high expectations individuals have regarding their affective reactions align with earlier research findings, indicating that immersive VR can positively influence engagement and motivation [4,6]. This is in line with larger trends in the adoption of educational technology, in which emerging tools tend to initially produce a surge of excitement before plateauing into more lasting and meaningful applications [27]. However, this paper provides some sense of the impact of novelty by demonstrating that teachers are aware of it and emphasize the necessity of lasting pedagogical value.

Findings can also be interpreted using the CAMIL [20], particularly regarding the constructs of interest, motivation, embodiment, and self-regulation. The descriptions of immersion, interaction, and a deeper understanding of the space provided by participants indicate the embodiment pathway proposed by CAMIL, in which learners experience the virtual environment and think more profoundly. In the meantime, this paper also illuminates important issues such as cognitive load and usability. The respondents reported visual overload and difficulty navigating the VR interface, particularly at the beginning. Even though CAMIL also acknowledges the influence of cognitive load in the context of immersion, these findings suggest that the extraneous cognitive load may also be driven by a lack of experience with the corresponding technology. This confirms the importance of preparatory measures, such as orientation or pretraining, that can enhance usability and reduce mental load [28].

The focus on curriculum alignment and scaffolded integration can be viewed as a valuable addition to the literature. Although the literature shows that VR is effective in enhancing learning outcomes [5,6], very few studies have examined how VR should be integrated into learning programs. The respondents in this research reported a progressive format in which VR implementation must be introduced in basic subjects, that is, anatomy, during the early years, and its application in more complex simulations of procedures during the later stages. This is consistent with competency-based education systems such as the South African National Qualifications Framework (SANQF) [29], which focus on developmental advancement and conformity with professional skills.

The perceived value of VR as a safe environment for practice and error aligns with the SBL literature [30,31]. Nonetheless, teachers in the present research considered VR as a complement, not a replacement, for conventional teaching and simulation techniques. Although VR has its benefits in terms of scalability and accessibility [12,32], its weaknesses include the lack of haptic feedback, which may limit its ability to substitute for hands-on clinical training.

Lastly, the barriers identified in this study, such as cost, infrastructure, and technical complexity, are similar to those reported in past studies [13,16]. However, this study also highlights that these issues are not only technical but also involve institutional readiness and teachers' abilities, which

XSL•FO
RenderX

justifies the need to examine them holistically for implementation.

Future Directions

The future directions of research in VR are as follows:

1. Future research should focus on longitudinal and experimental studies, as this will provide stronger evidence of VR's effectiveness relative to traditional and blended learning paradigms. The perceived effect of VR on learning, skills, and patient safety has been determined [33-35].
2. Further research could explore a design principle or a set of instructions for integrating VR into the curriculum, given the importance of aligning the learning outcomes with the VR application.
3. From an implementation standpoint, research should be conducted to explore scalable, cost-effective delivery models for integrating VR, especially in low-resource environments. This may involve alternative VR delivery options, such as desktop or mobile platforms, as well as other infrastructure models.
4. Finally, there is a need for future research on integration at the curriculum level, exploring the methodical integration of VR into HPE curricula and mapping it to competency frameworks. This work is critical to support the progression from VR as an innovative tool to VR as an embedded tool in HPE.

Limitations

There are several limitations to consider.

1. To begin with, this research was conducted at a single institution and used a small sample (N=18), which may limit the generalizability of the results to other geographical or institutional settings. Nevertheless, the concept of information power, in which sample adequacy is judged by the relevance and richness of the information in relation to the study aim, was used to assess sample adequacy. In this study, the specificity of the educator sample and the thoroughness of the RTA indicated that the data held sufficient information power to support meaningful interpretation. In addition, participants' years of experience were not collected, even though this could have informed their curriculum and VR-teaching practices.
2. The perceptions were gathered after a brief demonstration; therefore, the responses could also have been influenced by novelty bias rather than by long-term pedagogical usefulness.
3. This research did not measure learning outcomes; although perceptions provide a clear picture of feasibility, further research is needed to quantify objective skills acquisition.
4. Since only 3 applications were shown, the results might not be representative of the overall capabilities of more sophisticated or custom VR systems that are under development.

Conclusion

To sum up, health professions educators view VR as an engaging and potentially disruptive learning tool when designed structurally in line with educational objectives. Although VR offers distinct benefits in motivation, immersion, and scalability, it should be used as a strategic addition to current teaching and simulation techniques. The key to successful implementation in resource-constrained settings lies in shifting the current trend of momentary enthusiasm toward a systematic examination of curricular constraints and intended evaluation methods. VR has great potential to enhance the quality and safety of HPE by addressing institutional barriers and offering pedagogical support.

Acknowledgments

The use of artificial intelligence, ChatGPT (GPT-4; OpenAI), was solely to ensure that the human analysis did not contain blind spots and gaps. It was used as a secondary coder (in addition to the researcher's analysis) for post hoc quality assurance.

Funding

Funding was received from the University of the Western Cape as part of the University Capacity Development Programme under the Department of Higher Education, Science and Technology PhD support initiative.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to the privacy of the participants (university staff) but are available from the corresponding author upon reasonable request.

Authors' Contributions

All authors have contributed to the study. TWY was the primary investigator and conducted the study, analysis, and manuscript writing. MR and JF contributed to the conceptualization, methodology development, project supervision, and writing (review and editing).

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Virtual reality application details.
[DOCX File, 3642 KB - xr_v3i1e52925_app1.docx]

Multimedia Appendix 2
Semistructured interview guide.
[DOCX File, 3644 KB - xr_v3i1e52925_app2.docx]

Multimedia Appendix 3
Faculty of Community and Health Science.
[DOCX File, 3647 KB - xr_v3i1e52925_app3.docx]

Multimedia Appendix 4
Themes, subthemes, and codes.
[DOCX File, 3648 KB - xr_v3i1e52925_app4.docx]

References

1. 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] [Medline: 36141551]
2. Minga C, Ghosh S. Student perceptions on ICT use in teaching and learning in public secondary schools in Mbeya District, Tanzania. J Educ Soc Behav Sci 2024;37(6):26-39. [doi: 10.9734/jesbs/2024/v37i61325]
3. Alaker M, Wynn GR, Arulampalam T. Virtual reality training in laparoscopic surgery: a systematic review & meta-analysis. Int J Surg 2016 May;29:85-94. [doi: 10.1016/j.ijsu.2016.03.034] [Medline: 26992652]
4. Moro C, Štromberga Z, Raikos A, Stirling A. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. Anat Sci Educ 2017 Nov;10(6):549-559. [doi: 10.1002/ase.1696] [Medline: 28419750]
5. Kyaw BM, Saxena N, Posadzki P, et al. Virtual reality for health professions education: systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959. [doi: 10.2196/12959] [Medline: 30668519]
6. Hamilton D, McKechnie J, Edgerton E, Wilson C. Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. J Comput Educ 2021 Mar;8(1):1-32. [doi: 10.1007/s40692-020-00169-2]
7. Samala AD, Rawas S, Rahmadika S, Criollo-C S, Fikri R, Sandra RP. Virtual reality in education: global trends, challenges, and impacts—game changer or passing trend? Discov Educ 2025;4(1):229. [doi: 10.1007/s44217-025-00650-z]
8. Oje AV, Hunsu NJ, May D. Virtual reality assisted engineering education: a multimedia learning perspective. Comput Educ: X Real 2023 Dec;3:100033. [doi: 10.1016/j.cexr.2023.100033]
9. Reen FJ, Jump O, McEvoy G, et al. Student informed development of virtual reality simulations for teaching and learning in the molecular sciences. J Biol Educ 2025 Aug 8;59(4):604-620. [doi: 10.1080/00219266.2024.2386250]
10. Mastrolembo Ventura S, Castronovo F, Nikolić D, Ciribini ALC. Implementation of virtual reality in construction education: a content-analysis based literature review. J Inf Technol Constr 2022 Jan 1;27:705-731. [doi: 10.36680/j.itcon.2022.035]
11. Choi J, Thompson CE, Choi J, Waddill CB, Choi S. Effectiveness of immersive virtual reality in nursing education: systematic review. Nurse Educ 2022;47(3):E57-E61. [doi: 10.1097/NNE.0000000000001117] [Medline: 34657101]
12. Ravichandran RR, Mahapatra J. Virtual reality in vocational education and training: challenges and possibilities. J Dig Learn Edu 2023;3(1):25-31. [doi: 10.52562/jdle.v3i1.602]
13. Nemani S. Barriers and enablers to adopting virtual reality in lower secondary STEAM curricula. J Adv Res Educ 2025 Mar;4(2):1-14. [doi: 10.56397/JARE.2025.03.01]
14. Soltani N, Rostami A. Immersive learning at scale: exploring the feasibility of VR in education. Presented at: IMX ’25: ACM International Conference on Interactive Media Experiences; Jun 3-6, 2025. [doi: 10.1145/3706370.3727859]
15. Faloye ST, Ajayi N. Understanding the impact of the digital divide on South African students in higher educational institutions. Afr J Sci Technol Innov Dev 2022 Nov 10;14(7):1734-1744. [doi: 10.1080/20421338.2021.1983118]
16. Mondal H, Mondal S. Adopting augmented reality and virtual reality in medical education in resource-limited settings: constraints and the way forward. Adv Physiol Educ 2025 Jun 1;49(2):503-507. [doi: 10.1152/advan.00027.2025] [Medline: 40136005]
17. Annamalai N, Uthayakumaran A, Zyoud SH. High school teachers’ perception of AR and VR in English language teaching and learning activities: a developing country perspective. Educ Inf Technol 2023 Mar;28(3):3117-3143. [doi: 10.1007/s10639-022-11275-2]
18. Sørensen JL, Østergaard D, LeBlanc V, et al. Design of simulation-based medical education and advantages and disadvantages of in situ simulation versus off-site simulation. BMC Med Educ 2017 Jan 21;17(1):20. [doi: 10.1186/s12909-016-0838-3] [Medline: 28109296]

XSL•FO
RenderX

19. McIntosh C, Macario A, Flanagan B, Gaba DM. Simulation: what does it really cost? Simul Healthc 2006;1(2):109. [doi: 10.1097/01266021-200600120-00041]
20. 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]
21. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
22. Creswell JW, Miller DL. Determining validity in qualitative inquiry. Theory Pract 2000 Aug;39(3):124-130. [doi: 10.1207/s15430421tip3903_2]
23. Llanos-Ruiz D, Abella-García V, Ausín-Villaverde V. Virtual reality in higher education: a systematic review aligned with the sustainable development goals. Societies (Basel) 2025;15(9):251. [doi: 10.3390/soc15090251]
24. Alfalah SFM. Perceptions toward adopting virtual reality as a teaching aid in information technology. Educ Inf Technol 2018 Nov;23(6):2633-2653. [doi: 10.1007/s10639-018-9734-2]
25. Clapper TC. Creating the safe learning environment. PAILAL 2010;3(2):1-6 [FREE Full text]
26. Yerushalmi E, Polingher C. Guiding students to learn from mistakes. Phys Educ 2006 Nov 1;41(6):532-538. [doi: 10.1088/0031-9120/41/6/007]
27. Dedehayir O, Steinert M. The hype cycle model: a review and future directions. Technol Forecast Soc Change 2016 Jul;108:28-41. [doi: 10.1016/j.techfore.2016.04.005]
28. Jung J, Shin Y, Chung H, Fanguy M. The effects of pre-training types on cognitive load, self-efficacy, and problem-solving in computer programming. J Comput High Educ 2025 Sep;37(3):809-830. [doi: 10.1007/s12528-024-09407-3]
29. Ensor P. The National Qualifications Framework and higher education in South Africa: some epistemological issues. J Educ Work 2003 Sep;16(3):325-346. [doi: 10.1080/1363908032000099476]
30. Daltoè T, Ruth-Herbein E, Brucker B, et al. Immersive insights: unveiling the impact of 360-degree videos on preservice teachers' classroom observation experiences and teaching-quality ratings. Comput Educ 2024 May;213:104976. [doi: 10.1016/j.compedu.2023.104976]
31. Edgar AK, Macfarlane S, Kiddell EJ, Armitage JA, Wood-Bradley RJ. The perceived value and impact of virtual simulation-based education on students' learning: a mixed methods study. BMC Med Educ 2022 Nov 30;22(1):823. [doi: 10.1186/s12909-022-03912-8] [Medline: 36451222]
32. Nubi OJ, Vincent OR. Virtual reality: a pedagogical model for simulation based learning. Presented at: 2020 International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS); Mar 18-21, 2020. [doi: 10.1109/ICMCECS47690.2020.240895]
33. King D, Tee S, Falconer L, Angell C, Holley D, Mills A. Virtual health education: scaling practice to transform student learning: using virtual reality learning environments in healthcare education to bridge the theory/practice gap and improve patient safety. Nurse Educ Today 2018 Dec;71:7-9. [doi: 10.1016/j.nedt.2018.08.002] [Medline: 30205259]
34. Kim HY, Chun J. Effects of a patient experience–based virtual reality blended learning program on nursing students. Comput Inform Nurs 2022 Jul 1;40(7):438-446. [doi: 10.1097/CIN.0000000000000817] [Medline: 34723872]
35. Oh JW, Kim JE. Effectiveness of a virtual reality application-based education programme on patient safety management for nursing students: a pre-test-post-test study. Nurs Open 2023 Dec;10(12):7622-7630. [doi: 10.1002/nop2.2001] [Medline: 37767936]

Abbreviations

AI: artificial intelligence
CAMIL: Cognitive Affective Model of Immersive Learning
HPE: health professions education
RTA: reflexive thematic analysis
SANQF: South African National Qualifications Framework
SBL: simulation-based learning
VR: virtual reality

Edited by I Steenstra, T Leung; submitted 06.Feb.2024; peer-reviewed by A Hassan, R Bidkar, S Güngör, T NaseriBooriAbadi, Y Lei; revised version received 02.Apr.2026; accepted 02.Apr.2026; published 26.Jun.2026.

Please cite as:
Yu TW, Rowe M, Frantz J
Virtual Reality in Health Professions Education: Qualitative Descriptive Study of Educators' Perspectives
JMIR XR Spatial Comput 2026;3:e52925
URL: https://xr.jmir.org/2026/1/e52925
doi:10.2196/52925

XSL•FO
RenderX

© Tak Wing Yu, Michael Rowe, Jose Frantz. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 26.Jun.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predictive Factors of Augmented Reality–Based Clinical Task Performance Among Novice Users: Cross-Sectional Quantitative Study

Amogh J Vellore[1], BS; Shovan Bhatia[1], BS; Michael R Kann[1,2], BE; Nicolás M Kass[3], MD; Regan M Shanahan[1], BA; Jacquelyn Jardini[1], BS; Jayne Miner[1]; Sohail R Daulat[1], BS; Griffin Hurt[4], BPhil; Rishi Basdeo[5], MS; Nicole Don[1], MA; Jacob T Biehl[4], PhD; Edward G Andrews[1], MD

[1]Department of Neurological Surgery, University of Pittsburgh Medical Center, 200 Lothrop St, STE B-400, Pittsburgh, PA, United States

[2]Department of Orthopaedic Surgery, University of Pittsburgh Medical Center, Pittsburgh, PA, United States

[3]Department of Plastic Surgery, University of Pittsburgh Medical Center, Pittsburgh, PA, United States

[4]Department of Computer Science, School of Computing and Information, University of Pittsburgh, Pittsburgh, PA, United States

[5]Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh, PA, United States

Corresponding Author:
Edward G Andrews, MD
Department of Neurological Surgery, University of Pittsburgh Medical Center, 200 Lothrop St, STE B-400, Pittsburgh, PA, United States

Abstract

Background: Augmented reality (AR) can provide risk-free training for medical trainees, yet little is known about which learner characteristics facilitate adoption or inform training design.

Objective: We aimed to identify which learner characteristics predict AR performance in novices. We hypothesized that higher visuospatial ability and greater video game experience would be associated with faster completion times and fewer errors.

Methods: In this cross-sectional study, 21 undergraduate, graduate, and medical students (median age 22, IQR 21-24 years) without previous AR experience were recruited between June and December 2024. Participants completed a technology experience survey, the mental rotation task (MRT) for visuospatial ability, a standardized 7-task AR protocol mimicking clinical use on the Microsoft HoloLens 2 (hologram manipulation, orbit tracing, anatomical plane visualization, and hologram-to-object registration), and the National Aeronautics and Space Administration Task Load Index for cognitive load assessment. Outcome measures included completion time, slips (unintentional errors), and tracing quality.

Results: All analyses used a significance of α=.05. MRT scores did not predict baseline performance time (Pearson r=0.15, 95% CI −0.32 to 0.55; P=.54) or error rates (r=0.18, 95% CI −0.27 to 0.57; P=.43). Participants with extensive video game experience (>5 hours/week) made fewer slips (unpaired t test; mean difference −2.62 slips, 95% CI −5.19 to −0.04; P=.047), without faster completion times (Mann-Whitney test; median difference −22 seconds, 95% CI −7.00 to 57.00; P=.24). Video game experience did not predict baseline performance time (Pearson r=−0.35, 95% CI −0.69 to 0.13; P=.14). Significant learning effects emerged in unadjusted analyses: completion times decreased on attempts 2 and 3 compared with attempt 1 (mixed-effects analysis: mean difference 28.75 seconds, 95% CI 12.98-44.52; P<.001; 28.00 seconds, 95% CI 10.75-45.25; P=.002, respectively) with fewer slips (Friedman test: χ^2_2=17.8; P<.001; Dunn post hoc: P=.008 and P<.001, respectively). Orbit tracing (Wilcoxon test: median difference −5 seconds; P=.004) and virtual landmark placement times improved (Friedman test: χ^2_3=14.6; P=.002; Dunn post hoc; P=.009 and P=.02), but physical landmark placement did not. Covariate-adjusted models revealed no significant trial-by-covariate interactions.

Conclusions: Visuospatial ability does not predict clinically relevant AR performance, while extensive video game experience was associated with fewer errors. Despite previous studies emphasizing inherent learner characteristics in laparoscopy and endoscopy, covariate-adjusted models showed that AR learning curves were not significantly modified by MRT or video game experience. These findings suggest that early AR performance improvements among novice users are primarily driven by learning rather than visuospatial ability, supporting training approaches that emphasize structured practice, although the modest sample size limits detection of smaller effects.

(JMIR XR Spatial Comput 2026;3:e81236) doi:10.2196/81236

XSL•FO
RenderX

KEYWORDS

mixed reality; augmented reality; virtual reality; mental rotation task; visuospatial ability; medical education; video games

Introduction

The rise in augmented reality (AR) and virtual reality (VR) technology has greatly impacted a range of industries, including education, entertainment, and medicine [1,2]. AR enables the supplementation of real-world visibility with digital information, which can be shown through projections onto head-mounted displays (HMD) on headsets, smart glasses, or tablet-based displays. Within medicine, AR and VR applications continue to grow. While outcomes research remains preliminary given AR's relative infancy, studies have found that AR subjectively increases surgeon confidence in delineating tumor margins [3]. This observation was validated by a multicenter randomized controlled trial (n=113), which observed that AR-guided robotic prostatectomies were associated with a significant decrease in subsequent positive surgical margins, a key prognostic indicator for patient survival [4]. Other measured improvements have included decreased fluoroscopy time needed to navigate difficult tissue structures [5,6]. This technology has also expanded patient education [7,8] by helping patients gain a deeper understanding of their bodies and diseases while also demonstrably decreasing procedural anxiety [8,9] and improving satisfaction [8].

As applications of AR and VR continue to expand across specialties, these technologies hold tremendous potential as risk-free training modalities, allowing medical students and resident physicians to practice procedures without jeopardizing patient safety [10]. Recent literature has shown that AR can help resident physicians learn to identify aneurysms in surgical videos [11], support medical student and resident education as a reliable and predictive simulation-based medical education modality [12-14], and minimize mental workload while simultaneously improving learning capacity [15,16].

Despite this promise, there are still some important factors to consider. Although previous studies indicate that AR can increase mental resource availability [15], enhance working memory capacity [16], and facilitate long-term information storage [16], it may also serve as a distraction for some learners [17]. Research has shown broad educational benefits, from early childhood learning in preschool [18] to secondary education [19] and postgraduate medical education [13]. However, the extent of AR integration in medical education remains varied [20].

Within medical education specifically, previous studies have indicated mixed learning outcomes. For example, AR can be beneficial for anatomy learning compared to virtual dissection tables, but not when compared to the conventional atlas method [21]. Similarly, other studies have found no difference in learning among stereoscopic 3D AR models, monoscopic 3D desktop models, or conventional atlas learning [17]. Further complicating its role, evidence suggests that individuals who have lower spatial ability, as measured by mental rotation tasks (MRTs), may benefit more from AR than their peers with higher MRT scores [21,22]. These findings indicate that the mixed effects of AR within medical education may be explained by individual differences in spatial ability.

Despite the importance of spatial ability across industries, including STEM [23-25] (science, technology, engineering, and math) and medicine [26-29], and the growing adoption of AR within medicine [30], there is still a critical gap in our understanding of how novice AR users learn to use the technology. Previous experiences, such as video game experience, have been shown to play a role in spatial ability [31] as well as in medically relevant tasks [32,33]. More recently, studies have demonstrated that video game experience is a strong predictor of baseline skills in gastrointestinal endoscopy learners [34] and of baseline performance in nonmedical VR tasks [35].

However, it remains unclear which learner characteristics (eg, visuospatial ability and previous video game experience) support the efficient adoption of AR in clinical applications and whether short, targeted exposure is sufficient for novice users to reach proficiency. This study addresses this gap by quantifying novice performance and short-term learning on a neurosurgical AR navigation task and examining how these outcomes relate to individual differences in mental rotation ability and video game experience. We hypothesize that individuals with higher visuospatial ability and, specifically, more video game experience will complete AR-based neurosurgical navigation tasks more quickly and with fewer errors. These results may indicate whether specific learner characteristics confer an advantage in AR or whether novice performance in AR is primarily influenced by learning.

Methods

Research Design

This study used a cross-sectional framework in which participants were recruited using convenience sampling to complete a pretest demographics survey and an assessment of visuospatial ability, followed by a series of standardized AR tasks and a posttest National Aeronautics and Space Administration Task Load Index (NASA-TLX) survey to assess subjective mental load. This paper was prepared in accordance with the Journal Article Reporting Standards [36].

Inclusion and Exclusion Criteria

Participants comprised undergraduate, graduate, and medical students at the University of Pittsburgh between June 2024 and December 2024. Participants who had previous experience using AR were excluded.

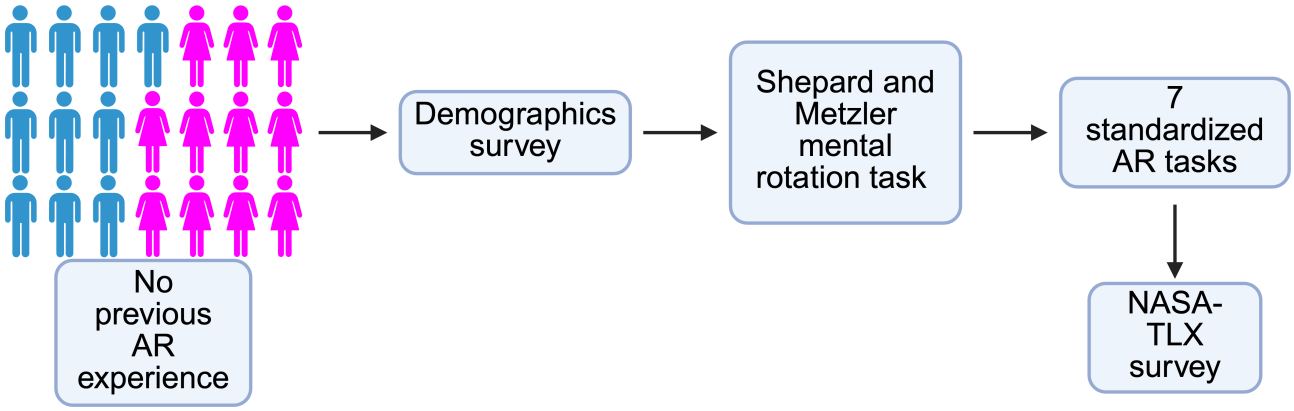
Ethical Considerations

Participants gave their informed consent for participation in the study, for their performance to be recorded for analysis, and for any secondary analyses without additional consent. Participants were not compensated. The authors confirm that there are no images or identifiable features within this manuscript. All participant information was deidentified, and study data were

XSL•FO
RenderX

stored in an encrypted location. This study received institutional review board approval from the University of Pittsburgh (STUDY22040182). The study workflow is shown in Figure 1.

Figure 1. Study design. In total, 23 participants were recruited for this study between June 2024 and December 2024. Two participants did not successfully complete all tasks and were excluded from the analysis, resulting in a final cohort of 21 participants. There were 11 female (pink) and 10 male (blue) participants with no previous experience with augmented reality (AR). The demographics survey collected information such as experience with video games, comfort with new technology, and educational background. All participants then performed a series of mental rotation tasks before completing 7 standardized AR tasks. Following completion of the tasks, participants were given a posttest National Aeronautics and Space Administration Task Load Index (NASA-TLX) survey to assess workload.



Surveys

Two pretest tasks were administered. The first task was a survey that collected demographic and experience information such as age, sex, level of education, experience with video games, comfort with new technology, and experience with surgical devices. The second pretest task was the MRT, a standardized paper-and-pencil measure of 3D spatial visualization derived from the mental rotation paradigm by Shepard and Metzler [37]. The MRT requires participants to decide whether comparison figures are rotated versions or mirror images of a target 3D object, providing a robust index of individual differences in mental rotation ability. Classic psychometric work has shown that the MRT has high internal consistency as indicated by the Kuder-Richardson Formula 20 (Kuder-Richardson Formula 20=0.88), which estimates how consistently dichotomously scored items measure the same underlying construct. Classic psychometric work has also demonstrated that the MRT has high test-retest reliability (r=0.83) [38], and subsequent reviews describe it as one of the most commonly used and well-validated measures of spatial ability [39,40]. Moreover, mental rotation tests such as the MRT are routinely incorporated into spatial ability batteries and reliably predict performance in applied visuospatial tasks (eg, engineering design, navigation, and surgical endoscopy) [29,41,42]. Because our experimental tasks required participants to infer 3D relationships from 2D displays and mentally transform object orientations, we selected the MRT as the primary measure of visuospatial ability.

This task was composed of 2 sets of 12 problems. Participants were allotted 3 minutes to complete each set of questions. During this task, participants were given a warning when their remaining time reached 2 minutes, 1 minute, 30 seconds, and 10 seconds. Following AR testing, participants were given the NASA-TLX survey, a clinically validated metric for measuring mental load [43].

Experimental Procedure

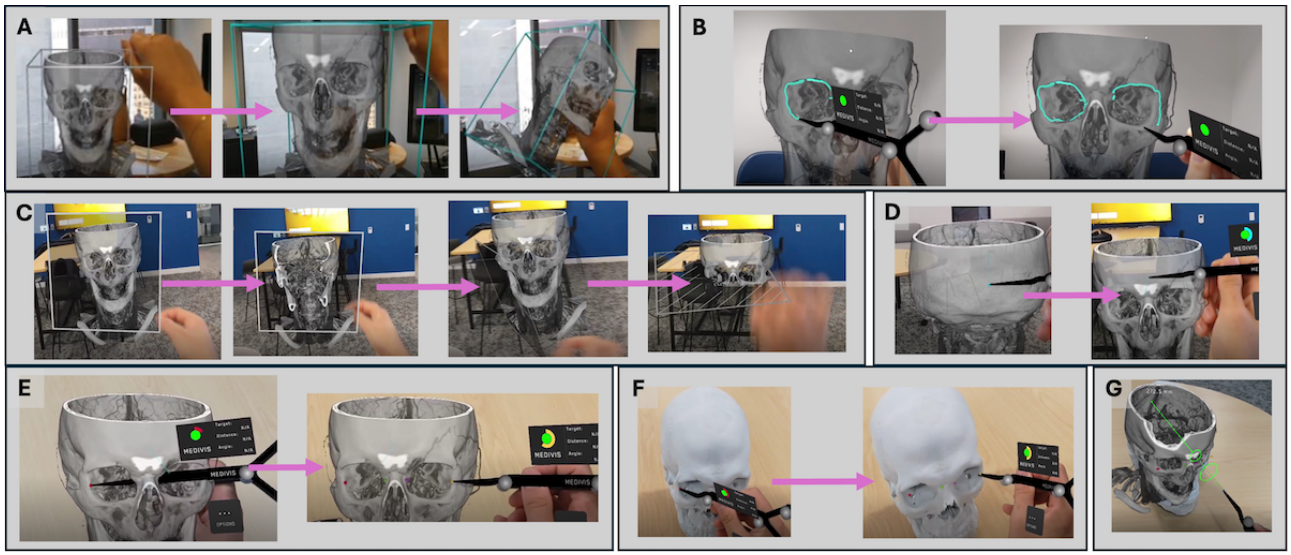
This study conducted AR-based tasks using SurgicalAR (version 1.6.1; Medivis Inc) software on Microsoft HoloLens 2. SurgicalAR is a surgical guidance system that volumetrically renders Digital Imaging and Communications in Medicine data and projects it onto an HMD, allowing for direct registration to patients. Participants were shown a generic, deidentified computerized tomography angiogram of the head. For tasks that required a stylus or pointer, a stylus tracked by the SurgicalAR system was used.

Participants were given 7 different AR-based tasks that were deliberately selected to resemble the clinical workflow steps that a neurosurgeon would perform in the operating room. Specifically, tasks 1 to 3 mimicked basic hologram interactions that may be performed while visualizing key structures or planning an operative approach. Tasks 4 to 7 were designed to follow a standard hologram-to-object registration in which 4 corresponding points were placed on the hologram and the physical object. Then, a 3D transformation was computed using the method by Horn [44] to complete the registration.

Before participants began using AR, the study moderator demonstrated the task using the HoloLens 2 while participants viewed the task through the SurgicalAR system cart monitor, which was positioned near the moderator. Then, the moderator gave and adjusted the headset on each participant and instructed them on basic gesture interactions. All tasks were performed on and recorded using the HoloLens 2. Videos were analyzed for performance using predefined metrics, as defined below. A description of each task is provided below, and representations of the tasks can be seen in Figure 2.

XSL•FO
RenderX

Figure 2. Series of augmented reality tasks that participants were required to complete. (A) Baseline performance: resizing and rotating a hologram of a human skull model; (B) orbit tracing: outlining the orbital rims on the hologram; (C) plane visualization: viewing coronal, sagittal, and axial planes of the hologram; (D) anterior-posterior trajectory point: placing virtual trajectory markers on the hologram; (E) virtual landmark placement: placing 4 virtual landmarks on the hologram; (F) physical landmark placement: placing 4 physical landmarks on the 3D-printed human skull model; and (G) trajectory alignment: performing trajectory alignment.



- Task 1 (Figure 2A): participants resized and rotated a hologram of a human skull model. This task required participants to unanchor the hologram, detaching it from its fixed position and allowing free movement. They then needed to make the hologram larger (zoom in) and smaller (zoom out) and rotate the hologram 360°. Finally, participants reanchored the hologram, locking it back into its original orientation and size. This task was repeated 3 times. Task performance was measured by time taken to complete and by number of slips. Slips were defined as unintentional errors or mistakes [45,46].
- Task 2 (Figure 2B): participants outlined the orbits (eye sockets) of the hologram. Participants were instructed to perform the orbit tracing in one continuous motion for each orbit, without retracting the areas they had already outlined. Performance was measured by a qualitative analysis of orbit tracing quality.
- Task 3 (Figure 2C): participants moved a cut-plane tool fully through the hologram of computerized tomography angiogram of the head in 3 directions—coronal, sagittal, and axial. They were instructed to perform the task while keeping their body facing the front of the hologram. Performance was measured by the number of slips, defined as instances in which a person intends to do one action but unintentionally does something else [45,46].
- Task 4 (Figure 2D): participants placed 2 virtual trajectory landmarks. The first point was placed midline on the lambdoid suture, and the second point was placed midline on the frontal bone. Performance was measured by the time required to successfully place the posterior point and anterior point.
- Task 5 (Figure 2E): participants placed 4 virtual landmark points on the bilateral lateral and medial parts of the hologram's orbit. They began with the lateral left orbit and worked from left to right, finishing with the lateral right orbit. Performance was measured by the time to place each virtual landmark point.
- Task 6 (Figure 2F): participants placed 4 physical landmark points on a 3D-printed skull model, matched to the same locations as the virtual landmarks. They began with the lateral left orbit and worked from left to right, finishing with the lateral right orbit. Performance was measured by the time to place each physical landmark point.
- Task 7 (Figure 2G): participants registered the holographic computerized tomography projection onto the physical skull and then activated the trajectory alignment tool. To accomplish this, participants used the stylus to make the anterior-posterior trajectory turn green, indicating successful alignment. Performance was measured by time taken to align trajectory.

Statistics

Descriptive statistics were used to summarize demographic variables and baseline characteristics. Group comparisons were performed using independent samples 2-tailed *t* tests for continuous variables and Pearson χ^2 tests for categorical variables, where appropriate. To assess learning effects, mixed-effects models with Tukey multiple comparisons were used for completion times. Friedman tests with Dunn post hoc comparisons were used for error counts and landmark placement times, and Wilcoxon signed-rank tests were used for paired comparisons. The overall effect of trial on performance was evaluated using a 1-way repeated-measures ANOVA with Greenhouse-Geisser correction. To evaluate whether learning effects were modified by MRT scores or video game experience, a covariate-adjusted repeated-measures general linear model was used. Linear regression was used to evaluate the predictive relationship between MRT scores and baseline task performance. Participants were stratified based on video game experience (>5 hours/week vs ≤5 hours/week) to assess group differences in task outcomes. Significance was set at α=.05 for all comparisons. All statistical analyses were conducted using GraphPad Prism (version 10.0.0; GraphPad Software Inc).

XSL•FO
RenderX

Results

Overview

In total, 23 participants with no previous experience with AR were recruited for this observational study. Of these, participants 3 and 42 (8.69%) did not successfully complete all tasks and were excluded from the analysis, resulting in a final cohort of 21 (91.3%) participants. Within the final cohort, there were 11 (52.4%) female participants, and the median age was 22 (IQR 21-24) years. There were 15 (71.4%) participants who were undergraduate students. In total, 13 (61.9%) participants spent between 0 to 5 hours per week playing video games, and 10 (47.6%) participants spent between 0 to 10 hours per week interacting with a touch screen device or computer. Furthermore, 13 (61.9%) participants were completely comfortable with new technology. Specific demographic information and comfort with new technology are presented in Table 1. Results of the NASA-TLX are presented in Table 2.

XSL•FO
RenderX

Table . Participant demographics (N=21).

Variable	Value
Sex, n (%)	
Female	11 (52.4)
Male	10 (47.6)
Age (years), median (range; IQR)	22 (19-25; 21-24)
Level of training, n (%)	
Undergraduate student	15 (71.4)
Medical student	5 (23.8)
Master's student	1 (4.76)
Time spent playing video games per week (hours), median (range; IQR)	5 (0-55; 1.5-21)
Weekly video games use (hours), n (%)	
0-5	13 (61.9)
6-10	1 (4.76)
11-15	2 (9.52)
16-20	0 (0)
≥21	5 (23.8)
Time spent interacting with touch screen device or computer (hours), median (range; IQR)	15 (0-63; 5-40)
Weekly touch screen devices or computer use (hours), n (%)	
0-10	10 (47.6)
11-20	1 (4.76)
21-30	3 (14.3)
31-40	3 (14.3)
≥41	4 (19.0)
Comfort with new technology (scale 1-5), n (%)	
Totally comfortable (5)	13 (61.9)
Very comfortable (4)	4 (19.0)
More or less comfortable (3)	3 (14.3)
Not very comfortable (2)	1 (4.76)
Not comfortable at all (1)	0 (0)
Experience with other forms of surgical guidance, n (%)	
Endoscopy[a]	2 (9.52)
DaVinci[b]	1 (4.76)
Microsurgery	0 (0)

[a]Average experience with endoscopy was 13 (SD 4.49) hours.

[b]Total experience with DaVinci was 6 hours.

XSL•FO
RenderX

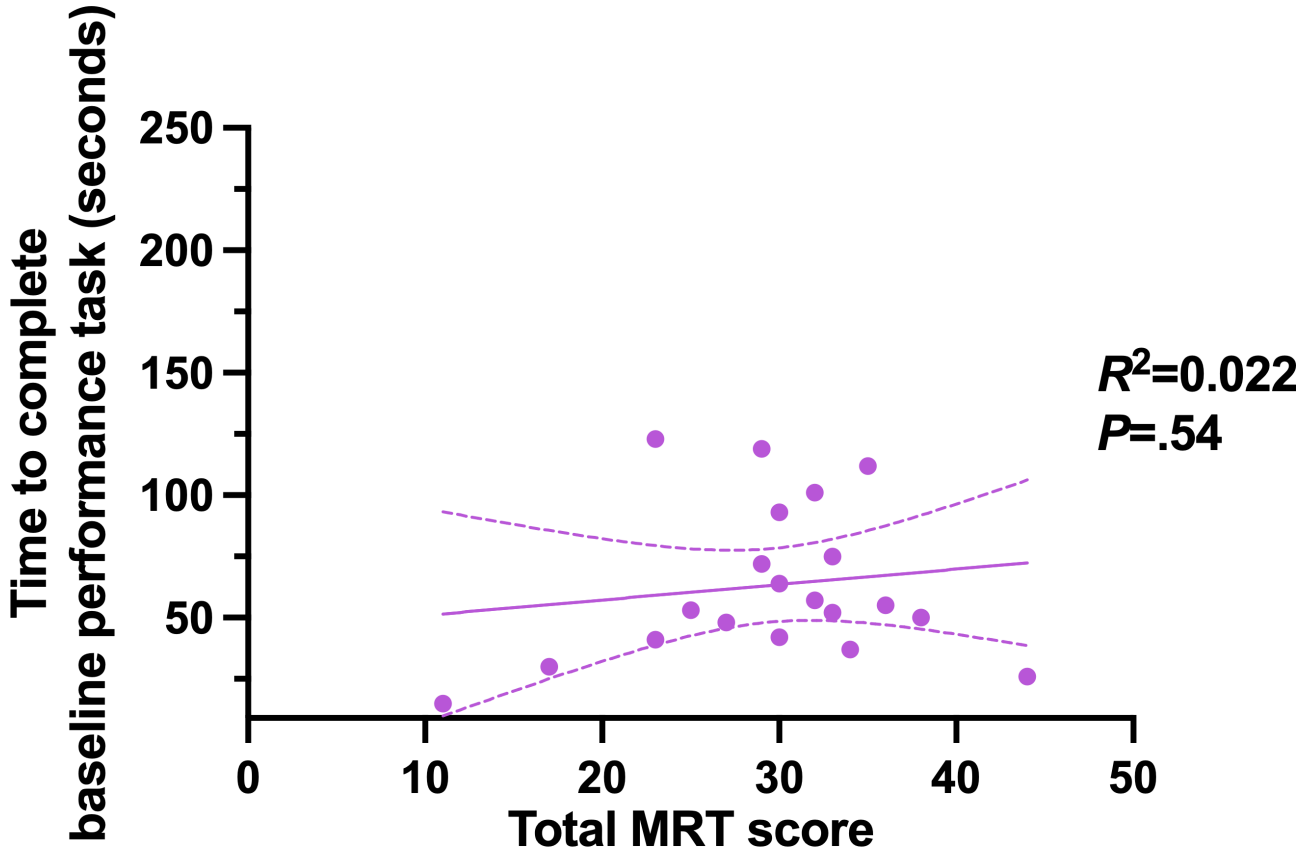
Table . National Aeronautics and Space Administration Task Load Index scores.

Category	Value, median (range; IQR)
Mental demand	50 (0-80; 25-67.5)
Physical demand	15 (0-70; 5-30)
Temporal demand	35 (0-75; 10-50)
Performance (lower is better)	50 (20-85; 37.5-67.5)
Effort	50 (0-90; 30-67.5)
Frustration	30 (0-85; 12.5-60)

Overall Performance on MRT

Visuospatial ability, as measured by the MRT, did not predict the time taken to complete the baseline performance task (Pearson r=0.15, 95% CI −0.32 to 0.55; R^2=0.022; P=.54; Figure 3). Similarly, MRT scores did not predict error rates on the baseline performance task (r=0.18, 95% CI −0.27 to 0.57; R^2=0.034; P=.43). There were no statistically significant differences in baseline performance time (P=.65) or number of slips (P=.62) between individuals with MRT scores ≥30 and those with scores <30.

Figure 3. Association between visuospatial ability, measured by the mental rotation task (MRT), and baseline task completion time. Each dot represents an individual participant's MRT score and corresponding completion time. The solid line indicates the linear regression fit, and the dashed lines represent the 95% CI of the regression. MRT score does not predict the ability to complete the baseline performance task (R^2=0.022; P=.54).



Video Game Performance

Participants were split into 2 groups based on video game experience (group with "extensive" experience of >5 hours/week and group with "minimal" experience of ≤5 hours/week) for analysis. This distribution was determined empirically to yield approximately equal participants per group (8 and 13, respectively). Participants with extensive video game experience did not demonstrate faster completion times compared to those with minimal video game experience (Mann-Whitney test; median difference −22 seconds, 95% CI −7.00 to 57.00; P=.24). However, individuals who had extensive video game experience made fewer slips on average than those who had minimal video game experience (mean 4.00, SD 2.27 slips, 95% CI 2.10-5.90 vs mean 6.61, SD 3.36 slips, 95% CI 4.59-8.64; unpaired t test; mean difference −2.62 slips, 95% CI −5.19 to −0.04; P=.047).

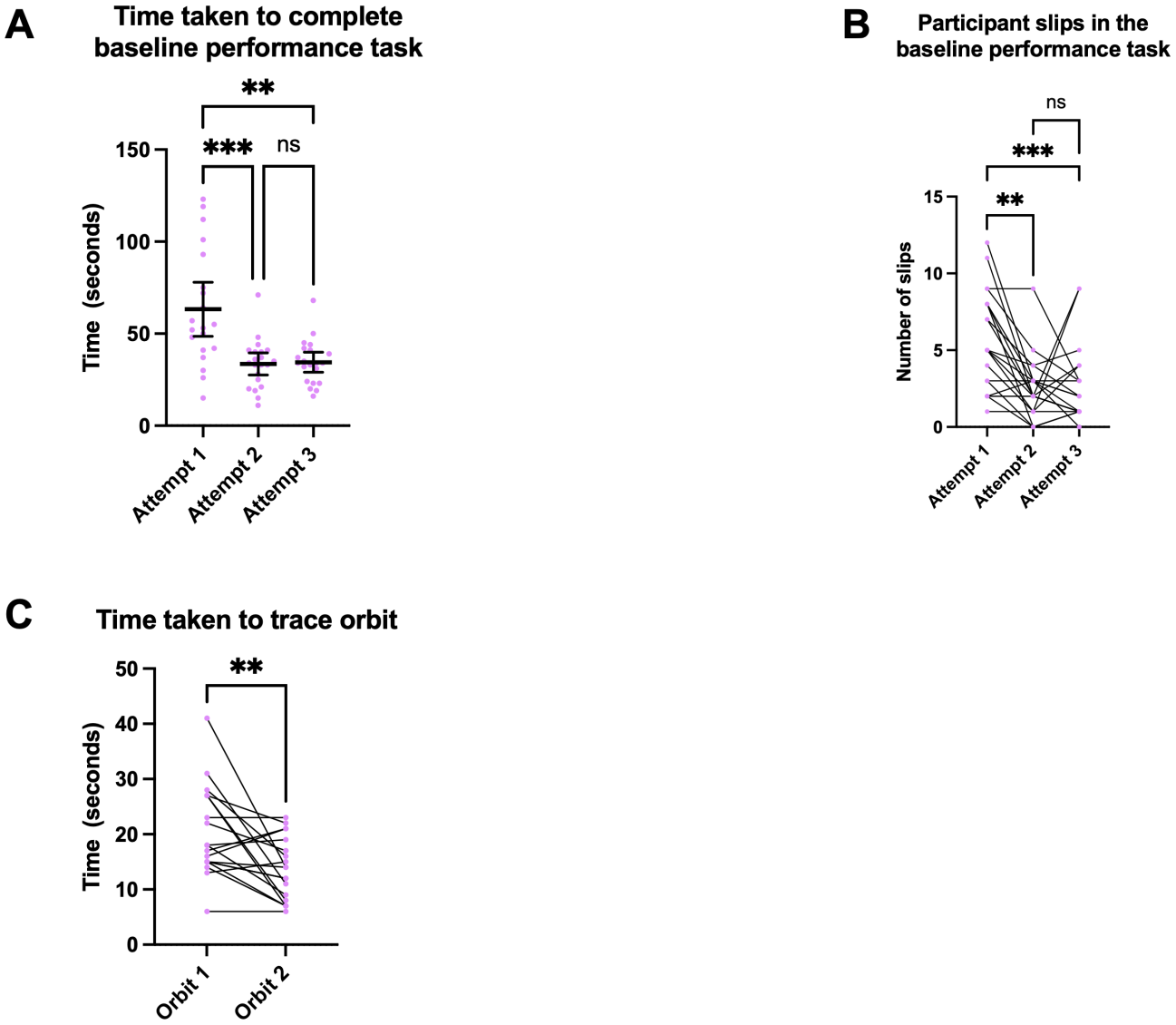
Learning

Participants learned to perform the baseline performance task (task 1) in a significantly shorter time between attempts 1 and 2 (mean 63.3, SD 31.4 seconds, 95% CI 48.5-78.0 vs mean 33.6, SD 13.1 seconds, 95% CI 27.6-39.6; P<.001) and attempts

XSL•FO
RenderX

1 and 3 (mean 63.3, SD 31.4 seconds, 95% CI 48.5-78.0 vs mean 34.5, SD 12.0 seconds, 95% CI 29.0-40.0; *P*=.002), but not between attempts 2 and 3 (mean 33.6, SD 13.1 seconds, 95% CI 27.6-39.6 vs mean 34.5, SD 12.0 seconds, 95% CI 29.0-40.0; *P*>.99; Figure 4A).

Figure 4. Learning effects observed across various augmented reality tasks. (A) Time taken to complete baseline performance task across attempts. Each dot represents an individual participant, and the horizontal lines indicate the mean with 95% CIs. Participants performed the baseline performance task in a significantly shorter time between attempts 1 and 2 [***P<.001] and attempts 1 and 3 [**P=.002]. The comparison between attempts 2 and 3 was not significant [ns; P>.99]. (B) Number of slips in the baseline performance task across attempts. Each dot represents an individual participant, and the connecting lines track each participant's performance across attempts. Participants improved in the accuracy of completing the baseline performance task, as demonstrated by fewer slips between attempts 1 and 2 [**P=.008] and attempts 1 and 3 [***P<.001]. The comparison between attempts 2 and 3 was not significant [ns; P>.99]. (C) Time taken to trace orbits. Each dot represents an individual participant, and the connecting lines track each participant's performance between orbits. Participants significantly improved the time to trace the orbits on the second attempt [**P=.004].



A 1-way repeated-measures ANOVA indicated that there was a significant effect of trial time, consistent with improved performance across trials (Greenhouse-Geisser $F_{1.135,22.691}$=11.890; *P*=.002). However, a covariate-adjusted repeated-measures general linear model that included weekly gaming hours and MRT score indicated that the trial effect was not significant (Greenhouse-Geisser $F_{1.112,20.021}$=0.050; *P*=.85), and there was no evidence that trial-related changes depended on either covariate (trial×gaming hours: *P*=.62; trial×MRT: *P*=.34).

In addition to faster task completion times between attempts 1 and 2 and 1 and 3, participants also performed the task more accurately. There was a decrease in the number of slips between attempts 1 and 2 (mean 5.62, SD 3.2 slips, 95% CI 4.2-7.1 vs mean 2.48, SD 2.02 slips, 95% CI 1.6-3.4; *P*=.008) and attempts 1 and 3 (mean 5.62, SD 3.2 slips, 95% CI 4.2-7.1 vs mean 2.57, SD 2.52 slips, 95% CI 1.4-3.7; *P*<.001; Figure 4B). Furthermore, participants completed the orbit tracing more quickly between the first and second orbit (mean 20.4, SD 8.15 seconds, 95% CI 16.5-24.4 vs mean 13.9, SD 5.52 seconds, 95% CI 11.31-16.5; *P*=.004) without a change in quality of orbit tracing (*P*=.77; Figure 4C).

Additionally, participants required less time to place virtual landmark points 2 and 3 (mean 7.2, SD 4.82 seconds, 95% CI

XSL•FO
RenderX

5.0-9.4 vs mean 4.2, SD 1.81 seconds, 95% CI 3.4-5.0; *P*=.009) and points 2 and 4 (mean 7.2, SD 4.82 seconds, 95% CI 5.0-9.4 vs mean 4.0, SD 1.82 seconds, 95% CI 3.2-4.8; *P*=.02). There was no learning effect observed for placing physical landmarks (Friedman test; Dunn post hoc: all pairwise *P*≥.64). There was no significant difference in the learning curve between participants with MRT scores ≥30 and those with scores <30 (*P*=.87). Furthermore, there was no difference in learning curves between participants with extensive video game experience and those with minimal video game experience (*P*=.81).

Predictive Variables

MRT performance did not predict baseline performance, as measured by task 1 (*P*=.54; Figure 3). Additionally, video game experience was not a predictor of baseline performance (Pearson *r*=−0.35, 95% CI −0.69 to 0.13; R^2=.12; *P*=.14); however, it did predict the number of slips (*P*=.046).

Discussion

Principal Findings

As AR technology continues to improve and integrate within health care and other industries, it becomes increasingly important to understand which factors contribute to technological proficiency among novice AR users. By identifying these factors, product designers can address the scarcity of implementation models that is hindering the widespread adoption of AR and VR in clinical settings [30] and develop programs to help guide novice users through more complex AR-based interactions, thereby proactively addressing areas of difficulty, minimizing the user learning curve, and increasing user adoptability. To address this growing need, our study aimed to identify predictors of performance in novice AR users. Our findings suggest that visuospatial ability does not predict AR task completion time, though extensive video game experience was associated with greater accuracy. Despite this result, neither visuospatial ability nor video game experience corresponded with an improved learning curve.

Predictive Variables of Performance Gains

Existing literature has placed a strong emphasis on visuospatial ability as a predictor of performance in various clinical settings, including ultrasound [27], laparoscopic [28], and endoscopic procedures [29], as well as in nonclinical settings [47,48] and learning [26]. Given that factors such as depth perception and stereovision undoubtedly contribute to an individual's visuospatial ability [49], our study used one of the most popular validated ways of evaluating spatial ability, the MRT [50,51]. In our study, we found no relationship between MRT scores and baseline performance. This suggests that AR proficiency may be influenced by more nuanced visual processing skills that are not captured by the MRT.

Höhler et al [49] and Martin-Gomez et al [52] have suggested that depth perception and stereoacuity affect individuals' ability to estimate distances of objects in AR. Given the importance of interacting with virtual elements in AR, estimating the depth and position of these objects may play a larger role than previously thought and could account for the visual processing skills that are not captured by the MRT.

The observed result that increased video game experience was correlated with increased accuracy in AR tasks may be explained by the beneficial effect of gaming on spatial cognition. Work by Bavelier and Green [53] indicates that specifically action video game play enhances spatial cognition; however, other literature has indicated that these cognitive improvements are not unique to only action games [54]. This indicates that the relationship between video game experience and accuracy in AR may be due to the cognitive benefits of extensively playing video games, regardless of genre.

The literature indicates that video game experience may be a positive predictor of performance in surgical tasks with respect to errors and time [55-57]. Our findings suggest that this relationship may extend to AR-based applications with respect to errors; however, more research is needed to evaluate its effect on performance time.

Learning How to Use AR

One of the reasons AR can be challenging for novice users is the variability in the learning process [58]. However, as with other skills, increased AR exposure is associated with improved performance. Our unadjusted analyses demonstrated a rapid learning effect, with the most pronounced gains occurring during early task exposure. This suggests that novice AR users may rapidly familiarize themselves with the AR environment. However, covariate-adjusted models did not indicate that these improvements differed significantly based on user characteristics.

Users who initially performed tasks more slowly demonstrated the greatest improvement. Tasks requiring less depth perception showed more rapid learning, while those emphasizing higher depth perception and precision, such as the virtual landmark placement (task 5), improved more gradually. Notably, physical landmark placement (task 6) did not show a learning effect, possibly because participants could rely on tactile feedback from touching the skull with the stylus.

Given that covariate-adjusted models showed no significant influence of MRT or video game experience on learning, these findings suggest that inherent user characteristics, such as spatial ability, do not impact early AR learning capacity in novice users. However, given our modest sample size (N=21), the nonsignificant covariate terms and interactions should be interpreted cautiously.

Importance of Depth Perception With AR

There is a possibility that depth perception and stereoacuity play a larger role in novice AR performance due to inherent technological limitations of the HMD. The AR device used in this study, the Microsoft HoloLens 2, uses a traditional fixed plane optical display. Research with the HoloLens has supported that visual rendering factors such as shadows [59] and lighting conditions [60] may impact the depth perception of users. Additionally, binocular disparity and the occlusion of an object are other important cues for depth perception [61].

XSL•FO
RenderX

If a user attempts to interact with a virtual object in AR, they may experience an occlusion error, in which the object appears translucent despite the user's hand not being at the appropriate distance to interact with it. Uehira and Suzuki [61] identified that this depth perception error was highly varied between individuals, particularly at short distances where the difference in binocular disparity is especially pronounced. Most of the tasks in our study were performed at short distances, mimicking clinical interactions with AR. Our study did not quantify the distances of the virtual objects, nor did we measure how many times users missed targets due to misjudgment of depth. Given that interaction with virtual objects is a fundamental component of AR use, it is likely that individuals who have stronger depth perception abilities may outperform those with weaker depth perception [49].

Concurrently, these findings provide new evidence that traditional measures of visuospatial ability do not reliably predict novice AR performance, while unmeasured factors, including depth perception, may contribute more than previously thought. The early performance gains observed in unadjusted analyses suggest that novice AR proficiency can be rapidly developed, a result supported by short-format training within urology [14]. Importantly, these learning effects, combined with the scarcity of existing implementation models [30], suggest that successful AR adoption may benefit from short, targeted training programs that guide all novice users to a competency threshold rather than prioritizing users based on traits such as visuospatial ability or video game experience. Furthermore, this emphasizes that predictive measures of novice performance should be interpreted in the context of this rapid rate of improvement.

Limitations

This study has some notable limitations. The potential sampling bias introduced by the inclusion of only undergraduate and graduate students may limit the generalizability of the findings to broader populations, such as resident and attending physicians who represent actual AR users in health care settings. Given that our sample size was 21 nonsurgeon participants, we believe that further research evaluating the learning curve within intraoperative environments is necessary before concluding that task-specific guides will reduce the learning curve.

Additionally, as the sample was modest (N=21) and the covariate-adjusted model included multiple predictors (gaming hours and MRT), this study may be underpowered to detect small-to-moderate covariate effects and trial-by-covariate interactions. Accordingly, nonsignificant covariate terms (eg, gaming hours P=.80; MRT P=.17) and interaction terms (trial×gaming hours P=.62; trial×MRT P=.34, Greenhouse-Geisser corrected) should be interpreted cautiously. Additionally, video game experience was self-reported and categorized based on hours per week. Our study found it challenging to obtain the genre of video games played and therefore did not analyze whether different categories of video games influenced performance with AR. The cognitive demand effects and the type of video games were not collected; however, these factors may have an influence on how participants perform in the tasks we evaluated in this study. Furthermore, technical limitations of the Microsoft HoloLens 2 cannot be discounted, such as ambient lighting conditions in the room during experimentation, which may have affected hologram visual quality. Finally, some outcomes, such as orbit tracing quality, were evaluated qualitatively and may be subject to observer bias.

Conclusions

As AR technology continues to grow in adoption across different industries, there is an increased need to identify the factors that contribute to effective AR use. Our research found that extensive video game experience was correlated with decreased error frequency, while neither visuospatial ability nor video game experience predicted novice user performance time. We believe that future research should focus on how depth perception, stereoacuity, and learning play a role in novice user performance, while also evaluating the learning curve of surgeons in intraoperative environments. This area of research holds important promise and may shape how industry professionals and product developers design and train future users to adopt AR systems more effectively.

Acknowledgments

The authors attest that generative artificial intelligence was not used in the generation of any part of this manuscript.

Funding

This study is based on work supported by the National Science Foundation Graduate Research Fellowship (grants DGE2140739 and 2139321).

Data Availability

The data presented in this study are available from the corresponding author on reasonable request and with institutional approval.

Conflicts of Interest

EGA, JTB, and GH are shareholders of SymphonyMR Inc.

References

XSL•FO
RenderX

1. Villagran-Vizcarra DC, Luviano-Cruz D, Pérez-Domínguez LA, Méndez-González LC, Garcia-Luna F. Applications analyses, challenges and development of augmented reality in education, industry, marketing, medicine, and entertainment. Appl Sci 2023 Feb;13(5):2766. [doi: 10.3390/app13052766]
2. Syed TA, Siddiqui MS, Abdullah HB, et al. In-depth review of augmented reality: tracking technologies, development tools, AR displays, collaborative AR, and security concerns. Sensors (Basel) 2022 Dec 23;23(1):146. [doi: 10.3390/s23010146] [Medline: 36616745]
3. Vucicevic RS, Castonguay JB, Treviño N, et al. Surgeon perspectives on a virtual reality platform for preoperative planning in complex bone sarcomas. J Orthop 2024 Oct;62:43-48. [doi: 10.1016/j.jor.2024.10.012] [Medline: 39507951]
4. Porpiglia F, Checcucci E, Volpi G, et al. Artificial intelligence 3D augmented reality-guided robotic prostatectomy versus cognitive MRI intervention: results of the prospective randomized RIDERS trial. Eur Urol 2026 Mar;89(3):233-243. [doi: 10.1016/j.eururo.2025.09.4172] [Medline: 41087293]
5. Nadeem-Tariq A, Kazemeini S, Kaur P, et al. Augmented reality in spine surgery: a narrative review of clinical accuracy, workflow efficiency, and barriers to adoption. Cureus 2025 Jun;17(6):e86803. [doi: 10.7759/cureus.86803] [Medline: 40718258]
6. West K, Al-Nimer S, Goel VR, et al. Three-dimensional holographic guidance, navigation, and control (3D-GNC) for endograft positioning in porcine aorta: feasibility comparison with 2-dimensional x-ray fluoroscopy. J Endovasc Ther 2021 Oct;28(5):796-803. [doi: 10.1177/15266028211025026] [Medline: 34142900]
7. Urlings J, Sezer S, Ter Laan M, et al. The role and effectiveness of augmented reality in patient education: a systematic review of the literature. Patient Educ Couns 2022 Jul;105(7):1917-1927. [doi: 10.1016/j.pec.2022.03.005] [Medline: 35341611]
8. Wehrkamp K, Miksch RC, Polzer H, et al. The impact of virtual-, augmented- and mixed reality during preoperative informed consent: a systematic review of the literature. J Med Syst 2025 Jun 24;49(1):89. [doi: 10.1007/s10916-025-02217-9] [Medline: 40555846]
9. Bray L, Sharpe A, Gichuru P, Fortune PM, Blake L, Appleton V. The acceptability and impact of the Xploro digital therapeutic platform to inform and prepare children for planned procedures in a hospital: before and after evaluation study. J Med Internet Res 2020 Aug 11;22(8):e17367. [doi: 10.2196/17367] [Medline: 32780025]
10. Mondal H, Mondal S. Adopting augmented reality and virtual reality in medical education in resource-limited settings: constraints and the way forward. Adv Physiol Educ 2025 Jun 1;49(2):503-507. [doi: 10.1152/advan.00027.2025] [Medline: 40136005]
11. Liu X, Xiao W, Yang Y, Yan Y, Liang F. Augmented reality technology shortens aneurysm surgery learning curve for residents. Comput Assist Surg (Abingdon) 2024 Dec;29(1):2311940. [doi: 10.1080/24699322.2024.2311940] [Medline: 38315080]
12. Barteit S, Lanfermann L, Bärnighausen T, Neuhann F, Beiersmann C. Augmented, mixed, and virtual reality-based head-mounted devices for medical education: systematic review. JMIR Serious Games 2021 Jul 8;9(3):e29080. [doi: 10.2196/29080] [Medline: 34255668]
13. Suresh D, Aydin A, James S, Ahmed K, Dasgupta P. The role of augmented reality in surgical training: a systematic review. Surg Innov 2023 Jun;30(3):366-382. [doi: 10.1177/15533506221140506] [Medline: 36412148]
14. Bhardwaj M, Singhal A, Bhardwaj G, Sur H. Effectiveness of simulation-based training in urology: a systematic review of educational outcomes and clinical skill transfer. Cureus 2025 Dec;17(12):e98641. [doi: 10.7759/cureus.98641] [Medline: 41409365]
15. Jeffri NF, Awang Rambli DR. A review of augmented reality systems and their effects on mental workload and task performance. Heliyon 2021 Mar 8;7(3):e06277. [doi: 10.1016/j.heliyon.2021.e06277] [Medline: 33748449]
16. Alessa FM, Alhaag MH, Al-Harkan IM, Ramadan MZ, Alqahtani FM. A neurophysiological evaluation of cognitive load during augmented reality interactions in various industrial maintenance and assembly tasks. Sensors (Basel) 2023 Sep 6;23(18):7698. [doi: 10.3390/s23187698] [Medline: 37765755]
17. Mayer R, Makransky G, Parong J. The promise and pitfalls of learning in immersive virtual reality. Int J Hum Comput Interact 2023;39(11):2229-2238. [doi: 10.1080/10447318.2022.2108563]
18. Şimşek EE. The effect of augmented reality storybooks on the story comprehension and retelling of preschool children. Front Psychol 2024;15:1459264. [doi: 10.3389/fpsyg.2024.1459264] [Medline: 39439758]
19. Amores-Valencia A, Burgos D, Branch-Bedoya JW. Influence of motivation and academic performance in the use of augmented reality in education. A systematic review. Front Psychol 2022;13:1011409. [doi: 10.3389/fpsyg.2022.1011409] [Medline: 36304863]
20. Bölek KA, De Jong G, Henssen D. The effectiveness of the use of augmented reality in anatomy education: a systematic review and meta-analysis. Sci Rep 2021 Jul 27;11(1):15292. [doi: 10.1038/s41598-021-94721-4] [Medline: 34315955]
21. Bork F, Stratmann L, Enssle S, et al. The benefits of an augmented reality magic mirror system for integrated radiology teaching in gross anatomy. Anat Sci Educ 2019 Nov;12(6):585-598. [doi: 10.1002/ase.1864] [Medline: 30697948]
22. Bogomolova K, van der Ham IJ, Dankbaar ME, et al. The effect of stereoscopic augmented reality visualization on learning anatomy and the modifying effect of visual-spatial abilities: a double-center randomized controlled trial. Anat Sci Educ 2020 Sep;13(5):558-567. [doi: 10.1002/ase.1941] [Medline: 31887792]

XSL•FO
RenderX

23. 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]
24. Roca-González C, Martin Gutierrez J, GarcÍa-Dominguez M, Mato Carrodeguas MC. Virtual technologies to develop visual-spatial ability in engineering students. Eurasia J Math Sci Technol Educ 2017;13(2):441-468. [doi: 10.12973/eurasia.2017.00625a]
25. Contero M, Company P, Naya F, Saorin JL. Learning support tools for developing spatial abilities in engineering design. Int J Eng Educ 2006;22(3):470-477 [FREE Full text]
26. Yousuf MS, Alsawareah A, Alhroub A, et al. Investigation of factors that influence the relationship between mental rotation ability and anatomy learning. Morphologie 2024 Mar;108(360):100728. [doi: 10.1016/j.morpho.2023.100728] [Medline: 37988905]
27. Mulder TA, van de Velde T, Dokter E, et al. Unravelling the skillset of point-of-care ultrasound: a systematic review. Ultrasound J 2023 Apr 19;15(1):19. [doi: 10.1186/s13089-023-00319-4] [Medline: 37074526]
28. Pedersen H, Ståhl D, Ekelund M, et al. Visuospatial ability is associated to 2D laparoscopic simulator performance amongst surgical residents. Surg Open Sci 2023 Jan;11:56-61. [doi: 10.1016/j.sopen.2022.11.004] [Medline: 36545373]
29. Rogister F, Pottier L, El Haddadi I, et al. Use of Vandenberg and Kuse Mental Rotation Test to predict practical performance of sinus endoscopy. Ear Nose Throat J 2022 Feb;101(2_suppl):24S-30S. [doi: 10.1177/01455613211000599] [Medline: 33734883]
30. Iqbal AI, Aamir A, Hammad A, et al. Immersive technologies in healthcare: an in-depth exploration of virtual reality and augmented reality in enhancing patient care, medical education, and training paradigms. J Prim Care Community Health 2024;15:21501319241293311. [doi: 10.1177/21501319241293311] [Medline: 39439304]
31. Li L, Chen R, Chen J. Playing action video games improves visuomotor control. Psychol Sci 2016 Aug;27(8):1092-1108. [doi: 10.1177/0956797616650300] [Medline: 27485132]
32. van Dongen KW, Verleisdonk EJ, Schijven MP, Broeders IA. Will the Playstation generation become better endoscopic surgeons? Surg Endosc 2011 Jul;25(7):2275-2280. [doi: 10.1007/s00464-010-1548-2] [Medline: 21416186]
33. Sammut M, Sammut M, Andrejevic P. The benefits of being a video gamer in laparoscopic surgery. Int J Surg 2017 Sep;45:42-46. [doi: 10.1016/j.ijsu.2017.07.072] [Medline: 28733118]
34. Gugura R, Fischer P, Tan ău M, Tefas C. Just five more minutes, mom: why video games could make you a better endoscopist. Surg Endosc 2023 Sep;37(9):6901-6907. [doi: 10.1007/s00464-023-10167-x] [Medline: 37316677]
35. Katz D, Hyers B, Patten E, Sarte D, Loo M, Burnett GW. Relationship between demographic and social variables and performance in virtual reality among healthcare personnel: an observational study. BMC Med Educ 2024 Mar 4;24(1):227. [doi: 10.1186/s12909-024-05180-0] [Medline: 38439056]
36. Journal Article Reporting Standards (JARS). American Psychological Association. URL: https://apastyle.apa.org/jars [accessed 2026-03-07]
37. 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] [Medline: 5540314]
38. Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. Percept Mot Skills 1978 Oct;47(2):599-604. [doi: 10.2466/pms.1978.47.2.599] [Medline: 724398]
39. Caissie AF, Vigneau F, Bors DA. What does the Mental Rotation Test measure? An analysis of item difficulty and item characteristics. Open Psychol J 2009 Dec 11;2:94-102. [doi: 10.2174/1874350100902010094]
40. Hegarty M. Ability and sex differences in spatial thinking: what does the mental rotation test really measure? Psychon Bull Rev 2018 Jun;25(3):1212-1219. [doi: 10.3758/s13423-017-1347-z] [Medline: 28808983]
41. Rajeb M, Krist AT, Shi Q, Oyeniran DO, Wind SA, Lakin JM. Mental rotation performance: contribution of item features to difficulties and functional adaptation. J Intell 2024 Dec 30;13(1):2. [doi: 10.3390/jintelligence13010002] [Medline: 39852411]
42. Uttal DH, Meadow NG, Tipton E, et al. The malleability of spatial skills: a meta-analysis of training studies. Psychol Bull 2013 Mar;139(2):352-402. [doi: 10.1037/a0028446] [Medline: 22663761]
43. Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. Human Mental Workload: North-Holland Publishing Company; 1988:139-183.
44. Horn BK. Closed-form solution of absolute orientation using unit quaternions. J Opt Soc Am A 1987;4(4):629-642. [doi: 10.1364/JOSAA.4.000629]
45. Reason J. Human Error: Cambridge University Press; 1990.
46. Norman DA. The Design of Everyday Things: MIT Press; 2013.
47. Tandon S, Abdul-Rahman A, Borgo R. Measuring effects of spatial visualization and domain on visualization task performance: a comparative study. IEEE Trans Visual Comput Graphics 2023;29(1):668-678. [doi: 10.1109/TVCG.2022.3209491] [Medline: 36166560]
48. Kourtesis P, MacPherson SE. An ecologically valid examination of event-based and time-based prospective memory using immersive virtual reality: the influence of attention, memory, and executive function processes on real-world prospective memory. arXiv. Preprint posted online on Feb 23, 2021. [doi: 10.48550/arXiv.2102.11652]

XSL•FO
RenderX

49. Höhler C, Rasamoel ND, Rohrbach N, et al. The impact of visuospatial perception on distance judgment and depth perception in an augmented reality environment in patients after stroke: an exploratory study. J Neuroeng Rehabil 2021 Aug 21;18(1):127. [doi: 10.1186/s12984-021-00920-5] [Medline: 34419086]
50. Aguilar Ramirez DE, Blinch J, Takeda K, Copeland JL, Gonzalez CL. Differential effects of aging on spatial abilities. Exp Brain Res 2022 May;240(5):1579-1588. [doi: 10.1007/s00221-022-06363-1] [Medline: 35428943]
51. Tang Z, Liu X, Huo H, et al. Eye movement characteristics in a mental rotation task presented in virtual reality. Front Neurosci 2023;17:1143006. [doi: 10.3389/fnins.2023.1143006] [Medline: 37051147]
52. Martin-Gomez A, Weiss J, Keller A, Eck U, Roth D, Navab N. The impact of focus and context visualization techniques on depth perception in optical see-through head-mounted displays. IEEE Trans Vis Comput Graph 2022 Dec;28(12):4156-4171. [doi: 10.1109/TVCG.2021.3079849] [Medline: 33979287]
53. Bavelier D, Green CS. Enhancing attentional control: lessons from action video games. Neuron 2019 Oct 9;104(1):147-163. [doi: 10.1016/j.neuron.2019.09.031] [Medline: 31600511]
54. Oei AC, Patterson MD. Enhancing cognition with video games: a multiple game training study. PLoS One 2013;8(3):e58546. [doi: 10.1371/journal.pone.0058546] [Medline: 23516504]
55. Tsai CL, Heinrichs WL. Acquisition of eye-hand coordination skills for videoendoscopic surgery. J Am Assoc Gynecol Laparosc 1994 Aug;1(4, Part 2):S37. [doi: 10.1016/s1074-3804(05)80989-2] [Medline: 9073766]
56. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Impact of hand dominance, gender, and experience with computer games on performance in virtual reality laparoscopy. Surg Endosc 2003 Jul;17(7):1082-1085. [doi: 10.1007/s00464-002-9176-0] [Medline: 12728373]
57. Rosser JC Jr, Lynch PJ, Cuddihy L, Gentile DA, Klonsky J, Merrell R. The impact of video games on training surgeons in the 21st century. Arch Surg 2007 Feb;142(2):181-186. [doi: 10.1001/archsurg.142.2.181] [Medline: 17309970]
58. Khurana A, Glueck M, Chilana PK. Do i just tap my headset?: how novice users discover gestural interactions with consumer augmented reality applications. In: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies: Association for Computing Machinery; 2024, Vol. 7:1-28. [doi: 10.1145/3631451]
59. Adams H, Stefanucci JK, Creem-Regehr S, Bodenheimer B. Depth perception in augmented reality: the effects of display, shadow, and position. Presented at: 2022 IEEE on Conference Virtual Reality and 3D User Interfaces (VR); Mar 12-16, 2022. [doi: 10.1109/VR51125.2022.00101]
60. Ashtiani O, Guo HJ, Prabhakaran B. Impact of motion cues, color, and luminance on depth perception in optical see-through AR displays. Front Virtual Real 2023 Dec 6;4. [doi: 10.3389/frvir.2023.1243956]
61. Uehira K, Suzuki M. Depth perception for virtual object displayed in optical see-through HMD. Presented at: ACHI 2018: Eleventh International Conference on Advances in Computer-Human Interactions; Mar 25-29, 2018 URL: https://personales.upv.es/thinkmind/dl/conferences/achi/achi_2018/achi_2018_11_20_20009.pdf [accessed 2026-03-26]

Abbreviations

AR: augmented reality
HMD: head-mounted display
MRT: mental rotation task
NASA-TLX: National Aeronautics and Space Administration Task Load Index
STEM: science, technology, engineering, and math
VR: virtual reality

Edited by S Brini; submitted 24.Jul.2025; peer-reviewed by T Bullock, Y Yang; revised version received 12.Mar.2026; accepted 12.Mar.2026; published 10.Apr.2026.

Please cite as:
Vellore AJ, Bhatia S, Kann MR, Kass NM, Shanahan RM, Jardini J, Miner J, Daulat SR, Hurt G, Basdeo R, Don N, Biehl JT, Andrews EG
Predictive Factors of Augmented Reality–Based Clinical Task Performance Among Novice Users: Cross-Sectional Quantitative Study
JMIR XR Spatial Comput 2026;3:e81236
URL: https://xr.jmir.org/2026/1/e81236
doi:10.2196/81236

© Amogh J Vellore, Shovan Bhatia, Michael R Kann, Nicolás M Kass, Regan M Shanahan, Jacquelyn Jardini, Jayne Miner, Sohail R Daulat, Griffin Hurt, Rishi Basdeo, Nicole Don, Jacob T Biehl, Edward G Andrews. Originally published in JMIR XR and Spatial Computing (https://xr.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

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Experience Comparison of a Traditional Mannequin-Based and an Augmented Reality Lumbar Puncture Simulator in Geriatric Medicine Training: Questionnaire Study

Emma Combret[1,2], MSc; Alexis Bourgeais[2,3], MD; Frédéric Noublanche[1,2,3,4], PhD; Cédric Annweiler[1,2,3,4,5], MD, PhD

[1]Laboratoire de Psychologie des Pays de la Loire (UR4638), Angers, Pays de la Loire, France

[2]Department of Geriatric Medicine and Memory Clinic, Research Center on Autonomy and Longevity, Centre Hospitalier Universitaire d'Angers, 4 rue Larrey, Angers, Pays de la Loire, France

[3]Health Faculty, Université d'Angers, Angers, Pays de la Loire, France

[4]Gérontopôle Autonomie Longévité des Pays de la Loire, Nantes, Pays de la Loire, France

[5]Department of Medical Biophysics, Schulich School of Medicine and Dentistry, Robarts Research Institute, Western University, London, ON, Canada

Corresponding Author:
Emma Combret, MSc
Laboratoire de Psychologie des Pays de la Loire (UR4638), Angers, Pays de la Loire, France

Abstract

Background: Lumbar puncture (LP) is a diagnostic procedure in geriatric medicine but may be technically challenging in older adults with spinal deformities, osteoarthritis, frailty, or cognitive impairment. Simulation-based training can provide repeated practice without patient risk. Augmented reality (AR) systems may improve anatomical visualization and sensory feedback, but their added value compared with mannequin-based simulators and whether perceptions differ according to clinical experience remain uncertain.

Objective: This study aimed to compare the user experience (UX) of a traditional mannequin-based LP simulator (M43E; T-S) and an augmented reality simulator with haptic feedback (Sim&Care 2; AR-S) across levels of expertise.

Methods: We conducted a single-center, randomized-order, within-participant comparative UX study at Angers University Hospital, France, in June 2025. A total of 30 participants (n=7, 23.3% sixth-year medical students; n=16, 53.3% postgraduate medical students; and n=7, 23.3% graduated physicians) used both simulators for 15 minutes each, with the exposure order randomized. After each session, participants completed an exploratory 5-point Likert questionnaire assessing face validity (FV), usability, perceived educational value (PEV), and overall satisfaction. After both sessions, they reported simulator preference, reasons for preference, and intention to reuse each simulator; open-ended comments were collected. Dimension-level paired comparisons were performed using the Wilcoxon signed-rank test, expertise-group comparisons used the Kruskal-Wallis test, and categorical outcomes were analyzed using the chi-square test. Item-level analyses were considered exploratory.

Results: The AR-S achieved higher composite scores than the T-S for FV (median 3.8 [IQR 3.6-4.2] vs 3.4 [IQR 3.2-3.8]; *P*=.02) and PEV (median 4 [IQR 3.8-4.6] vs 3.6 [IQR 3.4-4.2]; *P*=.02). No significant differences were observed for usability (median 4.3 [IQR 4-4.95] vs 4.6 [IQR 4.4-4.8]; *P*=.11) or overall satisfaction (median 4 [IQR 4-5] vs 4 [IQR 3-4]; *P*=.16). Exploratory item-level analyses indicated higher ratings for the AR-S for anatomical realism (*P*=.03), tactile sensations (*P*=.003), and perceived acquisition of new knowledge (*P*=.005), whereas the T-S was rated higher for ease of use (*P*=.01), clarity of instructions (*P*=.01), and ease of operation during the procedure (*P*=.003). Overall preference was balanced: 53.3% (16/30) of participants preferred the AR-S, and 46.7% (14/30) of participants preferred the T-S. Preference was not significantly associated with expertise group (*P*=.09), although graduated physicians descriptively favored the AR-S and sixth-year students favored the T-S. In total, 76.7% (23/30) of participants reported willingness to reuse both simulators.

Conclusions: AR and traditional mannequin-based LP simulation appear to offer distinct and complementary educational advantages rather than one modality being uniformly superior. The AR-S was valued for anatomical and sensory realism and PEV, whereas the T-S retained advantages related to procedural familiarity and ease of use with conventional equipment. These findings support tailoring simulator selection to learner expertise and educational objectives. Larger studies incorporating objective procedural performance are needed to determine how the modalities should be integrated into LP curricula.

(JMIR XR Spatial Comput 2026;3:e98090) doi:10.2196/98090

KEYWORDS

spinal puncture; simulation training; patient simulation; augmented reality; user experience; medical education; geriatrics

XSL•FO
RenderX

Introduction

Lumbar puncture (LP) is a critical procedure in geriatrics, especially for the diagnostic workup of neurodegenerative diseases, including Alzheimer disease and related disorders [1]. However, it is often feared by physicians and medical students due to its invasive nature and its undesirable effects, such as post-LP headaches, bruising, and infections [2]. In older adults, LP may be technically more challenging due to spinal deformities, osteoarthritis, frailty, or behavioral disturbances related to cognitive impairment [3].

Good knowledge of the procedure is the most effective way to avoid adverse effects [4]. Nevertheless, the limited opportunities for medical students to perform LP in clinical settings may contribute to insufficient procedural experience and reduced self-confidence [5]. Notably, research shows that medical students have a higher level of subjective and physiological stress than physicians, affecting their psychomotor capacity and working memory [6], and increasing the risk of post-LP headaches in patients [7].

In France, LP training during medical school remains highly variable depending on the hospital department where students do their internship, the number of students per internship, and the availability of interns and physicians to teach the procedure [8]. The use of simulators, which are becoming more popular, is a promising opportunity to improve traditional LP training. These simulators allow learners to reproduce the main steps of the procedure using the same material and to experience the same sensations, including anatomical landmark palpation, needle insertion, and cerebrospinal fluid collection [8].

Several studies have investigated the benefits of LP simulation training among medical students and residents. They have reported improvements in theoretical knowledge, confidence, and transfer of skills to clinical practice, including higher first-attempt success rates during subsequent LP procedures performed on patients [8-10] and fewer traumatic LPs [11]. However, the benefit of simulation-based training may vary depending on the comparison method and outcome assessed. In a study comparing problem-based learning and simulator-based learning, both approaches improved learner performance, and no significant difference was observed between groups immediately after training [12]. More recently, a randomized controlled trial evaluating virtual reality (VR) training compared with conventional training alone found no significant difference in overall LP performance between groups [13]. In addition, traditional physical simulators have several limitations, including material wear, limited anatomical variability, and restricted representation of tissue biomechanical properties [14]. These limitations may reduce their ability to reproduce the diversity of clinical situations encountered in practice, particularly in older patients, where osteoarthritis and spinal deformities may complicate LP performance [15].

To overcome these limitations, technological advances have led to the development of virtual and augmented reality (AR)–based LP simulators incorporating haptic feedback [15]. While VR-based systems have been more extensively investigated, AR-based LP simulators remain scarce. The Sim&Care Spine simulator (InSimo) is one example of a commercially available AR-based platform designed to provide anatomical visualization and haptic feedback during LP training [16]. This innovative method allows a better visualization of the internal and external anatomy of the patient and has a positive impact on visuospatial capacities, cognitive load, and procedure time [17]. Studies evaluating AR simulators found no significant improvement in LP success rates compared with traditional bedside teaching, but learners reported increased confidence and comfort. Students also reported high educational value, particularly appreciating the haptic feedback and AR visualization [16]. In a study evaluating a mixed reality simulator, participants reported high perceived realism and educational value, particularly regarding anatomical understanding and identification of landmarks. Training with the simulator was also associated with improvements in procedural parameters, including needle positioning, number of attempts, and procedure duration [18].

A comparative study evaluating mixed reality LP training versus traditional high-fidelity mannequin-based simulation showed comparable procedural performance between the 2 approaches. However, mixed reality training was well accepted by learners, with most students considering it to be as effective as traditional simulation and expressing interest in further use of this technology [19]. Although no significant differences were observed in objective outcomes, such as the number of punctures, time to successful puncture, or distance from the target, learners trained with AR reported greater confidence [20]. In addition, VR-based training increased learner engagement, satisfaction, and interest in clinical skills, suggesting that immersive technologies primarily enhance the educational experience rather than procedural performance [13]. AR may enhance student engagement by sustaining learners' interest and willingness to use this technology beyond the initial training session [21]. Overall, current evidence suggests that AR- and VR-based simulators do not consistently demonstrate superior procedural performance compared with traditional mannequin-based simulation. However, immersive technologies appear to provide specific advantages in terms of anatomical visualization, sensory feedback, learner confidence, engagement, and satisfaction. Although students generally perceive AR simulators as realistic, easy to use, and useful [6,18,22], overall preference remains divided [19]. This suggests that simulator preference may depend on multiple dimensions of user experience (UX), including anatomical and procedural realism, usability, and perceived educational value (PEV). Moreover, these perceptions may vary according to learners' expertise, as novices and experienced practitioners may value different aspects of simulation, such as procedural familiarity or anatomical understanding.

The objective of this study was therefore to evaluate the UX, including perceived realism, usability, and educational value, of an augmented reality simulator (AR-S; Sim&Care 2) compared with a traditional simulator (T-S; M43E) across different levels of expertise. We hypothesized that participants would prefer the AR simulator for anatomical and sensory realism, while usability and procedural familiarity might favor the traditional model.

XSL•FO
RenderX

Methods

Study Design

This study was a single-center, randomized-order, within-participant comparative UX study conducted in the Department of Geriatric Medicine of Angers University Hospital, Angers, France, in June 2025. A single-center design was chosen because the study was conducted within the Angers simulation center, where the T-S was routinely used for training and where the AR simulator was made available for a dedicated evaluation session by the manufacturer. Each participant evaluated both the traditional and AR simulators, with the order of simulator exposure randomized between participants to minimize potential order effects. UX was assessed immediately after each simulator session using a UX questionnaire, followed by a comparative questionnaire after completion of both sessions. This design allowed direct comparison of the 2 simulators while limiting interindividual variability. The study was reported in accordance with the RATE-XR reporting guideline for augmented and VR studies [23].

Recruitment

Thirty participants were included: 7 (23.3%) graduated physicians, 16 (53.3%) postgraduate medical students, and 7 (23.3%) sixth-year medical students. Participants were recruited through email invitations sent to medical students, residents, and physicians affiliated with Angers University Hospital. Recruitment was voluntary. Inclusion criteria were being a sixth-year medical student, postgraduate medical student, or a graduated physician affiliated with the Department of Geriatric Medicine of Angers University Hospital, involved in medical training or clinical practice, and agreeing to participate in the study. These inclusion criteria were chosen to include participants with different levels of experience with LP. Exclusion criteria included inability to provide informed consent or motor impairment preventing simulator use. The sample size was based on the availability of eligible participants during the study period. No formal sample size calculation was performed because this was an exploratory UX study.

Ethical Considerations

The study protocol was reviewed and approved by the Ethics Committee of Angers University Hospital (*Comité d'Éthique du Centre Hospitalier Universitaire d'Angers*), which issued a favorable opinion on June 10, 2025 (reference number 2025‑140). The study was conducted in accordance with the principles of the Declaration of Helsinki. Participants received oral information about the study, and their nonobjection to participation was obtained before the beginning of the experiment. The data were pseudonymized using a unique alphanumeric code assigned to each participant to ensure anonymity. No compensation was provided to participants.

Materials and Protocol

This study was conducted in collaboration with the Angers simulation center (AllSims) and InSimo. The 2 simulators were selected because they represent complementary approaches to LP training: a traditional mannequin-based simulator and an AR-based simulator. The M43E simulator (Kyoto Kagaku) was already available at the Angers simulation center, whereas the Sim&Care 2 simulator (InSimo) was provided by the manufacturer for a dedicated evaluation session.

Main Characteristics of the Implemented Simulators

The T-S (M43E simulator) is a high-fidelity static LP and epidural procedure simulator designed for procedural training. It consists of an anatomically accurate lumbar region model with a puncture block reproducing relevant structures, including lumbar vertebrae, spinous processes, transverse processes, epidural space, and dura mater. The simulator allows training in landmark identification, needle insertion, and cerebrospinal fluid collection using a real LP needle. It can be positioned in both upright and lateral decubitus configurations, and its translucent puncture block allows visualization of the needle trajectory during the procedure.

The AR-S (Sim&Care 2 simulator) combines a physical interface with virtual anatomical overlay and force-feedback technology delivered via a stylus device and a Microsoft HoloLens 2 headset AR display. The setup consisted of a PC running the simulation software, a Microsoft HoloLens 2 headset for the AR environment, a Touch haptic interface (3D Systems) for force feedback, and a mannequin serving as the physical support for the AR environment. The simulation software provided virtual anatomical visualization and force feedback during needle progression. Sim&Care 2 offers 4 patient cases: standard anatomy, scoliosis, osteoarthritis, and a pediatric case. For each case, users can select between 2 common LP positions: sitting or lateral decubitus. Various customization features enhance the simulation experience, including patient positioning, needle selection, and adjustable difficulty levels. Users can display or hide internal anatomical structures through a transparency mode and activate a visual guidance tool showing the needle trajectory to assist with accurate needle placement within the spinal canal.

Simulation Protocol

Each participant completed 2 simulation sessions, one with each simulator (Figure 1). Participants were randomized to the simulator exposure order (group A or B). Each session lasted approximately 15 minutes, including a brief familiarization period with the simulator followed by the LP procedure. Participants were allowed to perform several attempts during each session, and the same duration was allocated to both simulators. Only standardized instructions were provided: "Perform a diagnostic lumbar puncture as in clinical practice." As the objective was to evaluate perceived UX rather than procedural performance, the instructor did not provide corrective feedback or assess technical success during the sessions.

XSL•FO
RenderX

Figure 1. Simulators used during the study sessions. (A) M43E traditional simulator. (B) Sim&Care 2 augmented reality simulator with headset, haptic stylus, and physical interface.

For the AR-S, participants initially performed the procedure using the standard patient configuration. The pediatric patient case was not used in this study. For both simulators, the seated position was used initially during the simulation session, and participants could subsequently practice in the lateral decubitus position if desired.

Questionnaires

After each simulation session, participants completed a UX questionnaire assessing the simulator they had just used. The questionnaire included 4 dimensions: face validity (FV; 5 items), usability (5 items), PEV (5 items), and overall satisfaction (1 item). Each item was rated on a 5-point Likert scale.

After completing both simulation sessions, participants completed a second comparative questionnaire. This questionnaire assessed overall preference for one of the 2 simulators, the reasons for their preference (realism, ease of use, educational value, sensations, or other), and their intention to reuse each simulator (yes or no). Participants were also invited to provide open-ended comments. Twelve participants provided comments.

Questionnaire Development

In the absence of a validated LP-specific UX instrument, we developed an exploratory structured questionnaire based on evaluation domains commonly used in simulation-based medical education, including FV [24], usability [25], educational value [26], and satisfaction. No formal validation process or expert assessment was performed before the study due to the exploratory nature of this UX evaluation. Internal consistency was not formally assessed due to sample size limitations. This limitation has been acknowledged in the Discussion section.

FV was assessed using 5 items on a 5-point Likert scale:

- FV1: “The simulator’s anatomy seemed realistic to me.”
- FV2: “The tactile sensations were comparable to a real lumbar puncture.”
- FV3: “Patient positioning was similar to real practice.”
- FV4: “Identification of anatomical landmarks was credible.”
- FV5: “Performing the procedure felt natural.”

Usability was assessed using 5 items on a 5-point Likert scale:

- Usability 1: “The simulator was easy to use.”
- Usability 2: “The instructions for using the simulator were clear.”
- Usability 3: “The simulator was ergonomic.”
- Usability 4: “I easily understood how to operate the simulator during the procedure.”
- Usability 5: “The equipment functioned properly without difficulty.”

PEV was assessed using 5 items on a 5-point Likert scale:

- PEV1: “This simulator is useful for learning lumbar puncture.”
- PEV2: “This simulator could improve my real-life practice.”
- PEV3: “I learned something new with this simulator.”
- PEV4: “The simulator allows progress with training.”
- PEV5: “I would recommend this simulator to other learners.”

Overall satisfaction was assessed using 1 item on a 5-point Likert scale: “I am overall satisfied with this simulator.”

Data Analysis

Statistical analyses were performed using JASP software. Descriptive analyses were conducted for each questionnaire item, including the median and quartiles. Dimension-level analyses were performed as the primary analysis to align with the study hypothesis. Composite scores were calculated as the mean of the items included in each predefined UX dimension (FV, usability, and PEV). Intraindividual comparisons between T-S and AR-S were performed using the Wilcoxon signed-rank

XSL•FO
RenderX

test. Item-level analyses were conducted as exploratory analyses to identify specific differences between the simulators. Given the exploratory nature of these comparisons and the limited sample size, item-level results were interpreted descriptively.

To compare results between the 3 groups based on experience level, the Kruskal-Wallis test was used. Participants' preferences and their intention to reuse each simulator were analyzed using the Pearson chi-square test. Additionally, a qualitative analysis of the open-ended comments was conducted using a thematic approach.

Two participants submitted incomplete UX questionnaires. Because the missing responses concerned different items, the UX score analyses were based on 29 or 30 participants depending on the item. However, because their comparative questionnaires were fully completed, their responses were included in the analysis of simulator preference and reuse intention.

Results

Comparison of UX Dimensions Between Simulators

Results for the UX dimensions are presented in Table 1. Two participants had missing responses for some UX items. Because these missing data concerned different items, the analyses were based on 29 or 30 participants depending on the item. Dimension-level analyses were performed by comparing the composite scores of each predefined UX domain (FV, usability, PEV, and overall satisfaction). The AR-S showed higher scores for FV (z=−2.323; P=.02) and PEV (z=−2.289; P=.02) compared with the T-S. No significant differences were observed for the overall usability score (z=−1.598; P=.11). Finally, for the overall satisfaction item, statistical analysis showed no significant difference in the median overall satisfaction scores between the simulators (z=−1.349; P=.16).

Table . Comparison of user experience dimensions between simulators using the Wilcoxon signed-rank test.

Items	M43E, median (IQR)	Sim&Care 2, median (IQR)	*P* value	r_{pb} [a]
Face validity (n=30)	3.4 (3.2-3.8)	3.8 (3.6-4.2)	.02	0.213
Usability (n=29)	4.6 (4.4-4.8)	4.3 (4-4.95)	.11	0.217
Perceived educational value (n=29)	3.6 (3.4-4.2)	4.0 (3.8-4.6)	.02	0.213
Overall satisfaction (n=29)	4.0 (3.0-4.0)	4.0 (4.0-5.0)	.16	0.269

[a] r_{pb}: point-biserial correlation coefficient.

Item-Level Exploratory Analysis

Item-level analyses were conducted as exploratory analyses to identify which specific simulator characteristics contributed to the observed dimension-level differences. These results are presented in Table 2. For the FV items, participants rated the AR-S (Sim&Care 2 simulator) as more realistic on the item "The simulator's anatomy seemed realistic to me" (FV1), z=−2.155; P=.03, and on the item "The tactile sensations were comparable to a real lumbar puncture" (FV2), z=−2.866; P=.003. Concerning the usability items, the T-S (M43E simulator) obtained a significantly higher median on item U1 ("The simulator was easy to use;" z=2.456; P=.01) and on item U2 ("The instructions for using the simulator were clear;" z=2.446; P=.01). A highly significant difference was also observed for item U4 ("I easily understood how to operate the simulator during the procedure;" z=2.934; P=.003), in favor of T-S. For the PEV items, statistical analysis showed a significant difference for item PEV3 ("I learned something new with this simulator"), with a higher median for the AR-S (z=−2.808; P=.005).

XSL•FO
RenderX

Table . Comparison of user experience items between simulators using the Wilcoxon signed-rank test.

Items	M43E, median (IQR)	Sim&Care 2, median (IQR)	*P* value	r_{pb} [a]
FV[b]1 (n=30)	3 (3-4)	4 (3.25-5)	.03	0.244
FV2 (n=30)	3 (2-4)	4 (3.25-4)	.003	0.225
FV3 (n=30)	4 (3-4)	4 (3-4)	.10	0.328
FV4 (n=30)	4 (3-4)	4 (3-4)	.42	0.269
FV5 (n=30)	4 (3-4.75)	4 (3-4)	.65	0.244
Usability 1 (n=30)	5 (4-5)	4 (3.25-5)	.01	0.277
Usability 2 (n=30)	5 (5-5)	5 (4-5)	.01	0.342
Usability 3 (n=30)	4 (4-5)	4 (4-5)	.83	0.25
Usability 4 (n=29)	5 (5-5)	5 (4-5)	.003	0.328
Usability 5 (n=29)	5 (3-5)	5 (4-5)	.15	0.277
PEV[c]1 (n=29)	4 (4-5)	4 (4-5)	.92	0.269
PEV2 (n=29)	3.1 (3-5)	4 (4-5)	.47	0.256
PEV3 (n=29)	2.5 (2-3)	4 (3-4)	.005	0.239
PEV4 (n=29)	4 (3-5)	4 (4-5)	.45	0.256
PEV5 (n=29)	4 (4-5)	4 (4-5)	.45	0.269

[a]r_{pb}: point-biserial correlation coefficient.

[b]FV: face validity.

[c]PEV: perceived educational value.

Group Effect

Results for group differences are presented in Table 3. The “group” factor showed a significant effect on one realism item for the AR-S, FV4 (“Identification of anatomical landmarks was credible;” Kruskal-Wallis test, *P*=.04), with a lower median among graduated physicians. For the T-S, the “group” factor showed a significant effect on the realism item FV3 (“Patient positioning was similar to real practice;” Kruskal-Wallis test, *P*=.048), with a higher median among medical students.

XSL•FO
RenderX

Table . Effect of participant group on user experience questionnaire items for the M43E and Sim&Care 2 simulators using Kruskal-Wallis test.

Items	M43E, *P* value	Sim&Care 2, *P* value
FV[a]1	.12	.95
FV2	.25	.43
FV3	.048	.06
FV4	.20	.04
FV5	.65	.52
Usability 1	.92	.72
Usability 2	.14	.83
Usability 3	.05	.25
Usability 4	.19	.81
Usability 5	.89	.22
PEV[b]1	.99	.42
PEV2	.11	.10
PEV3	.93	.68
PEV4	.37	.15
PEV5	.29	.38
Overall satisfaction	.22	.37

[a]FV: face validity.

[b]PEV: perceived educational value.

Participants' Preference and Intention to Reuse

Among the 30 participants, 16 (53.3%) reported a preference for the AR-S, while 14 (46.7%) preferred the T-S. Among graduated physicians, 85.7% (6/7) preferred the AR-S. Among sixth-year medical students, 71.4% (5/7) preferred the T-S. For postgraduate medical students, the 2 simulators were equally preferred (8/16, 50% each). However, simulator preference was not significantly associated with participant group (χ^2_2=4.8; *P*=.09).

The reasons given in favor of the AR-S mainly concerned "better sensations" (11/30, 36.7%) and greater educational value (10/30, 33.3%). For the T-S, the most frequently cited arguments were greater realism (6/19, 31.6%) and greater ease of use (5/19, 26.3%).

Regarding the intention to reuse, 76.7% (n=23) of participants were willing to use both simulators again. A total of 10% (n=3) were willing to reuse only the AR-S, and the same proportion were willing to reuse only the T-S. Finally, 3.3% (n=1) did not intend to use an LP simulator again. There was no association between the intention to reuse the simulators and participant group (χ^2_6=8.0; *P*=.24).

Discussion

Principal Findings

The aim of this study was to compare the UX of sixth-year second cycle medical students, postgraduate medical students, and graduated physicians using 2 LP simulators, a traditional one (M43E simulator) and an AR one (Sim&Care 2 simulator). Our hypothesis was that participants would prefer the AR model due to its greater realism and perceived pedagogical value. However, our hypothesis was also that these results could be influenced by the ease of use of the traditional model and the realism of the procedure, particularly due to the use of an actual LP needle. Overall, our findings partially confirmed this hypothesis. Dimension-level analyses showed that the AR simulator was associated with higher perceived FV and PEV compared with the T-S. However, no significant differences were observed for usability or overall satisfaction. To further understand the differences observed between simulators, complementary item-level analyses and qualitative data were examined to identify specific characteristics that may explain participants' perceptions.

Regarding anatomical and technical realism, participants in the study reported higher scores for the AR-S. It was perceived as having greater anatomical realism and providing sensations comparable to those experienced in real life. In the open-ended comments, 2 participants noted that the AR-S offered a more pleasant and realistic sensation during ligament passage and needle insertion. However, statistical analyses did not show a significant difference between the simulators for item FV5 ("Performing the procedure felt natural"). To explain this, participants reported that the stylus of the AR-S did not provide the same handling sensation as a real LP needle, mainly because of its larger diameter than that of the needle used with the T-S. Among groups, graduated physicians gave lower scores to the AR-S on the item "Identification of anatomical landmarks was credible," which may be explained by their greater expertise. Conversely, sixth-year second cycle medical students gave higher scores for both simulators on FV3 ("Patient positioning

XSL•FO
RenderX

was similar to real practice”), which may be due to their limited clinical experience.

Regarding usability, no significant difference was observed between the 2 simulators at the overall dimension level. However, item-level analyses and participants’ comments suggested some practical advantages for the T-S. This may be partly explained by participants’ prior exposure to the T-S during their medical training, which may have made them more familiar with its use. Participants who preferred the T-S cited ease of use and realism as reasons, as it is simpler to set up and the procedure is closer to real-life practice. However, 2 residents noted difficulties in locating anatomical landmarks (especially for osteoarthritis simulation) and noted mannequin wear at landmark areas—issues inherent to this type of simulator compared with AR-based simulators.

Regarding PEV, participants in the study reported higher scores for the AR-S. However, item-level analysis showed a higher score for the AR-S on PEV3 (“I learned something new with this simulator”). In the open-ended comments, 1 resident described the AR-S as an “excellent learning tool,” that provided a better 3D representation of anatomy than the T-S.

However, preferences were split (16/30, 53.3% preferred the AR-S and n=14, 46.7% preferred the T-S). Descriptively, preferences differed across participant groups: graduated physicians mostly preferred the AR model, and sixth-year second cycle medical students preferred the T-S. These results may be explained by experience: medical students, having rarely or never performed an LP, may value a simulator that is easy to use and provides realistic procedural execution (including needle handling) during their initial LP training. Graduated physicians, with expertise in the procedure, preferred the AR-S, which may have allowed them to better appreciate its sensations and anatomical realism. Regarding the intention to reuse the simulators, results were similar for both, showing no marked preference.

These results are consistent with and complementary to the existing literature comparing traditional and immersive simulation for LP training. While several studies have shown that AR, mixed reality, or VR simulators improve learner confidence, engagement, satisfaction, and anatomical understanding, they generally report procedural performance comparable to that obtained with conventional mannequin-based simulation rather than clear superiority of immersive technologies [13,16,18,19]. Our results complement these findings by showing that these technologies are not necessarily perceived as superior across all dimensions of the user experience. Instead, participants appeared to distinguish between different aspects of realism and educational value. The AR-S was particularly valued for anatomical visualization, sensory realism, and understanding of anatomical landmarks, whereas the T-S remained preferred for procedural realism and ease of use. These findings may explain why previous studies reported high satisfaction with immersive simulators despite the absence of consistent improvements in procedural performance.

Similar findings have been reported in other fields of medical simulation, where students preferred the T-S for learning clinical procedures, despite considering virtual simulators more enjoyable and immersive [27].

These observations suggest that AR-S and T-S should be considered complementary rather than competing educational tools. Traditional mannequin-based simulation may remain particularly appropriate for novice learners, who need to become familiar with real equipment and procedural workflow. In contrast, AR simulation may provide additional value for more advanced learners by enhancing anatomical understanding, spatial orientation, and sensory realism, especially in anatomically complex situations such as those frequently encountered in older patients. In this way, complementary use of both approaches may be considered to combine traditional mannequin simulations, which allow the use of real medical equipment particularly for novice medical students, with AR simulations, which offer dynamic scenarios and more realistic sensations for postgraduate medical students or graduated physicians who already have substantial experience with LPs.

This study has some limitations. First, the sample size limited the statistical power of subgroup analyses and the generalizability of the findings across different levels of expertise [28]. Second, some participants, particularly residents and graduated physicians, had previous experience with the T-S, which may have influenced their perception of usability and realism in favor of the T-S. Similarly, previous LP experience was not systematically collected. This may have affected participants’ perceptions of realism and educational value, particularly among medical students with limited exposure to the procedure. Third, some dimensions of realism were difficult to capture using Likert-scale evaluations. While participants perceived the AR-S as more realistic in terms of anatomical visualization and sensory feedback, some considered the procedural performance less realistic due to the use of a stylus rather than conventional LP equipment. This highlights the multidimensional nature of realism in simulation-based education. In addition, PEV may have varied according to participants’ level of expertise. Novice learners may value simulators for acquiring basic procedural knowledge, whereas more experienced participants may assess educational value according to more advanced objectives, such as refining procedural skills or managing complex clinical situations. Furthermore, some items may overlap between UX dimensions, reflecting the multidimensional nature of user experience assessment. Finally, multiple item-level comparisons increase the risk of type I error, and these results should therefore be interpreted as exploratory.

Despite these limitations, this study provides insight into how different levels of expertise may influence simulator perception and highlights the importance of considering specific educational objectives when selecting simulation modalities. Beyond the technology itself, the educational value of immersive simulation also depends on its integration into the curriculum, with learning activities adapted to learners’ levels of expertise and educational objectives [29]. Future studies with larger cohorts and objective performance assessments are needed to determine how these complementary simulation modalities can be integrated into LP training pathways.

XSL•FO
RenderX

Conclusions

Our hypothesis was partially confirmed. The AR-S was perceived as providing greater anatomical and sensory realism, while the T-S remained valuable for procedural familiarization using real equipment. However, these differences did not translate into an overall preference for the AR-S.

These findings suggest that LP simulation should not rely on a single modality but rather on selecting the appropriate simulator according to learners' levels of expertise and educational objectives. Traditional mannequin-based simulation remains relevant for novice learners who need to develop procedural skills and familiarity with real equipment, whereas AR simulation may provide additional value for more advanced learners by enhancing anatomical understanding and sensory feedback.

Future research should evaluate how these complementary simulation modalities can be optimally integrated into medical curricula and assess their impact on procedural performance, learner confidence, and ultimately patient care.

Acknowledgments

The authors thank the learners for their engagement and participation in the training program, which made this study possible. The authors used the generative AI tool ChatGPT (OpenAI) solely for English language editing and improving the clarity of the manuscript. The authors reviewed and revised all content and remain fully responsible for the final manuscript.

Funding

This study received no financial support. The InSimo team graciously provided 2 simulators for the study day, and the All'Sims simulation center of the University of Angers and the University Hospital of Angers, France, made conventional lumbar puncture simulators available. No financial compensation or material consideration was provided for these contributions. The sponsors had no access to the study data and played no role in the study design, data collection, analysis, interpretation, manuscript preparation, or the decision to submit the manuscript for publication.

Data Availability

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

Authors' Contributions

Conceptualization: CA
Formal analysis: EC
Investigation: AB, EC, FN
Methodology: CA, FN
Project administration: CA, FN
Supervision: CA
Validation: AB, CA
Visualization: EC
Writing—original draft: EC
Writing—review and editing: AB, FN, CA

Conflicts of Interest

None declared.

References

1. Wang MY, Chen KL, Huang YY, et al. Clinical utility of cerebrospinal fluid Alzheimer's disease biomarkers in the diagnostic workup of complex patients with cognitive impairment. Transl Psychiatry 2025 Apr 7;15(1):130. [doi: 10.1038/s41398-025-03345-z] [Medline: 40195333]
2. Reis AE, Spano M, Davis-Hayes C, Salama GR. Lumbar puncture complications: a review of current literature. Curr Pain Headache Rep 2024 Aug;28(8):803-813. [doi: 10.1007/s11916-024-01262-2] [Medline: 38776003]
3. Courtois-Amiot P, Cloppet-Fontaine A, Poissonnet A, et al. Hypnosis for pain and anxiety management in cognitively impaired older adults undergoing scheduled lumbar punctures: a randomized controlled pilot study. Alzheimers Res Ther 2022 Sep 2;14(1):120. [doi: 10.1186/s13195-022-01065-w] [Medline: 36056417]
4. Duits FH, Martinez-Lage P, Paquet C, et al. Performance and complications of lumbar puncture in memory clinics: results of the multicenter lumbar puncture feasibility study. Alzheimers Dement 2016 Feb;12(2):154-163. [doi: 10.1016/j.jalz.2015.08.003] [Medline: 26368321]
5. von Cranach M, Backhaus T, Brich J. Medical students' attitudes toward lumbar puncture-and how to change. Brain Behav 2019 Jun;9(6):e01310. [doi: 10.1002/brb3.1310] [Medline: 31058449]

XSL•FO
RenderX

6. Ríos-Hernández M, Jacinto-Villegas JM, Zemiti N, Vilchis-González AH, Padilla-Castañeda MA, Debien B. Development of a lumbar puncture virtual simulator for medical students training: a preliminary evaluation. Int J Med Robot 2023 Dec;19(6):e2572. [doi: 10.1002/rcs.2572] [Medline: 37671968]
7. Henriksen MJ, Wienecke T, Kristiansen J, Park YS, Ringsted C, Konge L. Opinion and special articles: stress when performing the first lumbar puncture may compromise patient safety. Neurology 2018 May 22;90(21):981-987. [doi: 10.1212/WNL.0000000000005556] [Medline: 29784718]
8. Lilamand M, Vrillon A, Gonzales-Marabal L, et al. Lumbar puncture training with healthcare simulation improves self-confidence and practical skills of French medical residents in geriatrics. Eur Geriatr Med 2023 Jun;14(3):615-623. [doi: 10.1007/s41999-023-00769-y] [Medline: 36976436]
9. Barsuk JH, Cohen ER, Caprio T, McGaghie WC, Simuni T, Wayne DB. Simulation-based education with mastery learning improves residents' lumbar puncture skills. Neurology 2012 Jul 10;79(2):132-137. [doi: 10.1212/WNL.0b013e31825dd39d] [Medline: 22675080]
10. Gaubert S, Blet A, Dib F, et al. Positive effects of lumbar puncture simulation training for medical students in clinical practice. BMC Med Educ 2021 Jan 6;21(1):18. [doi: 10.1186/s12909-020-02452-3] [Medline: 33407416]
11. Lydon S, Reid McDermott B, Ryan E, et al. Can simulation-based education and precision teaching improve paediatric trainees' behavioural fluency in performing lumbar puncture? A pilot study. BMC Med Educ 2019 May 10;19(1):138. [doi: 10.1186/s12909-019-1553-7] [Medline: 31077216]
12. Sun C, Qi X. Evaluation of problem- and simulator-based learning in lumbar puncture in adult neurology residency training. World Neurosurg 2018 Jan;109:e807-e811. [doi: 10.1016/j.wneu.2017.10.093] [Medline: 29107161]
13. Zou X, Chen L, Shi Y, et al. Virtual reality training on lumbar puncture quality (VIRTUAL): a randomized controlled trial. BMC Med Educ 2026 Feb 13;26(1):470. [doi: 10.1186/s12909-026-08762-2] [Medline: 41689025]
14. Mirbagheri A, Owlia M, Khabbazan M, Moradi M, Mohandesi F. Introducing a robotic lumbar puncture simulator with force feedback: LP Sim. Front Biomed Technol 2020;7(1):74-81. [doi: 10.18502/fbt.v7i1.2727]
15. Masud T, Ogliari G, Lunt E, et al. A scoping review of the changing landscape of geriatric medicine in undergraduate medical education: curricula, topics and teaching methods. Eur Geriatr Med 2022 Jun;13(3):513-528. [doi: 10.1007/s41999-021-00595-0] [Medline: 34973151]
16. Felten R, Bigaut K, Wirth T, et al. Advancing medical training with augmented reality and haptic feedback simulator: outcomes of a randomized controlled trial on lumbar puncture. BMC Med Educ 2025 Aug 30;25(1):1231. [doi: 10.1186/s12909-025-07536-6] [Medline: 40885925]
17. McBain KA, Habib R, Laggis G, Quaiattini A, M Ventura N, Noel GP. Scoping review: the use of augmented reality in clinical anatomical education and its assessment tools. Anat Sci Educ 2022 Jul;15(4):765-796. [doi: 10.1002/ase.2155] [Medline: 34800073]
18. Huang X, Yan Z, Gong C, et al. A mixed-reality stimulator for lumbar puncture training: a pilot study. BMC Med Educ 2023 Mar 22;23(1):178. [doi: 10.1186/s12909-023-04173-9] [Medline: 36949483]
19. Meliagros P, Diener-Brazelle J, Garber A. Using mixed reality to teach medical students lumbar punctures. Am J Med 2024 Dec;137(12):1272-1275. [doi: 10.1016/j.amjmed.2024.08.003] [Medline: 39151683]
20. Sarkar K, Coote J, Abdul R, Slayden A. Feasibility, adoption, and utility of augmented reality in lumbar puncture simulation by medical students: a pilot study. J Med Ext Real 2024;1(1):191-202. [doi: 10.1089/jmxr.2024.0001] [Medline: 42558114]
21. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161. [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
22. Färber M, Hummel F, Gerloff C, Handels H. Virtual reality simulator for the training of lumbar punctures. Methods Inf Med 2009;48(5):493-501. [doi: 10.3414/ME0566] [Medline: 19448881]
23. Vlake JH, Drop DL, 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] [Medline: 39612482]
24. Wilson E, Hewett DG, Jolly BC, Janssens S, Beckmann MM. Is that realistic? The development of a realism assessment questionnaire and its application in appraising three simulators for a gynaecology procedure. Adv Simul (Lond) 2018 Nov 8;3:21. [doi: 10.1186/s41077-018-0080-7] [Medline: 30455991]
25. Mørk G, Bonsaksen T, Larsen OS, Kunnikoff HM, Lie SS. Virtual reality simulation in undergraduate health care education programs: usability study. JMIR Med Educ 2024 Nov 19;10:e56844. [doi: 10.2196/56844] [Medline: 39560982]
26. Elfrink VL, Kirkpatrick B, Nininger J, Schubert C. Using learning outcomes to inform teaching practices in human patient simulation. Nurs Educ Perspect 2010;31(2):97-100. [Medline: 20455366]
27. Macnamara AF, Bird K, Rigby A, Sathyapalan T, Hepburn D. High-fidelity simulation and virtual reality: an evaluation of medical students' experiences. BMJ Simul Technol Enhanc Learn 2021 Jun 16;7(6):528-535. [doi: 10.1136/bmjstel-2020-000625] [Medline: 35520971]
28. Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. BMJ 2015 Nov 4;351:h5651. [doi: 10.1136/bmj.h5651] [Medline: 26537915]
29. Yu TW, Rowe M, Frantz J. Virtual reality in health professions education: qualitative descriptive study of educators' perspectives. JMIR XR Spat Comput 2026 Jun 26;3:e52925. [doi: 10.2196/52925] [Medline: 42367747]

XSL•FO
RenderX

Abbreviations

AR: augmented reality
AR-S: augmented reality simulator
FV: face validity
LP: lumbar puncture
PEV: perceived educational value
T-S: traditional simulator
UX: user experience
VR: virtual reality

Edited by I Steenstra; submitted 13.Apr.2026; peer-reviewed by BC Diego, JM Jacinto-Villegas; revised version received 02.Sep.2026; accepted 02.Sep.2026; published 21.Sep.2026.

Please cite as:
Combret E, Bourgeais A, Noublanche F, Annweiler C
User Experience Comparison of a Traditional Mannequin-Based and an Augmented Reality Lumbar Puncture Simulator in Geriatric Medicine Training: Questionnaire Study
JMIR XR Spatial Comput 2026;3:e98090
URL: https://xr.jmir.org/2026/1/e98090
doi:10.2196/98090

© Emma Combret, Alexis Bourgeais, Frédéric Noublanche, Cédric Annweiler. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 21.Sep.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

User Acceptance of a Virtual Reality Beer Pong Intervention in Postoperative Rehabilitation: Exploratory Mixed Methods Study

Tim Schneider[1*], MSc; Merle Schlender[1*], MSc; Verena Uslar[1], Dr rer nat; Sirko Pelzl[2], MSc; Daniela Salzmann[2], Dr phil; Nomi Anna Kreß[1]; Navid Tabriz[1], Dr med, PD; Dirk Weyhe[1], Prof Dr med

[1]Department for Human Medicine, University Hospital for Visceral Surgery, PIUS-Hospital, Faculty VI, University of Oldenburg, Ammerländer Heerstraße 114-118, Oldenburg, Germany

[2]apoQlar medical GmbH, Hamburg, Germany

[*]these authors contributed equally

Corresponding Author:
Tim Schneider, MSc
Department for Human Medicine, University Hospital for Visceral Surgery, PIUS-Hospital, Faculty VI, University of Oldenburg, Ammerländer Heerstraße 114-118, Oldenburg, Germany

Abstract

Background: Postoperative patients often face pain, reduced mobility, and motivational barriers, while health care systems increasingly struggle with shortages of physiotherapy staff, making the evaluation of digital rehabilitation tools highly relevant. In this context, virtual reality (VR)–based rehabilitation has already shown promise in neurological and musculoskeletal settings, but evidence regarding its acceptance and usability in postoperative inpatient rehabilitation, particularly in surgical settings with potentially older, multimorbid patients, as well as the general acceptance and usability of VR-supported physiotherapy, is still insufficiently investigated.

Objective: This exploratory mixed methods study investigated the usability and acceptance of a VR-based rehabilitation application among postoperative patients and physiotherapists in a clinical setting. Secondary objectives were to gather initial data on potential differences in usability across age and gender groups and to explore patients' and physiotherapists' attitudes toward VR as a supplement or replacement for conventional physiotherapy.

Methods: This cross-sectional exploratory mixed methods study with prospective participant recruitment and data collection included 35 postoperative inpatients (17 female participants and 18 male participants; age groups: <40, 40 - 60, and >60 years) and 7 physiotherapists (4 female participants and 3 male participants) recruited from visceral surgery and orthopedic departments. Participants used a VR-based rehabilitation demonstrator ("Beer Pong," apoQlar GmbH) delivered through a head-mounted display (Meta Quest 3) for a standardized 5-minute session involving upper-limb movements and squatting exercises. The System Usability Scale (SUS) and a self-developed questionnaire were used to assess the usability of the whole system and general attitudes toward VR-supported physiotherapy. Usability ratings were analyzed descriptively and using a 2-way ANOVA to test for effects of age group and gender. Descriptive statistics were used to summarize responses on motivation, distraction, supplementation, and preferences regarding VR use.

Results: The overall mean SUS score was 73.21 (SD 13.74, 95% CI 68.49-77.93). No significant differences in usability ratings were found between age groups or genders. Most patients (29/35, 83%) agreed that the VR scenario increases motivation, and 92% (32/35) viewed VR exercises as a useful supplement to conventional physiotherapy. Physiotherapists reported an average SUS score of 74.64 (SD 10.04, 95% CI 65.35-83.93) and expressed positive attitudes toward VR implementation, as assessed using a self-developed questionnaire. However, only 20% (7/35) of patients believed VR could replace traditional therapy entirely.

Conclusions: The results of this pilot study seem to indicate high acceptance of VR-based rehabilitation applications by patients and physiotherapists in the clinical context of a postoperative setting. As one of the first studies to evaluate VR usability specifically in early postoperative inpatient rehabilitation, this work extends existing evidence beyond previously studied rehabilitation domains. Taken together with the possible potential to increase patients' motivation while relieving work strain on physiotherapists, these findings support further investigation through larger controlled trials to assess long-term usability, clinical efficacy, and integration into routine postoperative clinical care, even for older people.

(JMIR XR Spatial Comput 2026;3:e82704) doi:10.2196/82704

KEYWORDS

immersive therapy; human-computer interaction; patient engagement; digital health; gamification; technology acceptance

XSL•FO
RenderX

Introduction

Virtual Reality in Postoperative Rehabilitation

Postoperative rehabilitation is of the greatest importance for restoring function and relieving pain after surgery. Successful rehabilitation is instrumental in improving recovery and quality of life in patients [1]. Systematic physical training tailored to an individual patient's postoperative condition, functional capacity, and surgical limitations is central to these efforts. Weyhe et al [2], for example, demonstrated in a randomized controlled trial that supervised physiotherapy or home-based exercise initiated early after the resection of pancreatic cancer is both feasible and beneficial in the postoperative course.

In recent years, virtual reality (VR) has emerged as a promising technology to enhance physical rehabilitation [3]. VR exposes people to computer-generated environments that provide multimodal sensory feedback in an attempt to increase motivation, engagement, and adherence to exercise routines [4]. Moreover, VR is increasingly regarded as a potential solution to support the delivery of high-quality physiotherapeutic care in light of the growing shortage of skilled rehabilitation professionals [5].

Evidence for the applicability of VR-assisted rehabilitation is increasing. Evaluating applicability in this context often includes measures of improvement in patients' performance and motivation, as well as the usability of the developed system. In fact, a recent systematic review by Brepohl and Leite [6] revealed several advantages with regard to those outcomes, including, for instance, improved balance, gait, postural control, and pain, as well as enhanced motivation and training performance [7-9]. However, several obstacles need to be overcome: costly technology, restricted availability of suitable rehabilitation games, technical issues like sensor inaccuracy, lack of standardization, and lower familiarity among physiotherapists may hinder extensive use [6,10-13].

To date, the majority of VR applications in rehabilitation have been introduced in neurology (54%), followed by orthopedics and trauma treatment (14%) [6]. VR-based treatments are generally intended to augment, rather than substitute for, traditional physiotherapy and are best customized to address individual patient requirements [6,14].

In patients who have experienced neurological impairments, evidence suggests that VR can support improvements in specific motor functions of the upper limb. For example, a systematic review by Darekar et al [15] found that, in children with cerebral palsy, VR interventions were compared with conventional physiotherapy—defined as therapist-led, non-VR upper-limb training, including activities such as reaching, grasping, and manipulation exercises measured with tasks like the Quality of Upper Extremity Skills Test and grip strength assessments—and were associated with positive changes in upper limb motor outcomes, including range of motion, coordination, and manual function across several randomized trials, although overall evidence certainty remains low and results varied among studies. Similarly, findings by Darekar et al [15] indicated that VR-based rehabilitation can lead to improvements in balance, mobility, walking speed, and the ability to navigate complex environments following a stroke. A recent systematic review by Wang et al [16] also identified enhancements in gait and balance in individuals with spinal cord injuries as a result of VR interventions. However, the authors emphasize that the current body of evidence remains limited, highlighting the need for further research in this area. In the context of Parkinson disease, Yu et al [17] reported a significant positive effect of VR-based therapy on balance. Moreover, a study by Lin et al [18] found that active video game–assisted therapy led to notable reductions in pain as well as improvements in mobility and functional ability in patients with musculoskeletal disorders.

With regard to motivation, it is known that patients perceive rehabilitation using video games as more enjoyable and less taxing compared to traditional programs [18,19]. This holds true even for an older patient population [20]. However, it is not known if this applies to user acceptance of VR applications as well, particularly in a surgical rehabilitation setting with early postoperative patients and clinical end users such as physiotherapists. In this context, a surgical rehabilitation setting refers to the early postoperative phase of inpatient hospital care, where physiotherapy is delivered under medical supervision and might possibly be constrained by surgical precautions, pain, fatigue, and limited mobility.

In addition to patient improvement and a decrease in motivation, an important aspect in the evaluation of VR-based rehabilitation applications is the usability of the application itself. This might encompass the hardware used for the application, as well as the application itself, or the combination of both. The System Usability Scale (SUS) has emerged as a simple tool for measuring the usability of VR applications [20], which has been validated in a number of contexts [21,22]. In recent research, the SUS was, for instance, used to evaluate a VR mirrored hand system for stroke patients or a VR application designed to work with an omnidirectional treadmill [23,24]. However, usability is mostly evaluated from the patients' point of view, whereas physiotherapists are seldom involved in the development of a VR application from the start. This might be at least one of the reasons why there is no widespread usage of VR applications [25]. Recent work has emphasized the importance of clinician involvement in the development of VR rehabilitation tasks. For example, Danso et al [26] reported on the codevelopment of a VR task with physiotherapists for poststroke visuospatial neglect, highlighting how therapist-driven design can enhance task relevance, usability, and user experience.

To summarize, especially in light of the decreasing availability of physiotherapists for postoperative mobilization, digital support systems based on VR applications seem to have the potential to increase rehabilitation success, as well as patient motivation. Therefore, they may play an increasing role in guiding rehabilitation exercises. However, currently, little is known about the applicability of and the general attitude toward VR rehabilitation applications in an early postoperative clinical setting, especially from the viewpoint of a mostly older multimorbid surgical patient population, as well as clinical physiotherapists. Therefore, before starting a larger intervention study at the University Hospital for Visceral Surgery, PIUS-Hospital Oldenburg, we aimed to evaluate whether a

XSL•FO
RenderX

typical surgical patient population in a university hospital would generally be willing to receive support during postoperative inpatient physiotherapy through a VR demonstrator and whether clinically working physiotherapists would be willing to offer such applications to their patients and recognize a potential added value in them. To help patients and physiotherapists better understand the concept, we tested the demonstrator "Beer Pong VR" by apoQlar GmbH.

Study Aim and Outcome Measures

VR-based rehabilitation can be understood as a complex intervention, as it combines technological components, user interaction, and clinical context. This study was therefore designed as an exploratory mixed methods test focusing on usability and acceptance as early-stage outcomes. In line with the early development and feasibility phases of complex intervention research, the aim was not to assess clinical effectiveness but to evaluate whether the intervention is usable, acceptable, and suitable for further investigation in postoperative rehabilitation settings in order to ascertain the feasibility of conducting a future full-scale clinical trial, which could then incorporate measurements of clinical benefit as well.

With this in mind, the primary outcome of this exploratory pilot study was the perceived usability of an exemplary integrated VR rehabilitation system as a whole, measured using the SUS, in patients depending on age and in physiotherapists.

Secondary descriptively analyzed outcomes were as follows: (1) general patients' and physiotherapists' attitudes toward VR-based exercises compared with video-based exercise formats, assessed using a self-designed questionnaire; (2) perceived reasons for the acceptance or rejection of VR-supported training; and (3) a comparison of pain levels before and during the exercise.

Methods

Ethical Considerations

The study was conducted as human participants' research in accordance with national regulations, the principles of the Declaration of Helsinki (1975), and the guidelines of Good Clinical Practice (ICH-GCP). Prior to study initiation, ethical approval was obtained from the responsible Medical Ethics Committee (reference number 2024‐164). In addition, approval was granted by the Data Protection and Information Security Management Unit at the University of Oldenburg and the staff council at PIUS-Hospital. Written informed consent was obtained from all participants prior to inclusion. Patients and physiotherapists were informed verbally and in writing about the study objectives, procedures, potential risks (eg, dizziness during VR use), data handling, and their right to withdraw at any time without consequences for their treatment or employment. No waivers of consent were applied. Participation was voluntary, and no compensation was provided. Recruitment was carried out by medical staff from the Department of Visceral Surgery at the University of Oldenburg, located at PIUS-Hospital in Oldenburg, through personal communication. To ensure privacy and confidentiality, all collected data were pseudonymized immediately after data acquisition. No directly identifiable personal information was stored together with study data. The pseudonymized dataset was stored on a password-protected computer system, with access restricted to authorized members of the research team only. Data processing complied with applicable data protection regulations. No images, videos, or supplementary materials included in the manuscript allow for the identification of individual participants. The figures presented depict anonymized data visualizations or screenshots of the virtual environment without personal identifiers. Therefore, no additional consent for the publication of identifiable images was required.

Participants

Data were collected between November and December 2024. A total of 35 postoperative patients hospitalized during this period (Table 1) and 7 physiotherapists (4 female participants and 3 male participants) participated in the study. Patients were categorized into 3 age groups for analysis: less than 40 years, 40 to 60 years, and 60 years and older.

XSL•FO
RenderX

Table . Patient characteristics of postoperative participants enrolled in a pilot study evaluating a virtual reality (VR)–based exergaming intervention (Beer Pong) in a clinical rehabilitation setting, including gender, physical activity, diagnosis, treatment type, and VR experience.

Patient characteristics	Less than 40 years (n=10)	40 - 60 years (n=12)	60 years and older (n=13)	Patients (N=35)
Gender				
Female	3	8	6	17
Male	7	4	7	18
Physical activity				
None	3	4	3	10
Various[a]	7	8	10	25
Category of condition (topographical)				
Abdominal[b]	5	6	7	18
Upper extremity[c]	1	1	3	5
Lower extremity[d]	3	4	2	9
Collum[e]	0	0	1	1
Regio gluteales[f]	1	1	0	2
Type of therapy (invasiveness)				
Minimally invasive	6	6	4	16
Open	3	5	8	16
Other[g]	1	1	1	3
Previous experience with VR[h]				
<5 min	1	0	0	1
>20 min	6	1	1	8

[a]Fitness, cycling, hiking, Hot Iron, Qi Gong, tennis, golf, swimming, yoga, rowing, squash, football, and Pilates.

[b]Appendicitis, pancreatic carcinoma, cholecystolithiasis, rectal carcinoma, malignant perforation of internal organs, gastric carcinoma, metastatic breast carcinoma, adrenal adenoma, and unclear abdominal pain.

[c]Shoulder joint arthrosis, biceps tendon rupture, and finger fracture.

[d]Femoral head necrosis, dislocated knee prosthesis, knee joint arthrosis, coxarthrosis, and ankle joint fracture.

[e]Thyroid nodules.

[f]Pilonidal sinus or abscess.

[g]Abscess incision.

[h]The remaining participants were categorized as having no previous experience with VR.

Among the physiotherapists, 3 reported prior experience with VR (2 for <5 min and 1 for >20 min). Recruitment proved challenging, primarily due to patients reporting a subjectively poor general condition, significant disease-related distress, or acute pain at the time of recruitment. Participants were recruited through consecutive sampling during routine postoperative inpatient care. Eligible patients were identified by clinical staff from the departments of visceral surgery and orthopedics and were approached in person during their hospital stay.

Patients were eligible for inclusion if they met the following criteria:

- Age 18 years and older
- Current postoperative inpatient care
- Ongoing physiotherapeutic treatment
- Ability to perform physical exercises
- Medical clearance for mobilization by a physiotherapist
- Provision of written informed consent

Patients were excluded if any of the following applied:

- Preexisting neurological conditions (eg, stroke, multiple sclerosis, Parkinson disease, amyotrophic lateral sclerosis, or epilepsy)
- Muscular disorders affecting physiotherapy participation (eg, muscular dystrophy or myasthenia gravis)
- Inability to tolerate VR headsets (eg, dizziness or nausea)
- Severe postoperative complications (eg, sepsis, significant bleeding requiring reoperation, bowel obstruction, or respiratory failure requiring intensive care)
- Insufficient German language proficiency
- Inability or unwillingness to provide written informed consent

Physiotherapists were eligible for inclusion if they fulfilled the following criteria:

- Provision of written informed consent

XSL•FO
RenderX

Physiotherapists were excluded if they met the following criteria:

- Unable to tolerate VR
- Insufficient German language proficiency
- Unwilling or unable to provide written informed consent

Procedure

Participants were introduced to the VR system, which consisted of the Meta Quest 3 headset and the "Beer Pong" demonstrator developed by apoQlar GmbH (Figure 1). This introduction included a structured familiarization phase during which the experimenter explained the VR headset, the application interface, and the task sequence in detail. Verbal instructions were provided in an explanatory and supportive manner but were not strictly standardized. Core task elements, including performing a squat before each throw, grasping the virtual ball, and throwing it into the cup, were explained consistently, while additional guidance was adapted to individual participants' needs and questions. Participants then completed 3 practice throws to familiarize themselves with the interaction mechanics. During this phase, the VR display was mirrored to a laptop to identify and resolve potential technical or usability issues. The participants wore the VR headset and engaged with the app for 5 minutes. In the game, users were required to grasp a virtual ball and throw it into cups. Before each throw, the application prompted them to perform a squat, with movement execution monitored and, if necessary, corrected by physiotherapists to address improper posture or unsafe movement patterns. Users interacted with the system using hand-based gesture control to grasp and throw the virtual ball. Motion tracking was performed with the inbuilt sensors of the Meta Quest 3 headset.

Figure 1. Virtual reality (VR)–based exergaming demonstrator ("Beer Pong") used for postoperative physiotherapeutic exercises. Illustration of the immersive VR exergame used in this pilot study. To interact with the virtual ball, users are required to perform a squat, thereby combining lower-body activation with goal-directed upper-body movements and coordination to support postoperative physiotherapeutic rehabilitation.

Immediately after the session, participants completed the SUS to assess perceived usability. In addition to the SUS, a custom-designed questionnaire was administered to assess acceptance and motivation among both patients and physiotherapists, using a 5-point Likert scale ranging from "strongly disagree" to "strongly agree."

XSL•FO
RenderX

System Usability Scale

The SUS is a validated method for measuring the subjective perceived learnability and user satisfaction of new technological applications. The SUS consists of 10 items, each rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). For odd-numbered items, the score is adjusted by subtracting 1, whereas for even-numbered items, the score is adjusted by subtracting the item rating from 5. This scoring procedure results in a value between 0 and 4 for each item. The scores of all items are then summed and multiplied by 2.5 to yield a total SUS score ranging from 0 to 100. This score is referred to as the system usability score [20]. It is important to note that the average SUS score across systems is 68, while scores of 80 or above fall within the top 10 percentile. Based on these thresholds, scores can be categorized as follows: 0 - 51.6=poor, 51.7 - 62.6=OK, 62.7 - 72.5=good, 72.6 - 78.8=excellent, and 78.9 - 100=the best imaginable [21]. The full SUS questionnaire is provided in Multimedia Appendix 1. In this study, the SUS was used to assess the perceived usability of the integrated VR rehabilitation system as a whole, rather than a single isolated component.

Specifically, participants were instructed to evaluate their overall experience with the combined system, including the Meta Quest 3, the Beer Pong demonstrator, and the basic interactions required to perform the exercises within a supervised clinical setting.

Usability ratings, therefore, reflect the participants' perception of the combined hardware-software system as experienced during the standardized intervention session and not the usability of the software application or hardware device in isolation.

Acceptance and Attitudinal Questionnaire

In addition to the SUS, a self-developed questionnaire was used to assess patient and physiotherapist attitudes toward VR-supported physiotherapeutic exercises. The questionnaire addressed motivation, perceived distraction, preferences compared with conventional exercise formats, the perceived role of VR as a supplement or replacement for standard physiotherapy, and willingness to use VR-based exercises in a home setting (Multimedia Appendix 1).

The questionnaire was designed as an exploratory instrument to capture context-specific user perceptions relevant to postoperative rehabilitation. Given the exploratory nature of the study and the limited sample size, the items were intentionally descriptive, focusing on clinically relevant aspects of user experience rather than formal theoretical constructs. Accordingly, the results derived from this questionnaire are intended to provide an initial descriptive overview of user acceptance and should be interpreted as exploratory and hypothesis-generating.

Statistics

Quantitative Analyses

The quantitative SUS data were analyzed descriptively and using ANOVA to examine the effects of age and gender. To explore within-subject differences in pain ratings before versus during the VR exercise session, a Wilcoxon signed-rank test was conducted. This nonparametric test was selected due to the ordinal nature of the pain scale and the paired structure of the data. Effect size was calculated as $r = Z/\sqrt{n}$. Given the pilot nature of the study and the limited sample size, these analyses were intended to be descriptive and hypothesis-generating rather than confirmatory; therefore, no formal power analysis was conducted. Data analysis was conducted using SPSS (version 28.0.1.0; IBM Corp), MATLAB (version R2022a; The MathWorks Inc), and Microsoft Excel (version 2408, Microsoft Corporation, 2024).

Qualitative Analyses

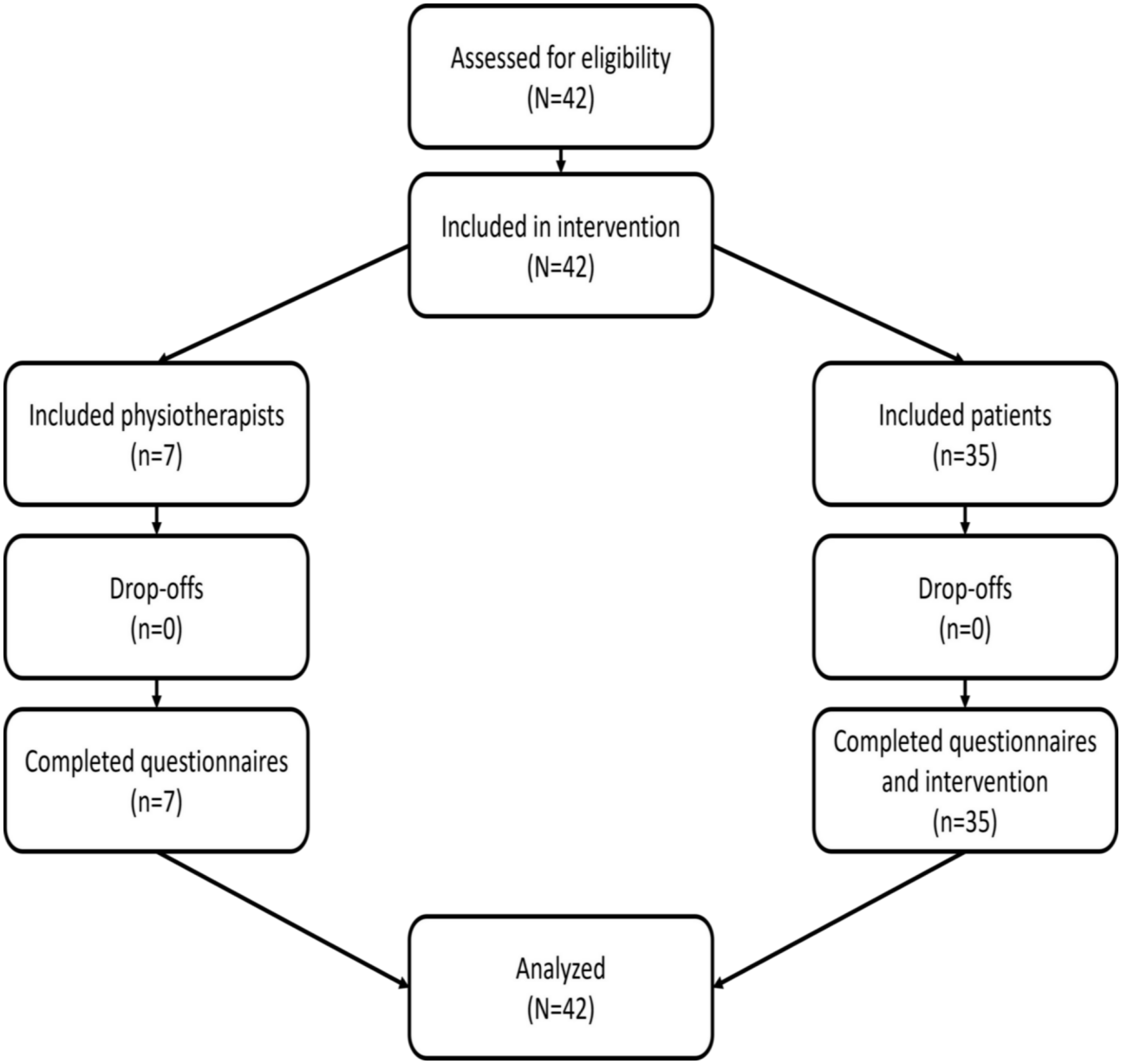
Open-ended responses were analyzed using a descriptive qualitative approach. Responses were reviewed, grouped by recurring themes (eg, motivation, perceived distraction, and usability-related comments), and summarized narratively. No formal qualitative coding framework or theory-driven qualitative methodology was applied for the 2 questions posed to patients ("Do you have any suggestions or requests?" and "Further comments?") or physiotherapists (the same questions and in addition: "In your opinion, what could the integration of VR-supported physiotherapy into everyday life look like?"). However, we roughly followed the thematic analysis framework [27]. Since our literature review prior to the study revealed no standardized or validated questionnaires that assess patients' or physiotherapists' attitudes toward the use of VR in physiotherapy, we developed our own questionnaire. The questions were initially developed by the study team with particular attention to clarity, relevance, and comprehensibility for the target groups. These questions were then presented to potential patients and the head of the physiotherapy department, who provided feedback and, if necessary, adjusted to improve clarity and ensure all relevant aspects were covered. The questionnaire was developed for exploratory purposes in this pilot study and did not undergo formal psychometric validation. Therefore, findings derived from these items should be interpreted cautiously and regarded as descriptive.

Results

The participant flowchart is presented in Figure 2. The SUS was used to assess the perceived usability of the VR application across all participants. The overall mean SUS score was 73.21 (SD 13.74, 95% CI 68.49-77.93). When stratified by age group, participants under 40 years reported a mean SUS score of 75.75 (SD 12.36, 95% CI 66.91-84.59), those aged 40 to 60 years had a mean score of 77.71 (SD 15.54, 95% CI 67.84-87.58), and participants over 60 years reported a mean score of 67.12 (SD 11.54, 95% CI 60.14-74.09). A visual comparison of SUS scores across age groups is provided in Figure 3. In addition, exploratory inferential analyses were conducted using a 2-way analysis of variance to examine potential effects of age group and gender. Homogeneity of variances was not violated (Levene test, P=.84). The exploratory analyses did not indicate statistically significant differences in usability ratings between age groups or genders ($F_{1,29}$=0.30, P=.59, partial η^2=0.010). Similarly, there was no statistically significant effect of age group on patients' SUS scores ($F_{2,29}$=2.72, P=.08, partial η^2=0.158).

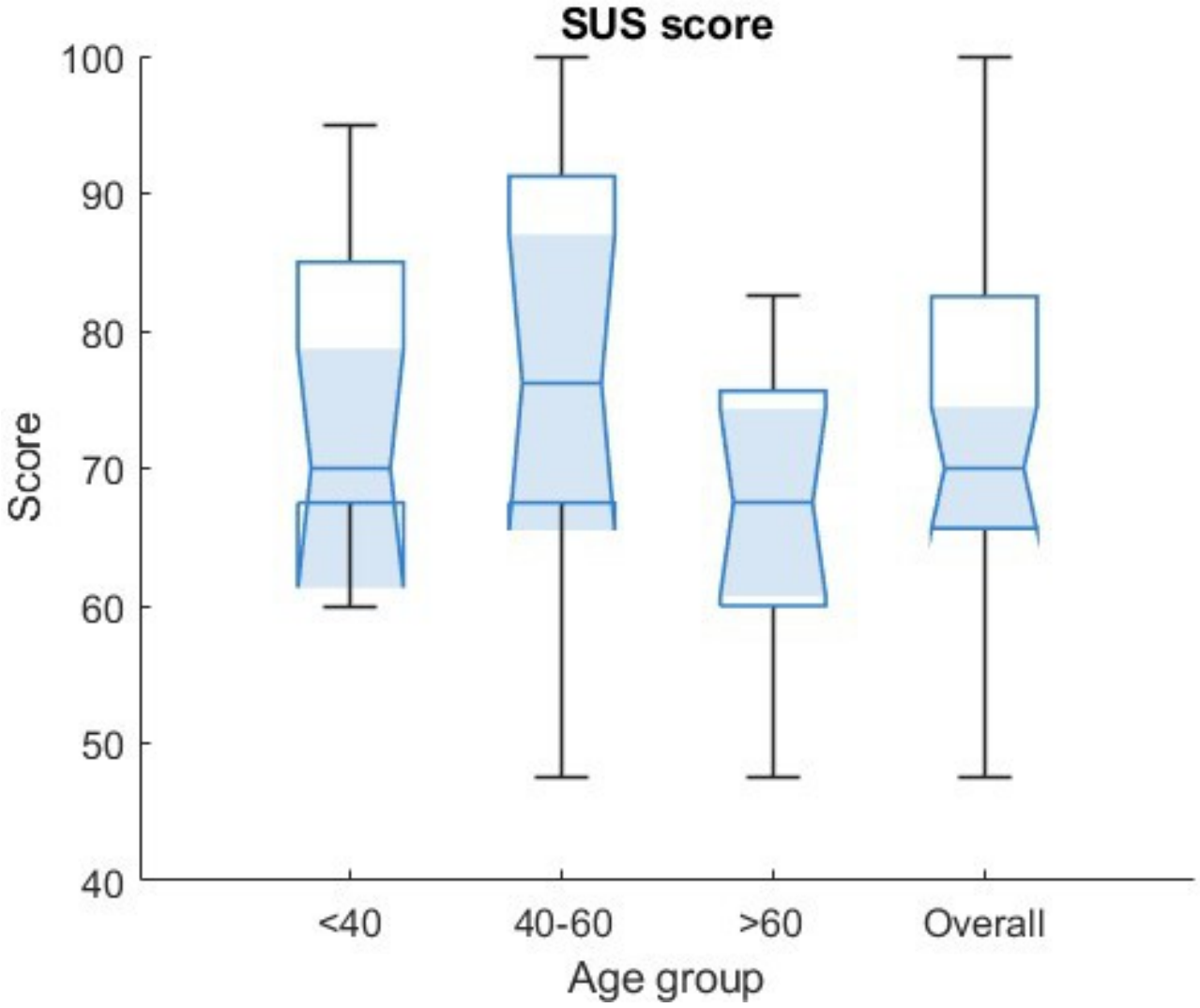
XSL•FO
RenderX

Figure 2. Flowchart of participant progression through the study, including eligibility assessment, allocation to intervention, completion status, and final analyzed sample for patients and physiotherapists.



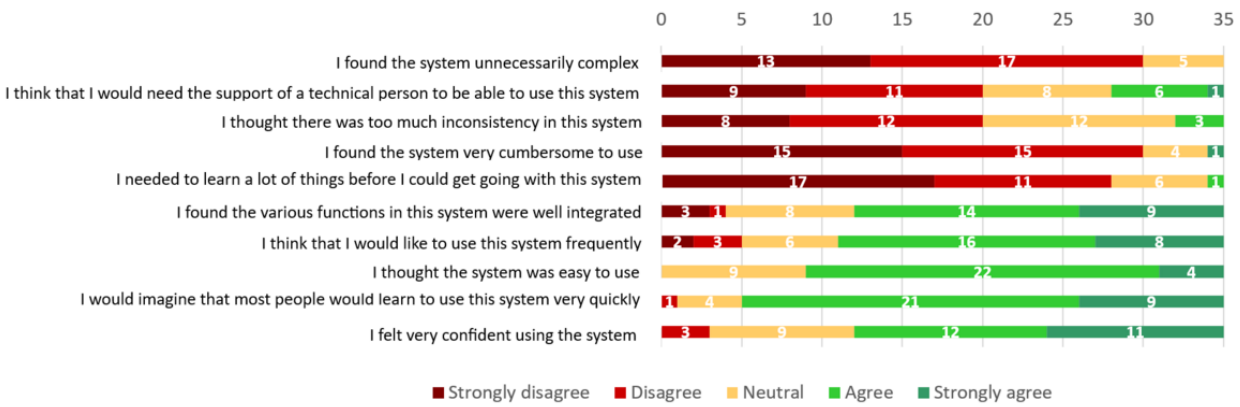
XSL•FO
RenderX

Figure 3. System Usability Scale (SUS) scores of the virtual reality–based exergaming application stratified by age group. Boxplots show SUS scores for postoperative patients grouped by age (<40, 40 - 60, >60 years) and for the overall study sample. Boxes represent the IQR, horizontal lines indicate median values, and whiskers depict the range of observed scores. Higher scores indicate greater perceived usability of the virtual reality application.



An overview of the response distribution for each of the 10 items of the SUS is presented in Figure 4. The bar chart differentiates between negatively worded items (items 1-5) and positively worded items (items 6-10). The figure illustrates the number of participants selecting each response option on the 5-point Likert scale, ranging from *strongly disagree* to *neutral* to *strongly agree*, thereby providing insight into specific aspects of perceived usability.

Figure 4. Distribution of patient responses to individual System Usability Scale (SUS) items for the virtual reality–based exergaming application. A stacked bar chart shows responses from postoperative patients (N=35) across all 10 SUS items. The upper 5 items are negatively worded, and the lower 5 items are positively worded; responses are displayed on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree." Higher agreement with positively worded items and lower agreement with negatively worded items generally indicate higher perceived usability. The figure illustrates an overall tendency toward favorable usability perceptions across most items.

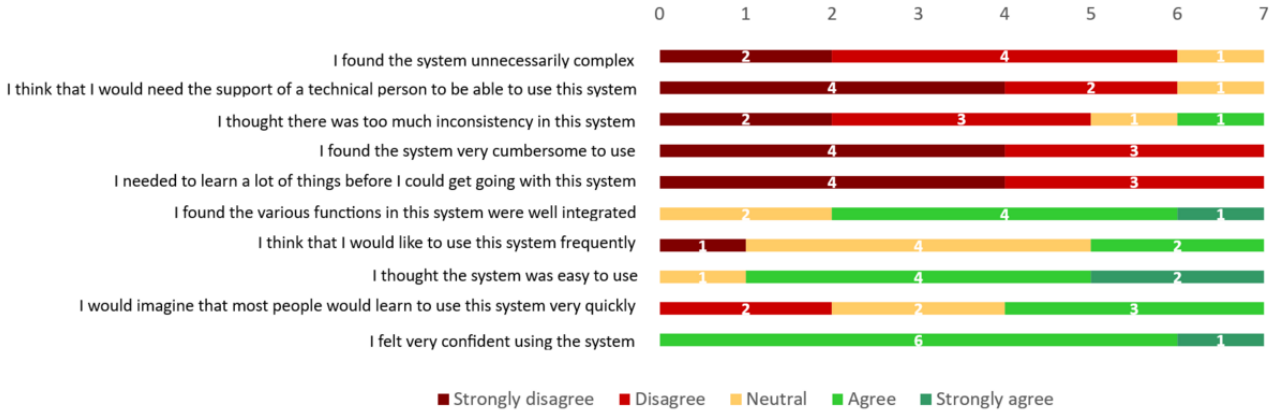


XSL•FO
RenderX

Thirty out of 35 (86%) patients disagreed with the statement that the system is unnecessarily complex. Twenty-three (66%) patients agreed that they felt confident using the system, while only 3 (9%) disagreed. Additionally, 24 (69%) patients expressed a willingness to use the system more frequently, with 6 (17%) remaining neutral.

The SUS scores obtained from the participating physiotherapists yielded a mean score of 74.64 (SD 10.04, 95% CI 65.35-83.93). The evaluation of the SUS responses from the participating physiotherapists is illustrated in Figure 5. Similar to the patient data, the first 5 items represent negatively worded statements, while the lower 5 items are positively phrased. The figure provides insight into how consistently the physiotherapists perceived the usability of the VR system across individual aspects, ranging from "strongly disagree" to "strongly agree." Six out of 7 (86%) disagreed with the statement that the system is unnecessarily complex. All 7 physiotherapists agreed that they felt confident using the system. Regarding the intention to use the system more frequently, 1 (14%) physiotherapist strongly disagreed, 2 (29%) physiotherapists agreed, while 4 (57%) remained neutral, neither agreeing nor disagreeing.

Figure 5. Distribution of physiotherapist responses to individual System Usability Scale (SUS) items for the virtual reality–based exergaming application. A stacked bar chart shows responses from physiotherapists (n=7) across all 10 SUS items. The upper 5 items are negatively worded, and the lower 5 items are positively worded; responses are displayed on a 5-point Likert scale ranging from "strongly disagree" to "strongly agree." Higher agreement with positively worded items and lower agreement with negatively worded items generally indicate higher perceived usability. The figure illustrates an overall tendency toward favorable usability perceptions across most items.



Twenty-nine out of 35 (83%) patients agreed or strongly agreed that diverse VR scenarios would enhance their motivation to practice. Thirty-three (94%) participants found VR exercises to be a welcome distraction from the hospital routine. Additionally, 32 (92%) patients could envision using VR-based exercises as a supplement to conventional physiotherapy. However, only 7 (20%) patients agreed that VR could replace traditional physiotherapy, with a considerable portion remaining neutral or disagreeing. Furthermore, 27 (77%) patients expressed willingness to perform VR-supported physiotherapeutic exercises at home. These results, obtained from a self-developed questionnaire, are summarized in Table 2.

XSL•FO
RenderX

Table . Patient-reported attitudes toward virtual reality (VR)–supported physiotherapeutic exercises in a postoperative pilot study (N=35), evaluating acceptance of a VR-based exergaming intervention for physiotherapeutic rehabilitation[a].

Question	Strongly disagree, n (%)	Disagree, n (%)	Neutral, n (%)	Agree, n (%)	Strongly agree, n (%)
Different VR scenarios for practicing would increase my motivation	1 (2.9)	1 (2.9)	4 (11.4)	19 (54.3)	10 (28.6)
Performing exercises using VR provides a welcome distraction from everyday hospital routines	—[b]	1 (2.9)	1 (0.03)	13 (37.1)	20 (57.1)
I can imagine using VR-based exercises as a supplement to conventional physiotherapy	—	1 (0.03)	2 (5.7)	15 (42.9)	17 (48.6)
I can imagine using VR-based exercises as a replacement for conventional physiotherapy	4 (11.4)	9 (25.7)	15 (42.9)	6 (17.1)	1 (2.9)
I prefer VR over traditional alternatives, such as watching exercise videos	3 (8.6)	4 (11.4)	13 (37.1)	13 (37.1)	2 (5.7)
I would also be willing to perform VR-supported physiotherapeutic exercises at home	—	2 (5.7)	6 (17.1)	20 (57.1)	7 (20)

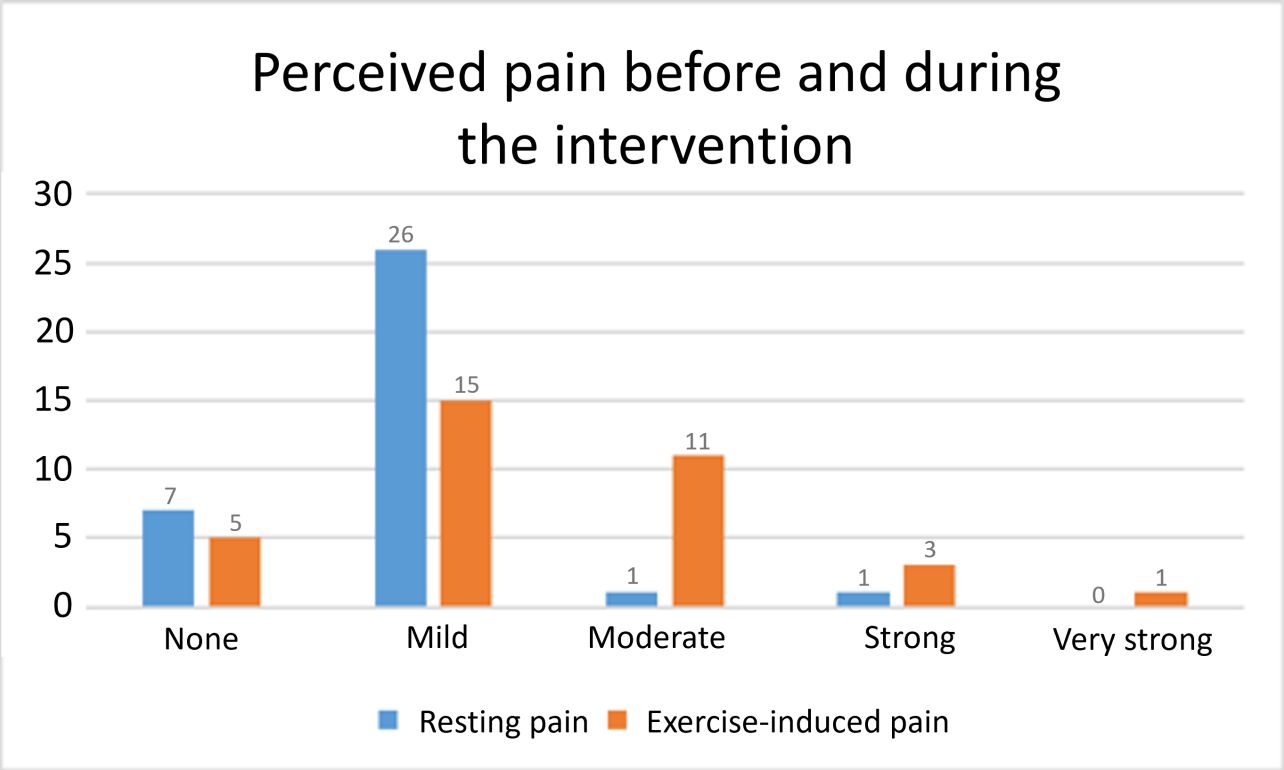
[a]Six statements assessed motivation, perceived distraction from hospital routines, suitability of VR as a supplement or replacement for conventional physiotherapy, preference compared with traditional exercise media, and willingness to use VR at home. Values are presented as absolute frequencies (n) and percentages (%) based on a 5-point Likert scale.

[b]Not applicable.

Figure 6 illustrates the distribution of patients' self-reported pain intensity levels before and during the VR-supported exercise session. Mean pain ratings increased from 0.89 (SD 0.58) before the exercise to 1.43 (SD 0.95) during the exercise. A Wilcoxon signed-rank test indicated a statistically significant within-subject difference between time points (z=3.49, P<.001, r=0.59), corresponding to a large effect size. Prior to the exercises, 26 out of 35 (74%) patients experienced mild pain, which decreased to 15 (43%) during the exercises.

XSL•FO
RenderX

Figure 6. Patient-reported pain intensity levels before and during the virtual reality–supported physiotherapeutic exercises. The bar chart shows the distribution of self-reported pain intensity levels among postoperative patients (N=35), assessed at rest before the exercise session and during the exercise session. Pain intensity was categorized as none, mild, moderate, strong, or very strong.



One out of 35 (3%) patients reported moderate pain before the exercises, increasing to 11 (31%) during the exercises. Similarly, 1 (3%) patient reported strong pain before the exercises, which rose to 3 (9%) during the exercises. Only 1 (3%) individual reported very strong pain during the exercises. Additionally, 7 (20%) patients reported no pain before the exercises, while 5 (14%) reported no pain during the exercises. The responses from the 7 participating physiotherapists to a self-developed questionnaire are summarized in Table 3. The majority expressed a positive attitude toward VR-supported physiotherapy applications, with 6 participants indicating openness to such technologies and 5 envisioning their use in clinical practice. Regarding the potential of VR exercises to alleviate staffing shortages, most responses were neutral, while opinions on whether VR could reduce daily clinical workload varied between neutral and disagreement. Most physiotherapists agreed that, following adequate patient introduction, they could provide appropriate care using VR physiotherapy. Responses concerning VR's ability to accommodate patient preferences were also predominantly neutral. Lastly, preferences between VR and more traditional alternatives, such as exercise videos, were mostly neutral among participants.

XSL•FO
RenderX

Table . Physiotherapists' attitudes toward virtual reality (VR)–supported physiotherapeutic applications in a postoperative clinical pilot study[a].

Question	Strongly disagree, n (%)	Disagree, n (%)	Neutral, n (%)	Agree, n (%)	Strongly agree, n (%)
I am open to a VR-supported physiotherapeutic application	—[b]	—	1 (14.3)	2 (28.6)	4 (58.1)
I can imagine working with VR in my practice	—	1 (14.3)	2 (28.6)	2 (28.6)	2 (28.6)
I believe that VR-supported exercises could help address understaffing within the team	1 (14.3)	1 (14.3)	4 (58.1)	—	1 (14.3)
I believe that VR-based physiotherapy could reduce workload in daily clinical routines	—	3 (42.9)	3 (42.9)	—	1 (14.3)
I believe that, following an appropriate patient introduction to VR physiotherapy, I could provide adequate patient care	—	—	5 (71.4)	1 (14.3)	1 (14.3)
I believe that VR physiotherapy could help accommodate patients' preferences	—	—	4 (58.1)	2 (28.6)	1 (14.3)
I would prefer a more traditional alternative (eg, an exercise video) over VR	—	1 (14.3)	5 (71.4)	1 (14.3)	—

[a]Responses from physiotherapists (n=7) were collected as part of a prospective pilot study. A self-developed questionnaire was used to assess attitudes toward usability, feasibility of implementation, perceived workload implications, and preferences regarding VR-supported physiotherapy. Data are presented as absolute frequencies (n) and percentages (%) based on a 5-point Likert scale.

[b]Not applicable.

Discussion

Main Findings

This exploratory mixed methods study provides initial evidence that VR-based rehabilitation might generally be well accepted in an early postoperative clinical setting by both postoperative patients and physiotherapists. The average SUS scores for the overall system for both groups indicated an "above average" to "excellent" level of usability, according to SUS interpretation conventions. These categories are heuristic and should be interpreted cautiously, particularly in the absence of confidence intervals and given the pilot nature of the study.

In addition, no significant differences in perceived usability were found between age groups or with respect to gender. The general attitude toward VR usage in a clinical rehabilitation setting seems favorable for patients as well as for physiotherapists. As an exploratory pilot study, the present findings should be interpreted as hypothesis-generating and do not allow conclusions regarding clinical benefits or causal mechanisms.

Compared with the existing literature, this study provides initial results in a research field that has previously received little or no attention. Previous studies have primarily focused on other clinical domains, such as neurological or musculoskeletal rehabilitation, rather than early postoperative settings [17,18,28]. Therefore, even these results from this exploratory pilot study are highly interesting from a clinical perspective, since early postoperative mobilization is an important component of recovery, especially after abdominal or orthopedic procedures [29]. The integration of engaging, technology-supported interventions may enhance adherence and motivation [25].

Even participants 60 years and older provided favorable ratings, contrasting with previous reports indicating that the adoption of digital health technologies among older adults remains limited and selective [30].

Furthermore, the application was also perceived positively by physiotherapists, who play a central role in selecting and delivering therapeutic interventions. This observation is consistent with recent survey evidence indicating generally favorable clinician attitudes toward immersive VR, despite limited real-world adoption. For example, a cross-sectional international survey found that most physiotherapists perceived VR as clinically useful and not overly complex, and nearly all current users intended to continue using it, although concerns

XSL•FO
RenderX

about cost and value for money were noted [31]. Similarly, a national survey in physiotherapy care reported that only a small minority of therapists currently use VR in practice, yet those who do report positive experiences and a willingness to recommend it, suggesting that adoption barriers may reflect structural rather than attitudinal limitations [25]. Their openness to VR in our study is encouraging, especially in light of the growing shortage of trained physiotherapy professionals in many health care systems. According to the World Health Organization and recent reports from national health agencies, the demand for rehabilitation services is increasing rapidly due to demographic aging, rising rates of chronic conditions, and a strained workforce [32-35]. Digital tools such as VR-based training might help in the future to bridge this gap by supporting therapists in repetitive or motivational aspects of care, thereby allowing them to allocate their time more effectively.

In addition to clinicians' perspectives, patient-reported outcomes in this study further support the feasibility of VR-assisted rehabilitation. Most participants expressed favorable attitudes toward the system, reporting increased motivation, enjoyment, and willingness to use VR-based exercises as a supplement to conventional therapy. Notably, this positive appraisal persisted despite a statistically significant increase in pain ratings during task execution. On the one hand, this pattern might suggest that transient movement-related discomfort did not undermine acceptance. On the other hand, our data are consistent with rehabilitation contexts where temporary exertion-associated pain is expected during early mobilization via standard physiotherapy [36]. Rather than indicating a negative response to the intervention, the findings presented here may reflect typical activity-related sensations during therapeutic exercise, highlighting that experiential engagement and perceived usefulness can coexist with short-term physical strain.

Limitations

Several methodological limitations should be noted. The small sample size and single-center design limit the generalizability of the findings. As SUS scores are influenced by system type and usage context, direct comparisons of absolute values across heterogeneous applications should be interpreted with caution [37].

Additionally, the study population may be subject to self-selection bias, as, first, individuals more comfortable with technology might have been more willing to participate, and second, only physically able patients were included. Both factors might potentially inflate usability ratings. The brief exposure time of 5 minutes to the VR application may not have allowed participants to fully familiarize themselves with the system, thus potentially impacting their subjective evaluations. Long-term acceptance and usability under routine clinical conditions remain to be explored.

Technical issues, including hardware limitations and software stability, were not extensively evaluated but may affect usability and integration into clinical workflows. Furthermore, adaptations might be necessary to accommodate diverse patient needs, particularly for older adults or those with physical impairments, to ensure equitable access and benefit.

Implications for Future Research

Successful clinical implementation of VR-based rehabilitation requires appropriate training for both physiotherapists and patients to maximize engagement and therapeutic benefits [25]. The positive reception among physiotherapists suggests an openness to integrating VR tools into existing rehabilitation protocols, which may help address workforce shortages by supporting repetitive or motivational tasks.

In addition to the inherent limitations mentioned above, the observed usability and acceptance may partly be explained by XR-specific interaction characteristics such as spatial computing, immersive multisensory feedback, and embodied task execution, which can facilitate intuitive interaction, enhance engagement, and strengthen perceived task relevance. Prior research suggests that immersive environments and embodied interaction can increase motivation and user involvement by creating a stronger sense of presence and agency during task performance—factors that are considered central mechanisms underlying effective XR experiences [38]. However, as this study did not address important aspects that are known to play a key role in the acceptance of VR, namely spatial computing, immersion, or embodied interaction, and as these constructs were therefore not systematically measured or analyzed in relation to usability outcomes, this should be addressed in the future. Future studies should also incorporate validated XR-specific instruments to examine how such experiential factors may moderate acceptance and performance in clinical rehabilitation contexts.

Despite the drawbacks mentioned above, this study nevertheless provides some important insights. First, while most of the existing literature focuses on clinical effectiveness, this study adds to the limited body of research addressing technology acceptance and usability in surgical rehabilitation settings. Second, by incorporating both patients and clinical staff in the evaluation process, the study highlights the dual necessity of acceptance on both ends for successful clinical implementation and the resulting fact that VR must be carefully designed to complement rather than disrupt standard care pathways. Future studies should therefore investigate the long-term efficacy and user acceptance of VR rehabilitation over extended periods among larger, more diverse populations. Randomized controlled trials comparing VR-based interventions with traditional and alternative digital therapies are warranted to establish clinical effectiveness and best practice guidelines.

Conclusions

This exploratory mixed methods study seems to indicate that VR-based rehabilitation could be generally well accepted by both postoperative patients and physiotherapists in the future, with usability ratings indicating an excellent level of user-friendliness across age groups and genders. The findings seem to suggest that VR applications can potentially serve as a feasible and engaging supplement to conventional physiotherapy in clinical settings. Notably, this study addresses a clinical context that has been underrepresented in prior VR rehabilitation research, namely early postoperative inpatient care. However, the limited sample size, brief exposure time, and single-center design indicate the need for further research.

XSL•FO
RenderX

Overall, the results of this pilot study warrant further investigation of VR applications, such as the one tested "Beer Pong" in larger clinical trials, with a focus on enhancing patient motivation and adherence, supporting physiotherapists, and potentially improving postoperative recovery outcomes. Larger randomized controlled trials with extended intervention periods are necessary to confirm these findings and evaluate long-term usability, clinical benefits, and cost-effectiveness aspects. By providing initial feasibility evidence from both patient and clinician perspectives under real clinical conditions, this study contributes preliminary insight into the practical implementation potential of immersive rehabilitation tools. In addition, larger trials should look into VR's potential to support clinical workflows, especially amid workforce shortages.

Acknowledgments

We sincerely thank the physiotherapy team at Pius-Hospital Oldenburg for their valuable support and collaboration throughout this study. We would also like to thank apoQlar GmbH for developing and providing the app, which was developed within the project: KI-SIGS—Artificial Intelligence Space for Intelligent Health Systems.

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the Generative AI Delegation Taxonomy (GAIDeT, 2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing and translation. The GAI tool used was ChatGPT 4.5. 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

Subproject: tool for human-assisted intelligent virtual agents and robots (KI-SIGS—Funding Code: 01MK20012K) was funded by the Federal Ministry for Economic Affairs and Climate Action.

Data Availability

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

Authors' Contributions

MS and TS contributed equally to the study design, conceptualization, data analysis, and writing of the manuscript. NAK was involved in the conception and design of the study, as well as in data collection. VU and NT contributed to the study design and concept. SP and DS were involved in the study design and, in particular, in the development of the app. DW was the primary idea generator for the study and was involved both in its conception and design. All authors contributed to the drafting of the manuscript and approved the final version for publication.

Conflicts of Interest

SP reports a potential conflict of interest as the founder and chief executive officer of apoQlar GmbH and DS is employed at apoQlar GmbH. All other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Multimedia Appendix 1
Questionnaires.
[PDF File, 206 KB - xr_v3i1e82704_app1.pdf]

References

1. Martínez-Fernández JL, Cuesta-Barriuso R. Postoperative rehabilitation and functional recovery after knee meniscectomy: an ambispective cohort study. Surgeries 2025 Sep;6(4):79. [doi: 10.3390/surgeries6040079]
2. Weyhe D, Obonyo D, Uslar V, Tabriz N. Effects of intensive physiotherapy on quality of life (QoL) after pancreatic cancer resection: a randomized controlled trial. BMC Cancer 2022 May 9;22(1):520. [doi: 10.1186/s12885-022-09586-1] [Medline: 35534822]
3. Howard MC. A meta-analysis and systematic literature review of virtual reality rehabilitation programs. Comput Human Behav 2017 May;70:317-327. [doi: 10.1016/j.chb.2017.01.013]
4. Li A, Montaño 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] [Medline: 21779307]
5. Elser A, Ohse M, Frankenstein C, et al. Usage of virtual reality technology in physiotherapy in Germany: results from a survey. Bioengineering (Basel) 2025 Jan 23;12(2):106. [doi: 10.3390/bioengineering12020106] [Medline: 40001626]
6. Brepohl PCA, Leite H. Virtual reality applied to physiotherapy: a review of current knowledge. Virtual Real 2023 Mar;27(1):71-95. [doi: 10.1007/s10055-022-00654-2]

XSL•FO
RenderX

7. Corbetta D, Imeri F, Gatti R. Rehabilitation that incorporates virtual reality is more effective than standard rehabilitation for improving walking speed, balance and mobility after stroke: a systematic review. J Physiother 2015 Jul;61(3):117-124. [doi: 10.1016/j.jphys.2015.05.017] [Medline: 26093805]
8. Norouzi-Gheidari N, Hernandez A, Archambault PS, Higgins J, Poissant L, Kairy D. Feasibility, safety and efficacy of a virtual reality exergame system to supplement upper extremity rehabilitation post-stroke: a pilot randomized clinical trial and proof of principle. Int J Environ Res Public Health 2019 Dec 23;17(1):113. [doi: 10.3390/ijerph17010113] [Medline: 31877910]
9. Sessoms PH, Fraser JJ, Bodell DM, et al. Clinical effectiveness of virtual reality versus conventional clinic-based vestibular physical therapy on balance and function in active duty service members. A pilot randomized controlled trial. Virtual Real 2023 Mar;27(1):263-276. [doi: 10.1007/s10055-021-00546-x]
10. Hoda M, Hoda Y, Alamri A, Hafidh B, Saddik AE. A novel study on natural robotic rehabilitation exergames using the unaffected arm of stroke patients. Int J Distrib Sens Netw 2015 Jun 1;11(6):590584. [doi: 10.1155/2015/590584]
11. Palaniappan SM, Duerstock BS. Developing rehabilitation practices using virtual reality exergaming. Presented at: 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT); Dec 6-8, 2018; Louisville, KY, USA p. 090-094. [doi: 10.1109/ISSPIT.2018.8642784]
12. Rutkowski S, Rutkowska A, Kiper P, et al. Virtual reality rehabilitation in patients with chronic obstructive pulmonary disease: a randomized controlled trial. Int J Chron Obstruct Pulmon Dis 2020;15:117-124. [doi: 10.2147/COPD.S223592] [Medline: 32021150]
13. Tsuda K, Sudo K, Goto G, et al. A feasibility study of virtual reality exercise in elderly patients with hematologic malignancies receiving chemotherapy. Intern Med 2016;55(4):347-352. [doi: 10.2169/internalmedicine.55.5275] [Medline: 26875958]
14. Cuthbert JP, Staniszewski K, Hays K, Gerber D, Natale A, O'Dell D. Virtual reality-based therapy for the treatment of balance deficits in patients receiving inpatient rehabilitation for traumatic brain injury. Brain Inj 2014;28(2):181-188. [doi: 10.3109/02699052.2013.860475] [Medline: 24456057]
15. Darekar A, McFadyen BJ, Lamontagne A, Fung J. Efficacy of virtual reality-based intervention on balance and mobility disorders post-stroke: a scoping review. J Neuroeng Rehabil 2015 May 10;12(1):46. [doi: 10.1186/s12984-015-0035-3] [Medline: 25957577]
16. Wang L, Zhang H, Ai H, Liu Y. Effects of virtual reality rehabilitation after spinal cord injury: a systematic review and meta-analysis. J Neuroeng Rehabil 2024 Oct 28;21(1):191. [doi: 10.1186/s12984-024-01492-w] [Medline: 39468617]
17. Yu J, Wu J, Lu J, et al. Efficacy of virtual reality training on motor performance, activity of daily living, and quality of life in patients with Parkinson's disease: an umbrella review comprising meta-analyses of randomized controlled trials. J Neuroeng Rehabil 2023 Sep 30;20(1):133. [doi: 10.1186/s12984-023-01256-y] [Medline: 37777748]
18. Lin YT, Lee WC, Hsieh RL. Active video games for knee osteoarthritis improve mobility but not WOMAC score: a randomized controlled trial. Ann Phys Rehabil Med 2020 Nov;63(6):458-465. [doi: 10.1016/j.rehab.2019.11.008] [Medline: 31981832]
19. Vernadakis N, Derri V, Tsitskari E, Antoniou P. The effect of Xbox Kinect intervention on balance ability for previously injured young competitive male athletes: a preliminary study. Phys Ther Sport 2014 Aug;15(3):148-155. [doi: 10.1016/j.ptsp.2013.08.004] [Medline: 24239167]
20. Brooke J. SUS: a "quick and dirty" usability scale. In: Usability Evaluation in Industry: CRC Press; 1996.
21. Gao M, Kortum P, Oswald FL. Multi-language toolkit for the system usability scale. Int J Hum Comput Interact 2020 Dec 13;36(20):1883-1901. [doi: 10.1080/10447318.2020.1801173]
22. Lewis JR, Sauro J. Item benchmarks for the system usability scale. J Usability Stud 2018;13:158-167 [FREE Full text]
23. Hsieh YW, Howe TH, Lee MT, Tai RY, Chen CC. Design and usability evaluation of an immersive virtual reality mirrored hand system for upper limb stroke rehabilitation. Sci Rep 2025 Feb 17;15(1):5781. [doi: 10.1038/s41598-025-90698-6] [Medline: 39962160]
24. Quigley A, Pierson D, Jeong J, et al. Feasibility and usability of an omnidirectional treadmill-based virtual reality rehabilitation game: a mixed-methods feasibility study. Phys Ther 2025 Nov 29;105(12):pzaf137. [doi: 10.1093/ptj/pzaf137] [Medline: 41263507]
25. Slatman S, Staal JB, van Goor H, Ostelo R, Soer R, Knoop J. Limited use of virtual reality in primary care physiotherapy for patients with chronic pain. BMC Musculoskelet Disord 2024 Feb 22;25(1):168. [doi: 10.1186/s12891-024-07285-5] [Medline: 38388377]
26. Danso A, Nijhuis P, Ansani A, et al. Development and user experiences of a novel virtual reality task for poststroke visuospatial neglect: exploratory case study. JMIR XR Spat Comput 2025;2:e72439. [doi: 10.2196/72439] [Medline: 41340838]
27. Ahmed SK, Mohammed RA, Nashwan AJ, et al. Using thematic analysis in qualitative research. J Med Surg Public Health 2025 Aug;6:100198. [doi: 10.1016/j.glmedi.2025.100198]
28. Alrashidi M, Wadey CA, Tomlinson RJ, Buckingham G, Williams CA. The efficacy of virtual reality interventions compared with conventional physiotherapy in improving the upper limb motor function of children with cerebral palsy: a systematic review of randomised controlled trials. Disabil Rehabil 2023 Jun;45(11):1773-1783. [doi: 10.1080/09638288.2022.2071484] [Medline: 35575755]

XSL•FO
RenderX

29. Okamoto T, Ridley RJ, Edmondston SJ, Visser M, Headford J, Yates PJ. Day-of-surgery mobilization reduces the length of stay after elective hip arthroplasty. J Arthroplasty 2016 Oct;31(10):2227-2230. [doi: 10.1016/j.arth.2016.03.066] [Medline: 27209333]
30. Heart T, Kalderon E. Older adults: are they ready to adopt health-related ICT? Int J Med Inform 2013 Nov;82(11):e209-e231. [doi: 10.1016/j.ijmedinf.2011.03.002] [Medline: 21481631]
31. Luc A, Slatman S, Everard G, Huygelier H, Pitance L, Detrembleur C. Knowledge, practices, and acceptance of physiotherapists regarding the use of immersive virtual reality in clinical practice: a cross-sectional online survey. Disabil Rehabil 2026 Feb 2:1-17. [doi: 10.1080/09638288.2026.2622898] [Medline: 41622875]
32. Mills JA, Gosney J, Stephenson F, et al. Development and implementation of the World Health Organization emergency medical teams: minimum technical standards and recommendations for rehabilitation. PLoS Curr 2018 Jul 9;10:ecurrents.dis.76fd9ebfd8689469452cc8c0c0d7cdce. [doi: 10.1371/currents.dis.76fd9ebfd8689469452cc8c0c0d7cdce] [Medline: 30050723]
33. Rauch A, Negrini S, Cieza A. Toward strengthening rehabilitation in health systems: methods used to develop a WHO package of rehabilitation interventions. Arch Phys Med Rehabil 2019 Nov;100(11):2205-2211. [doi: 10.1016/j.apmr.2019.06.002] [Medline: 31207218]
34. The World Health Report: health systems financing: the path to universal coverage. : World Health Organization; 2010 URL: https://iris.who.int/server/api/core/bitstreams/18e28fe3-d767-417f-811a-fd5f639a9ee4/content [accessed 2025-08-04]
35. Rehabilitation 2030: meeting report, Geneva, Switzerland, 10–11 July 2023. : World Health Organization; 2023 URL: https://iris.who.int/server/api/core/bitstreams/34d3c289-87c3-4ba3-86dc-199a94bb3e88/content [accessed 2026-05-30]
36. Kırtıl İ, Ongün P, Kişin Ş. Perceived barriers and factors affecting postoperative mobilization and related outcomes in surgical patients. J Basic Clin Health Sci 2026 Jan 31;10(1):49-63. [doi: 10.30621/jbachs.1671406]
37. Lewis JR, Sauro J. The factor structure of the system usability scale. In: Kurosu M, editor. Human Centered Design: Springer; 2009:94-103. [doi: 10.1007/978-3-642-02806-9_12]
38. 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] [Medline: 31798492]

Abbreviations

ICH-GCP: International Council for Harmonization Good Clinical Practice
SUS: System Usability Scale
VR: virtual reality

Edited by I Steenstra, T Leung; submitted 25.Aug.2025; peer-reviewed by A Danso, S Jiang; revised version received 26.May.2026; accepted 26.May.2026; published 06.Aug.2026.

Please cite as:
Schneider T, Schlender M, Uslar V, Pelzl S, Salzmann D, Kreß NA, Tabriz N, Weyhe D
User Acceptance of a Virtual Reality Beer Pong Intervention in Postoperative Rehabilitation: Exploratory Mixed Methods Study
JMIR XR Spatial Comput 2026;3:e82704
URL: https://xr.jmir.org/2026/1/e82704
doi:10.2196/82704

© Tim Schneider, Merle Schlender, Verena Uslar, Sirko Pelzl, Daniela Salzmann, Nomi Anna Kreß, Navid Tabriz, Dirk Weyhe. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 6.Aug.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Effects of Clinician Prescribing and Implementation Materials on Adoption of Virtual Reality Therapeutics: Randomized Feasibility Pilot Study

Ashlyn Zebrowski[1,2], PhD; R Jackson Fernandez[3], BS; Cameron Pinkelton[3], BS; Rebecca Kitzmiller[4], MHR, RN-BC, PhD; Adam W Kiefer[5], PhD; Laura Stanley[6], PhD; Gita Mody[3,7], MD, MPH; Carlton Moore[3], MD, MS; Lukasz Mazur[1,2], PhD

[1]Carolina Health Informatics Program, School of Information and Library Science, University of North Carolina at Chapel Hill, 335 South Columbia Street, Chapel Hill, NC, United States

[2]Division of Healthcare Engineering, Department of Radiation Oncology, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[3]Division of Cardiothoracic Surgery, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]School of Nursing, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[5]Department of Exercise and Sport Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[6]Department of Computer Science, Gianforte School of Computing, Montana State University, Bozeman, MT, United States

[7]UNC Lineberger Comprehensive Cancer Center, Chapel Hill, NC, United States

Corresponding Author:
Ashlyn Zebrowski, PhD
Carolina Health Informatics Program, School of Information and Library Science, University of North Carolina at Chapel Hill, 335 South Columbia Street, Chapel Hill, NC, United States

Abstract

Background: Virtual reality therapeutics (VRx) are emerging as a form of digital therapeutics (DTx) with growing evidence for improving symptoms such as anxiety, pain, and distress. However, real-world adoption remains limited because successful use depends not only on the therapeutic benefits but also on how the intervention is introduced, supported, and integrated into care, in addition to the therapeutic benefits. Existing literature has identified barriers related to clinician familiarity, prescribing pathways, onboarding, and independent use outside the clinic, but few studies have experimentally evaluated how implementation strategies influence early adoption outcomes.

Objective: This feasibility pilot study evaluated how different implementation strategies influence early adoption of a VRx intervention in a nonclinical setting by comparing unguided VRx use, self-directed VRx use with implementation materials, and VRx use with prescriber-led consultation and support.

Methods: An institutional review board–approved, 3-arm randomized feasibility pilot study was conducted following the Virtual Reality Clinical Outcomes Research Experts framework and reported using the CONSORT (Consolidated Standards of Reporting Trials) 2010 extension guidelines for pilot and feasibility studies. Participants were randomly allocated to 1 of 3 experimental conditions. Outcomes included technology acceptance, usability, fidelity, and tolerability. Quantitative measures were complemented by qualitative exit interviews. Analyses were exploratory, and the study was not powered for hypothesis testing.

Results: Across conditions, brief VRx exposure was associated with significant improvements in overall technology acceptance ($P<.001$), though postintervention scores remained in the neutral range, indicating early shifts rather than strong endorsement. Participants receiving provider-supported implementation demonstrated the most consistent pattern of favorable outcomes, including higher fidelity of use and significantly greater time-based adherence ($P=.04$). Usability scores were generally favorable across conditions. Tolerability was acceptable, with low levels of reported cybersickness and no discontinuations due to adverse effects. Qualitative findings highlighted the importance of guidance, expectation-setting, and perceived clinical legitimacy in shaping engagement.

Conclusions: This study demonstrates that implementation strategy is a key determinant of adoption for virtual reality–based DTx. By experimentally isolating onboarding and provider support components, this work extends prior efficacy-focused research to address implementation design. Findings suggest that brief provider involvement and structured onboarding may enhance engagement and adherence, supporting future evaluation of scalable DTx implementation models in clinical and community settings.

XSL•FO
RenderX

(JMIR XR Spatial Comput 2026;3:e90626) doi:10.2196/90626

KEYWORDS

virtual reality; digital therapeutics; human factors; usability; implementation science

Introduction

Background and Rationale

Anxiety, depression, and chronic pain are among the most prevalent and disabling health conditions globally, affecting hundreds of millions of people and contributing significantly to personal, economic, and societal burden [1-5]. Digital therapeutics (DTx), particularly virtual reality (VR)–based interventions (virtual reality therapeutics [VRx]), are emerging as clinically validated, nonpharmacological treatment options for a range of behavioral and neurological conditions [6-14]. Evidence indicates that VRx can reduce symptoms of anxiety, depression, and pain across diverse populations, including patients with cancer, older adults with chronic disease, and individuals with pain syndromes, in some cases, approaching the effectiveness of standard care [9-12,14-22]. By leveraging mechanisms such as cognitive distraction, immersive meditation, and behavioral engagement, VRx offers a scalable, patient-centered approach to symptom management that may improve access, adherence, and patient experience [9,14,23-29].

Despite promising evidence and reported acceptability, widespread adoption of DTx and VRx remains limited [21,30-35]. Barriers include insufficient provider training, unclear prescribing pathways, cost concerns, and inadequate support for patient onboarding and sustained use [12,25,27,28,36-38]. Additional challenges include workflow integration, technology limitations, and lack of structured implementation strategies [10,17,21,27,33,35,39,40]. While user-related factors such as engagement and personalization have shown positive influence on the adoption of digital interventions [32,39,41], provider behavior, organizational context, and implementation strategy are increasingly recognized as critical but underdeveloped drivers [28,34,42-44].

Many existing VRx studies do not incorporate implementation frameworks or assess perceptual and behavioral outcomes, focusing primarily on symptom efficacy [37,44-50]. As a result, there is limited insight into how patients and providers interact with VRx in real-world contexts, particularly during independent at-home use [7,11,51-54]. While prior studies suggest that patients can use VRx at home with initial training and remote support, most evidence is derived from structured or guided programs, leaving gaps in understanding unsupervised use, adherence, and integration into routine care [11,43,51,54-57]. Addressing these gaps requires experimental evaluation of implementation strategies that more closely reflect real-world use, particularly with respect to provider involvement and support structures.

Research in related areas underscores the importance of implementation structure for digital interventions. Factors such as guidance, integration into daily life, and social connectedness are key influences on engagement with digital health interventions [39], while provider training, resource allocation, and reimbursement remain barriers to VR adoption in clinical settings [58]. Organizational buy-in and workflow integration are also critical for facilitating successful implementation [59]. However, these studies provide limited insight into provider involvement or delivery models for at-home VRx use. Collectively, while evidence supporting digital and VR-based therapeutic interventions is growing, practical strategies for enabling scalable, real-world adoption remain insufficiently defined.

Implementation science provides a useful framework to address these challenges. The Behavior and Acceptance Framework (BEAR) supports systematic identification of barriers and facilitators influencing digital health implementation [60]. It integrates behavioral change theory and technology acceptance concepts to evaluate feasibility, acceptability, usability, and fidelity as key constructs for translating efficacy into real-world adoption [60]. Despite rapid growth in efficacy-focused VRx research, there remains limited experimental evidence evaluating how specific implementation strategies such as the BEAR affect real-world adoption.

Despite demonstrated clinical efficacy, many digital health technologies fail to scale due to barriers in implementation rather than therapeutic limitations. A key gap remains in understanding how provider engagement and structured implementation strategies impact DTx adoption. This study addresses this gap by evaluating how different implementation approaches affect acceptability, usability, and fidelity of VRx among healthy individuals in nonclinical settings. Participants were randomly allocated to one of three conditions: (1) unguided self-use, (2) self-directed use with asynchronous support materials, and (3) provider-supported onboarding with simulated prescription and guided setup. Acceptability and usability were assessed using the technology acceptance model (TAM) [61-63] and System Usability Scale (SUS) [64], and qualitative findings from exit interviews were mapped to BEAR [60] domains to identify implementation-related barriers and facilitators.

By systematically comparing these conditions, this study is, to our knowledge, one of the first to experimentally evaluate the impact of provider-led support versus self-directed implementation strategies on VRx adoption. These findings provide early evidence to inform scalable models for integrating DTx into routine care, particularly by clarifying the role of clinician involvement and structured support in enabling real-world use.

Objectives

The objective of this feasibility pilot study was to evaluate how different implementation strategies influence early adoption of a VRx intervention in a nonclinical setting. Specifically, the study compared unguided VRx use, self-directed VRx use with implementation materials, and provider-supported VRx use with implementation materials to assess differences in acceptability, usability, fidelity, and tolerability. We

XSL•FO
RenderX

hypothesized that participants receiving clinician-supported implementation would demonstrate more favorable adoption-related outcomes than those receiving self-directed support alone or no additional support.

Methods

Study Design

This was a 3-arm pilot study evaluating the acceptability, usability, and fidelity of a VR-based intervention (VRx) and accompanying implementation package across 3 experimental conditions. The study design was guided by the Virtual Reality Clinical Outcomes Research Experts framework [65], and the implementation package design was informed by the BEAR [60] and formative qualitative research with clinicians and end users. This study is reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) 2010 extension for pilot and feasibility studies [66]. The completed checklist is provided in Checklist 1. The CONSORT-eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist [67] was also completed and is provided in Checklist 2.

Participants were allocated to 1 of 3 experimental conditions. As a pilot study, a formal power calculation for hypothesis testing was not performed. Consistent with guidance for feasibility research, approximately 10 - 15 participants per study arm (total N=30 - 45) were targeted for enrollment, which was considered sufficient to assess study procedures, characterize variability in key outcomes, and explore patterns across implementation conditions, while remaining appropriate for an early-stage study [68,69]. Accordingly, the study was not powered to detect statistically significant differences between groups, and findings should be interpreted as exploratory.

A mixed methods approach was used to provide a comprehensive assessment of how implementation strategies and provider engagement influence early adoption. Quantitative measures, such as task completion rates and structured questionnaires, assessed acceptability, usability, and fidelity, and enabled comparison across groups. Qualitative insights from exit interviews contextualized findings by capturing user perspectives, identifying usability challenges, and highlighting perceived benefits and areas for improvement.

Participants

Overview

Study participants were recruited through university listservs, community announcements, and institutional study recruitment platforms. Recruitment and data collection took place between December 2024 and February 2025. Study data were collected and managed using REDCap (Research Electronic Data Capture; Vanderbilt University), hosted at the University of North Carolina at Chapel Hill [70,71].

Of the 39 participants enrolled, 31 (79.5%) completed the study and were included in the final analysis.

Eligibility Criteria

Participants were eligible for inclusion if they were adults aged 18 - 65 years, were in good general health with no safety concerns for VR use, and had no history of psychiatric or neurological conditions that could be exacerbated by VR exposure (eg, epilepsy or seizure disorders, severe motion sensitivity, vestibular disorders, or conditions associated with sensory overstimulation). Additional inclusion criteria included limited prior VR experience (defined as fewer than 2 lifetime uses), the ability to provide informed consent, and a motion sickness score below 21.6 on the prescreening assessment. Full screening instruments, including VR safety and prior experience questions used for eligibility screening, are provided in Multimedia Appendix 1.

Study Setting

This study was conducted outside of a clinical setting, in either (1) a human factors laboratory designed to simulate a home-like environment or (2) a community-based setting (eg, private library study spaces and private community center spaces). Measures were taken to ensure privacy, confidentiality, and consistent study procedures across all settings.

Adverse Events and Safety Monitoring

Screening procedures were used to minimize the risk of adverse events. Individuals who did not meet the inclusion criteria were not enrolled. No adverse events were reported during the study. Cybersickness susceptibility was assessed during screening using the Motion Sickness Susceptibility Questionnaire—Short Form [72]. Participants with severe susceptibility to motion sickness were excluded. Prior to VRx use and immediately following, participants were screened using the Cybersickness in Virtual Reality Questionnaire (CSQ-VR) [73] to minimize cybersickness risk and assess symptoms. Participants were instructed to discontinue use if symptoms occurred, with study personnel present to ensure safety.

Randomization

Participants were randomly allocated to 1 of 3 experimental conditions with a target 1:1:1 allocation ratio. The allocation sequence was generated in Microsoft Excel using the RAND() function prior to study initiation. Simple randomization was used; no formal blocking or stratification was applied. Assignments were made by study team members at the time of scheduling, with each enrolled participant assigned to the next available condition in the allocation sequence. Because recruitment, scheduling, and data collection occurred in parallel over a defined recruitment window (December 2024-February 2025), target group sizes were monitored during enrollment; when a scheduled participant did not attend their session, a subsequent enrolled participant was assigned to the affected condition to maintain approximate group sizes across arms. Self-reported technology comfort level and prior VR experience were monitored as baseline characteristics but were not used as formal allocation variables.

Allocation concealment from study personnel was not implemented, as schedulers required visibility of assignments to prepare condition-specific materials in advance of each session. Full blinding of participants and study personnel was

XSL•FO
RenderX

not feasible, given the unavoidably visible differences between study arms, for example, the presence of a provider in condition 3 and the presence of physical implementation materials in conditions 2 and 3. To minimize expectancy effects, participants were not informed of their assigned condition until arrival for their study appointment. The absence of allocation concealment and blinding is acknowledged as a limitation.

Intervention

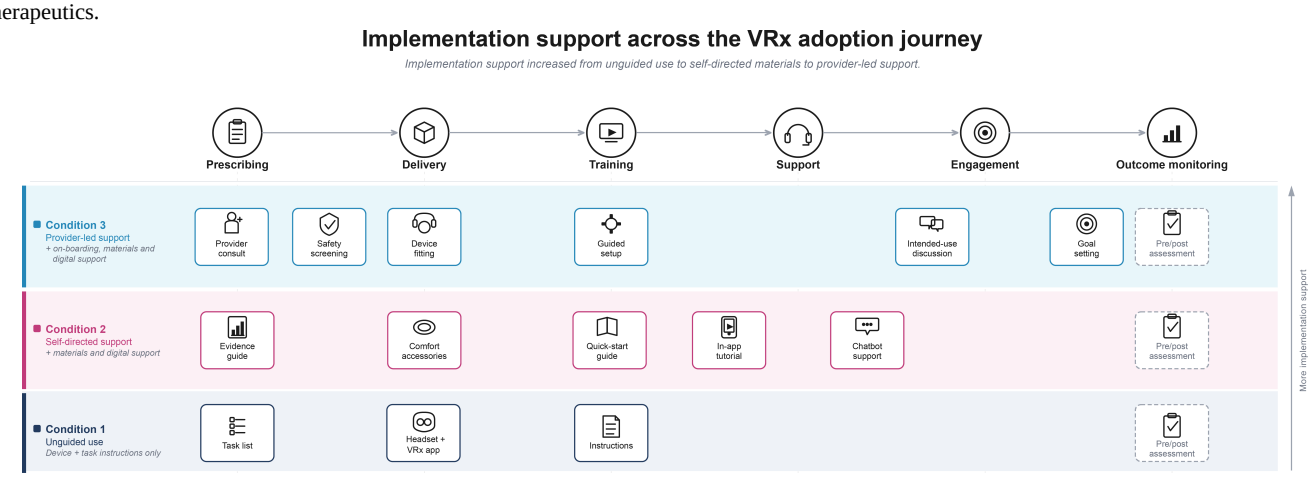
The intervention consisted of art therapy delivered using the application OpenBrush on an Oculus Quest 2 headset. Participants were instructed to complete a series of drawing tasks over a 20-minute session. Further information about OpenBrush and selection of the application is provided in Multimedia Appendix 2.

Condition-specific materials (eg, comfort accessories, tutorials, chatbot, or provider consultation) were provided to participants as outlined in Table 1. Detailed implementation package elements, equipment lists, and task protocols are provided in Multimedia Appendix 3. Differences between experimental conditions across stages of implementation are summarized in Figure 1. Representative screenshots of the VRx application interface are shown in Figure 2.

Table . Overview of experimental conditions and levels of implementation support in a 3-arm randomized feasibility pilot study of virtual reality therapeutics (VRx) conducted among healthy adults in nonclinical settings at the University of North Carolina at Chapel Hill (December 2024-February 2025).

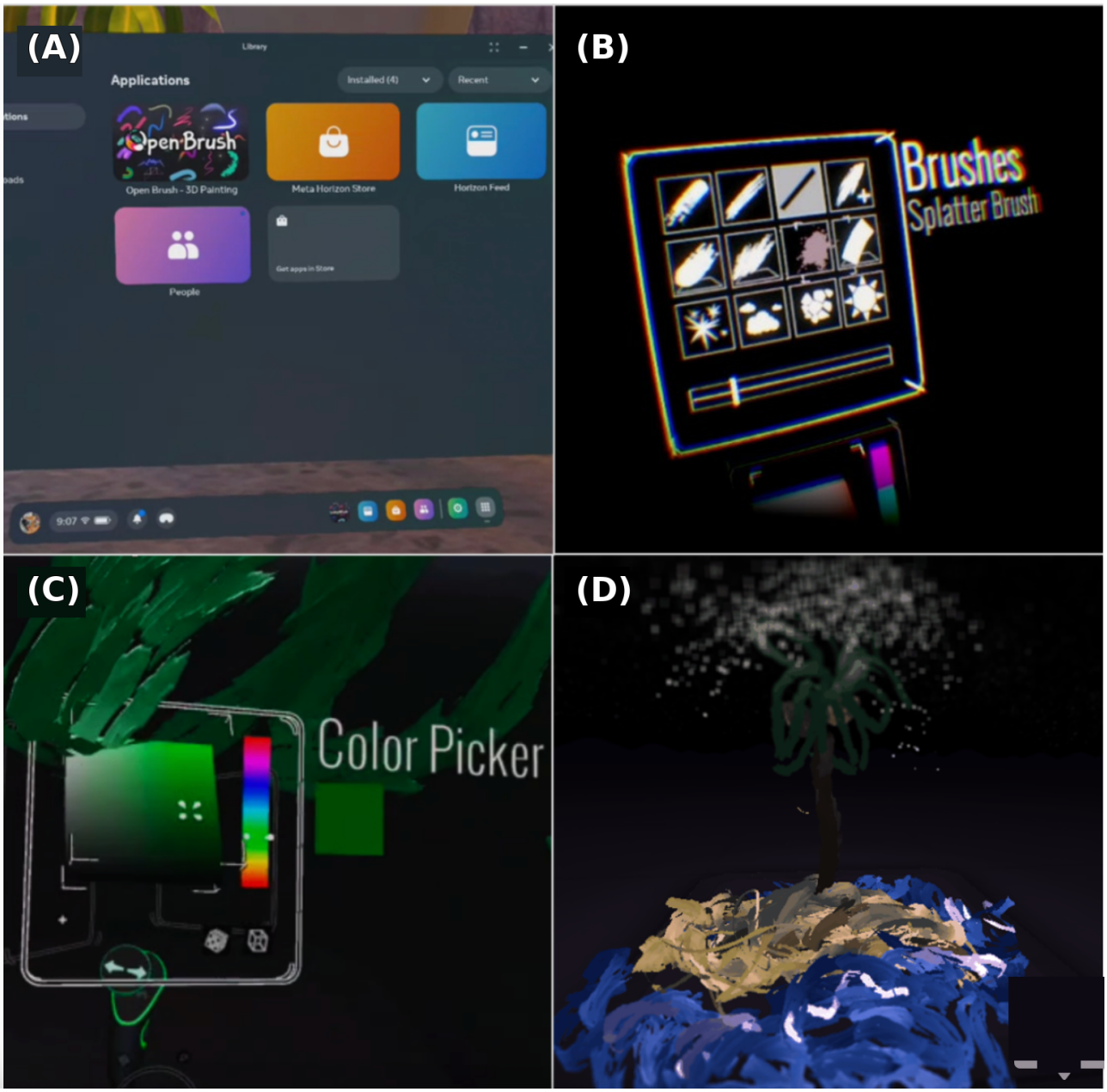
Condition	Description	Intervention details
1. Unguided use	Self-directed use with no onboarding or support materials	Virtual reality headset preloaded with VRx application, minimal instructions
2. Self-directed support	Self-directed use with asynchronous support materials	Headset with VRx application and comfort accessories, quick start guides, tutorial, chatbot support
3. Provider-led support	Implementation package plus a simulated provider encounter	Same as condition 2, plus a 5-minute scripted provider consultation including mock diagnosis, virtual reality fitness check, and guided first use

Figure 1. Conceptual overview of the implementation support solution components used across the VRx adoption journey in a 3-arm feasibility pilot study. All participants moved through the same 6 implementation stages, while support increased from unguided use to self-directed support materials and provider-led onboarding. Detailed condition-specific components are provided in Table 1 and Multimedia Appendix 3. VRx: virtual reality therapeutics.



XSL•FO
RenderX

Figure 2. Representative images of the OpenBrush virtual reality application used in the study. (A) OpenBrush application shown in the headset application library. (B) In-app brush selection menu. (C) In-app color picker tool. (D) Example 3D drawing created in OpenBrush.



The implementation package was developed based on prior qualitative research conducted by the study team exploring barriers to VRx adoption and was informed by the BEAR [60] and Virtual Reality Clinical Outcomes Research Experts framework [65] to address key implementation barriers across 6 stages of the implementation process. Package components included comfort accessories to improve device ergonomics, paper-based quick reference guides, a clinical evidence infographic, and an in-app step-by-step tutorial. The VR system menu was preconfigured for conditions 2 and 3 to display only the therapeutic application, reducing navigation complexity. A support chatbot was developed to provide real-time troubleshooting assistance. A complete description of all implementation package components, including photographs, equipment specifications, training documents, and the chatbot development process, is provided in Multimedia Appendix 3.

In the provider-supported condition, participants engaged in a standardized, scripted interaction delivered by a trained research team member with clinical training (eg, medical students). The interaction simulated a brief clinical encounter introducing VRx, including an explanation of the intervention, expected benefits, guidance on use, and a brief safety screening and device fitting. Interactions were time-limited (<5 minutes) and delivered using a predefined script to ensure consistency, as provided in Multimedia Appendix 3.

All individuals delivering the interaction were trained on the script and protocol prior to study initiation, and the script was followed closely to ensure standardized delivery. A checklist-based approach was used to ensure that all key components of the interaction were delivered consistently across participants. The provider consultation script and materials were developed based on prior qualitative research conducted by the

XSL•FO
RenderX

study team exploring clinician perspectives on DTx and VRx and were designed to simulate a realistic clinical interaction.

Data Collection

Participants completed questionnaires at session initiation and immediately following intervention use. All measurement instruments administered in this study, including screening tools and outcome measures, are provided in Multimedia Appendix 4. Observational data, including task performance, errors, help requests, and adherence, were recorded by trained observers using a structured checklist (Multimedia Appendix 5). For conditions 2 and 3, the use of implementation system components and the time spent were also recorded.

Exit interviews were conducted immediately following the session to capture user perceptions, barriers, and facilitators related to intervention use.

Primary outcomes included acceptability, usability, and fidelity; tolerability was assessed as a secondary outcome. Validated measurement instruments and techniques were used to collect both subjective and objective data, which are summarized in Table 2.

Table . Summary of measurement instruments used in a 3-arm pilot study of virtual reality therapeutics conducted from December 2024 to February 2025[a].

Construct	Measurement	Framework	Classification
Acceptability	TAM3[b]-based pre- or postquestionnaires, time in virtual reality, behavioral observations	TAM [61-63]	Subjective+objective
Usability	Task completion, error rates, help requests, time to complete, SUS[c] scores	SUS [64], task analysis	Objective+subjective
Fidelity	Adherence to protocol, task completion, duration, motivation (time engaged)	Adherence metrics	Objective
Tolerability	Pre- or postcybersickness questionnaire	CSQ-VR[d] [73]	Subjective
Qualitative insights	Exit interviews	BEAR[e] [60]	Subjective

[a]Measures collected subjective and objective data at 3 time points: during screening, at session initiation, and immediately following intervention use.

[b]TAM: technology acceptance model.

[c]SUS: System Usability Scale.

[d]CSQ-VR: Cybersickness in Virtual Reality Questionnaire.

[e]BEAR: Behavior and Acceptance Framework.

Data Analysis

Acceptability

Acceptability was assessed using a TAM-based questionnaire administered before and after the intervention, capturing perceived usefulness (PU), perceived ease of use (PEU), attitude toward technology (ATT), and behavioral intention to use (BI). Questions used a 5-point Likert scale for response.

Objective indicators of acceptability included time spent in VR and behavioral engagement during the session.

Usability

Usability was evaluated using both objective and subjective measures, including task completion rates, error frequency, help requests, time to complete tasks, and SUS scores [64,74,75].

A modified SUS was used, with 2 items removed due to the limited complexity of the system. Scores were rescaled to a 0 - 100 range. Given this modification, results are interpreted descriptively and are not directly comparable to standard SUS benchmarks.

Fidelity

Fidelity was assessed through adherence to the study protocol, including task completion, duration of engagement, and time-based adherence. Observational measures captured whether participants followed intended workflows versus engaging in exploratory behaviors.

Tolerability

Tolerability was assessed using the CSQ-VR [73], administered before and after the intervention. This measure assessed participant perceptions of discomfort and was used to monitor for potential adverse events and evaluate the role of preintervention screening in supporting safe VRx use.

Qualitative Insights

Exit interviews were conducted using a structured guide informed by the BEAR [60] to capture participant perceptions, barriers, and facilitators to VRx use. Following the approach of Braun and Clarke [76] to reflective thematic analysis, responses were evaluated using an iterative, inductive process including familiarization with the data through repeated reading of the interview notes, generation of initial codes, and grouping of codes into themes. Themes were then iteratively reviewed, refined, and defined to capture patterned meaning across

XSL•FO
RenderX

participant experiences. Resulting themes were mapped to constructs from the BEAR to contextualize findings relative to implementation-relevant barriers and facilitators.

Statistical Analysis

Overview

Data normality was assessed using Shapiro-Wilk tests. As several variables violated normality assumptions, nonparametric statistical tests were used.

Consistent with pilot study recommendations, analyses were structured to characterize the direction and magnitude of effects rather than to test formal hypotheses. Analyses were categorized as primary, secondary, or exploratory. All statistical analyses were conducted using R software (R Foundation for Statistical Computing).

Primary Analysis

Pre- and postchanges in overall TAM scores were assessed using Wilcoxon signed rank tests. Effect sizes were calculated as $r=Z/\sqrt{N}$, with 95% CIs estimated using bias-corrected and accelerated bootstrapping (2000 resamples).

Observational and behavioral data (eg, task performance, adherence, and engagement time) were summarized using descriptive statistics.

Secondary Analysis

Between-group differences in acceptability, usability, and fidelity outcomes were evaluated using Kruskal-Wallis tests. Where significant, post-hoc pairwise comparisons were conducted using Wilcoxon rank sum tests with Bonferroni-adjusted *P* values. TAM construct-level pre- and postchanges (PU, PEU, ATT, and BI) were assessed across the full sample using Wilcoxon signed rank tests. Overall TAM pre- and postchange was additionally examined within each experimental condition. Effect sizes were calculated as $r=Z/\sqrt{N}$, with 95% CIs estimated using bias-corrected and accelerated bootstrapping (2000 resamples).

Exploratory Analysis

Construct-level TAM analyses by experimental condition and SUS item-level analyses were conducted on an exploratory basis. SUS item-level analyses were conducted descriptively to characterize usability patterns across conditions. No corrections for multiple comparisons were applied to exploratory analyses, and findings should be interpreted based on overall patterns rather than individual construct-level comparisons. Additional exploratory subgroup analyses examined associations between participant characteristics and outcomes.

Modified SUS

A modified SUS [64] was used, with 2 items removed due to limited system complexity: “I found the various functions in this system were well integrated” and “I thought there was too much inconsistency in this system.” Scores were rescaled to a 0 - 100 range. Analyses were conducted at the total score level; item-level responses were interpreted descriptively. This modification may limit direct comparability to standard SUS scores and benchmarks and is acknowledged as a limitation of the study.

Multiple Comparisons

Bonferroni corrections were applied to post-hoc pairwise comparisons. No corrections were applied to within-group or exploratory analyses, consistent with the hypothesis-generating nature of pilot studies [77].

Exploratory subgroup analyses examining associations between participant demographic characteristics (eg, age, education level, technology comfort, prior VR experience) and primary outcomes were conducted to assess for potential confounding; these analyses are reported in Multimedia Appendix 6. The associated risk of inflated type I error for uncorrected comparisons is acknowledged in the Limitations section.

Ethical Considerations

This study was reviewed and approved by the UNC Institutional Review Board (IRB) at The University of North Carolina at Chapel Hill (IRB: 24 - 2894). The study was determined to meet criteria for approval under federal regulations governing human subjects research, and all procedures were conducted in accordance with applicable ethical guidelines. All participants provided informed consent prior to enrollment. The consent process included a detailed explanation of the study purpose, procedures, risks, and the voluntary nature of participation. The informed consent form is provided in Multimedia Appendix 1. To protect participant privacy and confidentiality, all data collected were deidentified and stored securely using REDCap [70,71]. Personally identifiable information was removed from all research records, and data were stored without participant names. No identifying information is disclosed in any publications resulting from this research. Figures depicting the VR intervention (Figure 2) show application screenshots only, and no photographs of participants were included in the multimedia appendices. Participants did not receive financial compensation for their participation in this study. This study was not prospectively registered, as it was designed as a pilot feasibility study focused on evaluating implementation strategies and process outcomes (eg, acceptability, usability, and fidelity) rather than clinical efficacy or health outcomes. This approach is consistent with guidance indicating that formative and feasibility studies examining implementation processes and user perceptions do not require trial registration. All study procedures and outcome measures were predefined in the IRB-approved protocol prior to data collection, and no outcomes were modified after enrollment began.

Results

Participant Characteristics

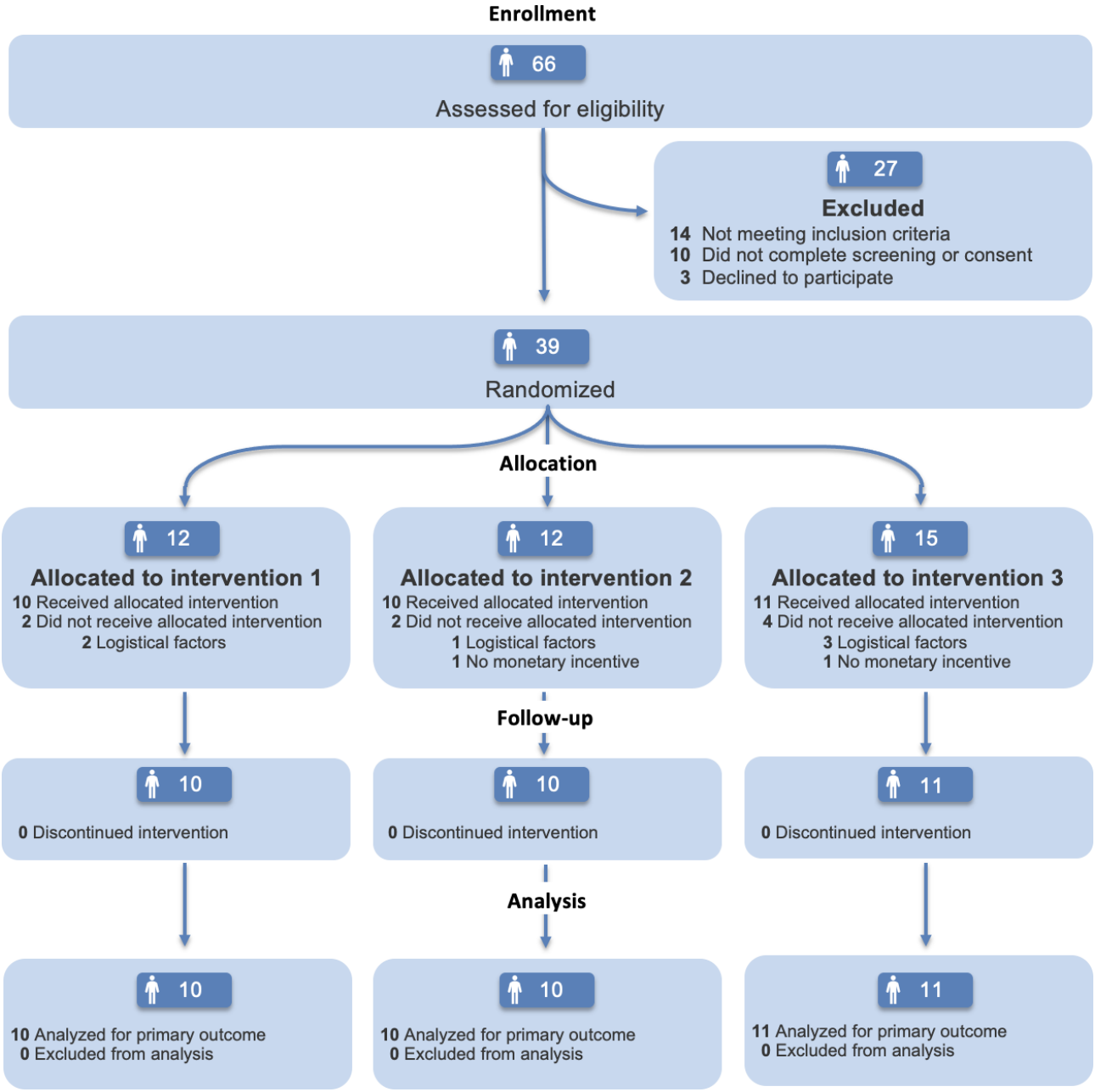
Of the 39 participants enrolled, 31 (79.5%) completed the study and were included in the final analysis. In total, 8 participants did not complete the study due to logistical factors (n=6) or study-related factors (n=2), specifically the lack of monetary incentive.

Participants had a mean age of 38.6 (SD 12.3; range 18 - 64) years, and the sample was predominantly female (20/31, 64.5%).

XSL•FO
RenderX

Figure 3 presents a CONSORT-style participant flow diagram, and Table 3 summarizes the demographic characteristics by experimental condition.

Figure 3. CONSORT flow diagram [78] illustrating participant progress through enrollment, intervention allocation, follow-up, and data analysis in a 3-arm pilot study of virtual reality therapeutics conducted from December 2024 to February 2025. CONSORT: Consolidated Standards of Reporting Trials.



XSL•FO
RenderX

Table . Demographic characteristics of study participants in a 3-arm pilot study of virtual reality (VR) therapeutics conducted from December 2024 to February 2025.

Characteristic	Overall (N=31)	Condition 1 (n=10)	Condition 2 (n=10)	Condition 3 (n=11)
Age (years), mean (SD)	38.6 (12.3)	37.3 (12.6)	40.9 (12.6)	37.6 (12.6)
Sex, n (%)				
Female	20 (64.5)	4 (40)	7 (70)	9 (81.8)
Male	10 (32.3)	6 (60)	3 (30)	1 (9.1)
Other	1 (3.2)	0 (0)	0 (0)	1 (9.1)
Education, n (%)				
High school diploma or GED	2 (6.5)	1 (10)	1 (10)	0 (0)
Some college, no degree	3 (9.68)	2 (20)	0 (0)	1 (9.1)
Associate degree	1 (3.3)	0 (0)	0 (0)	1 (9.1)
Bachelor degree	7 (22.6)	2 (20)	3 (30)	2 (18.1)
Master degree	13 (41.9)	3 (30)	4 (40)	6 (54.5)
Doctoral or professional degree	5 (16.1)	2 (20)	2 (20)	1 (9.1)
Employment, n (%)				
Employed full-time	17 (54.8)	6 (60)	5 (50)	6 (54.5)
Employed part-time	4 (12.9)	1 (10)	2 (20)	1 (9.1)
Self-employed	3 (9.7)	1 (10)	2 (20)	0 (0)
Student	5 (16.1)	2 (20)	1 (10)	2 (18.2)
Unemployed, looking for work	1 (3.2)	0 (0)	0 (0)	1 (9.1)
Unemployed, not looking for work	1 (3.2)	0 (0)	0 (0)	1 (9.1)
Overall technology comfort level, n (%)				
Low	16 (51.6)	4 (40)	6 (60)	6 (54.5)
Medium	15 (48.4)	6 (60)	4 (40)	5 (45.5)
Prior VR experience, n (%)				
Somewhat familiar	15 (48.4)	6 (60)	3 (30)	6 (54.5)
Never used	14 (45.2)	4 (40)	5 (50)	5 (45.5)
N/A—never heard of	2 (6.45)	0 (0)	2 (20)	0 (0)

Participant characteristics, including self-reported technology comfort and prior VR experience, were monitored during enrollment to assess balance across conditions. Participants were categorized as having low (16/31, 51.6%) or medium (15/31, 48.4%) technology comfort levels. Participants were assigned to experimental conditions according to a prespecified allocation sequence, with condition assignments made at the time of scheduling. Group comparisons indicated no significant differences in baseline demographic characteristics, including age, education level, or employment status.

Exploratory subgroup analyses examining associations between participant demographic characteristics and primary outcomes are reported in Multimedia Appendix 6. With the exception of acceptability scores, which differed significantly across age groups (χ^2_4=10.1; P=.04; ηH^2=0.235), no significant associations were identified. Given the small subgroup sizes and exploratory nature of this analysis, the age-related acceptability finding should not be interpreted as evidence of confounding and warrants further investigation in future adequately powered studies.

Acceptability Outcomes

Overall TAM Results

Overall, technology acceptance demonstrated statistically significant increases following intervention exposure across the full sample (Figure 4). Mean TAM scores increased from 2.84 (SD 0.32) before the intervention to 3.22 (SD 0.45) after the intervention, representing a statistically significant improvement

XSL•FO
RenderX

with a large effect size (*P*<.001; *r*=0.75, 95% CI 0.64-0.82; Table 4). However, absolute postintervention scores remained near the midpoint of the scale, indicating overall neutral to moderate levels of acceptance rather than high acceptance.

Figure 4. Pre- and postintervention TAM scores by experimental condition in a 3-arm randomized feasibility pilot study of VRx conducted from December 2024 to February 2025. Mean TAM scores increased across all conditions following intervention exposure; however, postintervention scores remained near the midpoint of the scale, indicating moderate rather than high levels of acceptance. TAM: technology acceptance model; VRx: virtual reality therapeutics.

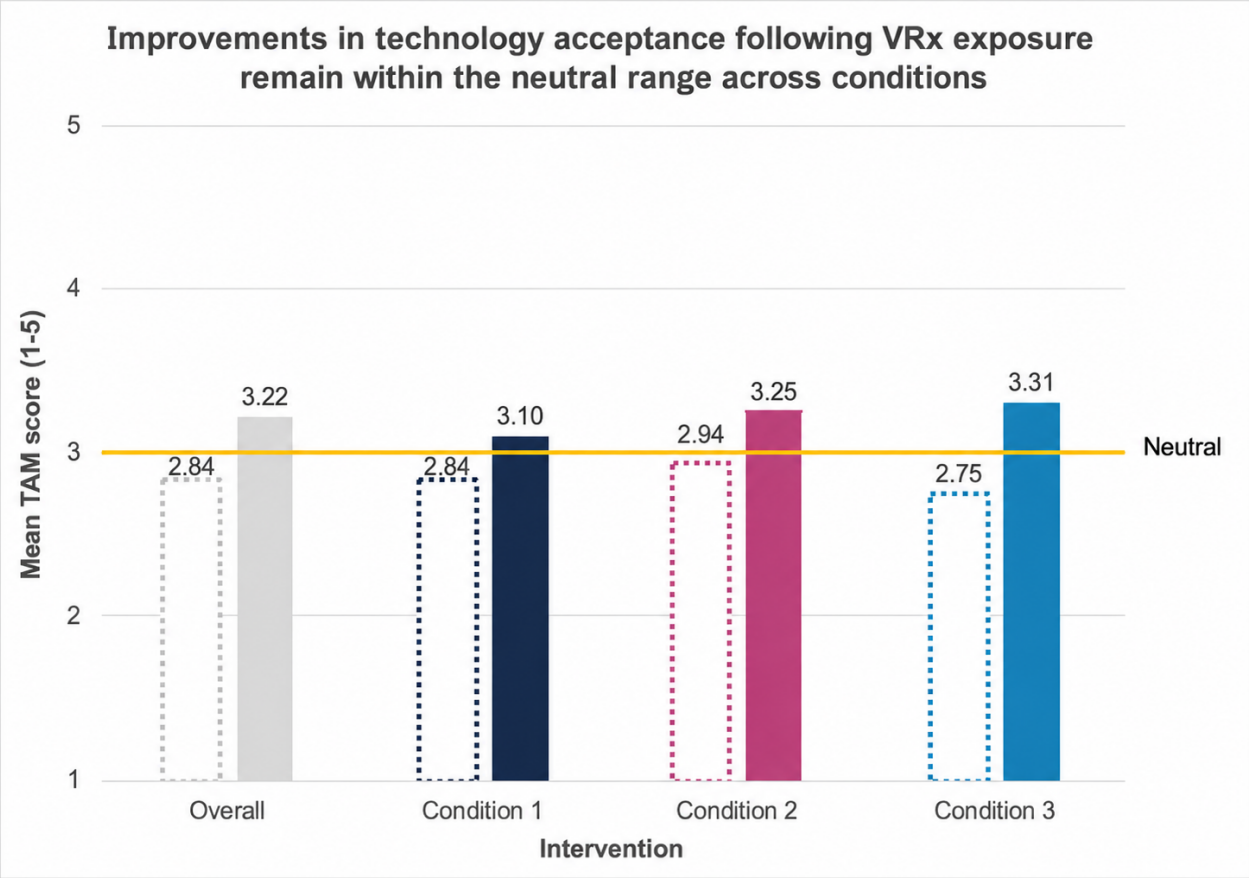


Table . Pre- and postchanges in overall mean technology acceptance model scores by experimental condition (N=31) in a 3-arm randomized feasibility pilot study of virtual reality therapeutics conducted from December 2024 to February 2025[a].

Condition	Values, n	Preintervention, mean (SD)	Postintervention, mean (SD)	Δ Mean	*W*	*P* value[b] (Wilcoxon)	Effect size, *r* (95% CI)
Condition 1 unguided use	10	2.84 (0.40)	3.10 (0.53)	+0.26	0	.03	0.70 (0.42-0.80)
Condition 2 self-directed support	10	2.94 (0.27)	3.25 (0.49)	+0.31	4	.03	0.70 (0.14-0.85)
Condition 3 provider-led support	11	2.75 (0.29)	3.31 (0.34)	+0.56	0	.005	0.85 (0.66-0.89)
Overall sample	31	2.84 (0.32)	3.22 (0.45)	+0.38	7	<.001	0.75 (0.64-0.82)

[a]Technology acceptance model scores are based on a Likert scale from 1 to 5. Pre- and postdifferences were assessed using Wilcoxon signed rank tests. Effect sizes are reported as $r=Z/\sqrt{N}$. 95% CIs were estimated using bias-corrected and accelerated bootstrapping (2000 resamples). Δ Mean represents the difference between preintervention and postintervention scores. This table reflects primary (overall sample) and secondary (condition-level) analyses.
[b]*P* values are reported for all comparisons; statistical significance was defined as *P*<.05.

At the intervention condition level, statistically significant improvements were observed across all 3 experimental conditions (Table 4). The largest improvement was observed in condition 3 (provider-led support), followed by condition 2 (self-directed support) and condition 1 (unguided use). These findings suggest that more structured implementation approaches were associated with greater improvements in technology acceptance. Despite these improvements, absolute TAM scores across all conditions remained near the midpoint of the 5-point scale, indicating that while acceptability shifted from slightly

XSL•FO
RenderX

below neutral to slightly above neutral, none of the conditions achieved a high level of acceptance.

TAM construct-level changes are summarized in Table 5; constructs examined included PU, PEU, ATT, and BI [61-63]. Across the full sample, all constructs demonstrated statistically significant improvements following intervention exposure. Effect sizes were large for PU, ATT, and BI (r=0.64, 0.63, and 0.67, respectively) and medium for PEU (r=0.48). Mean PU increased by 0.40 from 2.87 (SD 0.41) to 3.27 (SD 0.53), mean PEU increased by 0.30 from 2.89 (SD 0.53) to 3.19 (SD 0.56), mean ATT increased by 0.37 from 3.02 (SD 0.46) to 3.39 (SD 0.50), and mean BI increased by 0.45 from 2.58 (SD 0.41) to 3.03 (SD 0.64). The largest absolute improvement was observed in BI, followed by PU and ATT, while PEU demonstrated a more modest increase in comparison.

Table . Pre- and postchanges in overall mean technology acceptance model (TAM) scores at the construct level in a 3-arm randomized feasibility pilot study of virtual reality therapeutics conducted from December 2024 to February 2025[a].

TAM construct	Preintervention, mean (SD)	Postintervention, mean (SD)	Δ Mean	*W*	*P* value[b] (Wilcoxon)	Effect size, *r* (95% CI)
Perceived usefulness	2.87 (0.41)	3.27 (0.53)	+0.40	0	<.001	0.64 (0.51-0.74)
Perceived ease of use	2.89 (0.53)	3.19 (0.56)	+0.30	35	.007	0.48 (0.14-0.68)
Attitude toward technology	3.02 (0.46)	3.39 (0.50)	+0.37	0	.001	0.63 (0.49-0.72)
Behavioral intention to use	2.58 (0.41)	3.03 (0.64)	+0.45	16	<.001	0.67 (0.43-0.78)
Overall TAM	2.84 (0.32)	3.22 (0.45)	+0.38	7	<.001	0.75 (0.64-0.82)

[a]Values represent the mean pre- and postintervention TAM scores and corresponding mean change (Δ mean) within each TAM construct. TAM scores are based on a Likert scale (range 1-5). Pre- and postdifferences were assessed using Wilcoxon signed rank tests. Effect sizes are reported as $r=Z/\sqrt{N}$. 95% CIs were estimated using bias-corrected and accelerated bootstrapping with 2000 resamples. Δ Mean represents the difference between preintervention and postintervention scores. This table reflects TAM construct-level analyses conducted across the overall sample (N=31).

[b]*P* values are reported for all comparisons; statistical significance was defined as $P<.05$.

Despite these statistically significant improvements across all constructs, postintervention scores remained near the midpoint of the scale (~3.0 - 3.4), indicating moderate rather than high levels of acceptance. While improvements reflect a greater openness to VRx, postintervention results for intention to use remained comparatively lower than other constructs, suggesting that increased acceptance did not fully translate into strong intention to adopt.

Exploratory Between-Condition Construct-Level Results

Exploratory construct-level changes by experimental condition are presented descriptively in Table 6, with full inferential statistical results reported in Multimedia Appendix 6.

XSL•FO
RenderX

Table . Exploratory construct-level changes in mean technology acceptance model (TAM) scores by experimental condition in a 3-arm randomized feasibility pilot study of virtual reality therapeutics conducted from December 2024 to February 2025[a].

TAM construct	Condition 1			Condition 2			Condition 3		
	Preintervention, mean (SD)	Postintervention, mean (SD)	Δ Mean	Preintervention, mean (SD)	Postintervention, mean (SD)	Δ Mean	Preintervention, mean (SD)	Postintervention, mean (SD)	Δ Mean
Perceived usefulness	2.80 (0.42)	3.20 (0.63)	+0.40	3.10 (0.21)	3.30 (0.48)	+0.20	2.73 (0.47)	3.32 (0.51)	+0.59
Perceived ease of use	3.00 (0.33)	3.00 (0.47)	+0.00	2.70 (0.63)	3.05 (0.60)	+0.35	2.95 (0.57)	3.50 (0.50)	+0.55
Attitude toward technology	3.10 (0.61)	3.35 (0.53)	+0.25	3.15 (0.34)	3.45 (0.50)	+0.30	2.82 (0.34)	3.36 (0.50)	+0.54
Behavioral intention to use	2.45 (0.55)	2.85 (0.78)	+0.40	2.80 (0.35)	3.20 (0.63)	+0.40	2.50 (0.22)	3.05 (0.52)	+0.55
Overall TAM	2.84 (0.40)	3.10 (0.53)	+0.26	2.94 (0.27)	3.25 (0.49)	+0.31	2.75 (0.29)	3.31 (0.34)	+0.56

[a]Values represent mean pre- and postintervention TAM construct-level scores and corresponding mean change (Δ mean) within each condition. TAM scores are based on a Likert scale from 1 to 5. These analyses are exploratory and descriptive; inferential statistical results including *P* values and effect sizes are reported in Multimedia Appendix 6. No correction for multiple comparisons was applied to exploratory analyses; results should be interpreted in the context of the overall patterns.

In condition 1 (unguided use), improvements were observed primarily in BI and overall TAM scores, with smaller or negligible changes across PU, PEU, and ATT. Condition 2 (self-directed support) demonstrated modest and relatively consistent improvements across all constructs; however, the magnitude of change was generally limited and did not show clear differentiation across constructs.

In contrast, condition 3 (provider-led support) exhibited the largest and most consistent improvements across all TAM constructs. These findings suggest that more structured, human-directed implementation approaches may be associated with broader and more meaningful gains in technology acceptance, particularly in translating improvements across multiple acceptance domains.

As these analyses were exploratory and no correction for multiple comparisons was applied, findings should be interpreted in the context of overall patterns rather than individual construct-level differences.

Engagement and Satisfaction Results

Participants spent an average of 17.00 (SD 5.45) minutes engaging with the VRx intervention. Descriptive engagement and acceptability metrics by experimental condition are summarized in Table 7.

Table . Engagement and acceptability metrics in the overall sample and by experimental condition in a 3-arm feasibility pilot study of virtual reality therapeutics (VRx) conducted from December 2024 to February 2025.

Experimental condition	Values, n	Mean duration in VRx (SD) (minutes)	Mean task completion rate (SD) (%)	Mean virtual reality engagement (SD) (%)	Mean acceptability score (SD)	Mean acceptability (SD) (%)
Condition 1 unguided use	10	13.80 (6.66)	94.6 (4.5)	56.7 (41.7)	150.0 (28.3)	88.2 (16.6)
Condition 2 self-directed support	10	17.10 (4.82)	93.7 (4.7)	73.3 (34.4)	138.0 (32.4)	81.2 (19.1)
Condition 3 provider-led support	11	19.82 (3.03)	95.8 (3.7)	93.9 (13.5)	149.6 (16.7)	88 (9.8)
Overall sample	31	17.00 (5.45)	94.7 (4.3)	75.3 (34.4)	146.0 (26.1)	85.9 (15.3)

Engagement time varied by condition, with participants in condition 3 (provider-led support) demonstrating the longest duration in VRx (mean 19.82, SD 3.03), followed by condition 2 (self-directed support; mean 17.10, SD 4.82) and condition 1 (unguided use; mean 13.80, SD 6.66). A similar directional pattern was observed across VR engagement and acceptability measures, with condition 3 showing the highest levels overall.

A statistically significant difference across conditions was observed for time spent in VRx (Kruskal-Wallis χ^2_2=6.7; *P*=.04; ηH^2=0.222). Exploratory post-hoc pairwise comparisons with

XSL•FO
RenderX

Bonferroni correction suggested that participants in condition 3 spent significantly more time in VRx compared to those in condition 1 (adjusted P=.046) with no statistically significant differences observed between other condition pairs (Table 8).

No statistically significant between-group differences were observed for task completion rate, VR engagement percentage, or acceptability scores.

Table . Post-hoc Wilcoxon pairwise analysis with Bonferroni correction findings suggest that participants in condition 3 spent significantly more time in virtual reality therapeutics (VRx) compared to those in condition 1[a].

Experimental condition	N1	N2	Wilcoxon statistics	*P* value	*P*.adj
Condition 1 versus condition 2	10	10	34.5	.25	.76
Condition 1 versus condition 3[b]	10	11	20.5	.02	.046
Condition 2 versus condition 3	10	11	33.5	.13	.39

[a]Analysis is comparing time in VRx within experimental conditions in a 3-arm randomized feasibility pilot study of VRx conducted from December 2024 to February 2025.

[b]Significant difference (P<.05).

These findings suggest that more structured implementation approaches may support deeper engagement with VR interventions, although this effect was primarily reflected in time spent in VRx rather than broader engagement or acceptability measures. Given the exploratory nature of these analyses, the limited sample size, and the reliance on a single pairwise comparison, this finding should be interpreted with caution.

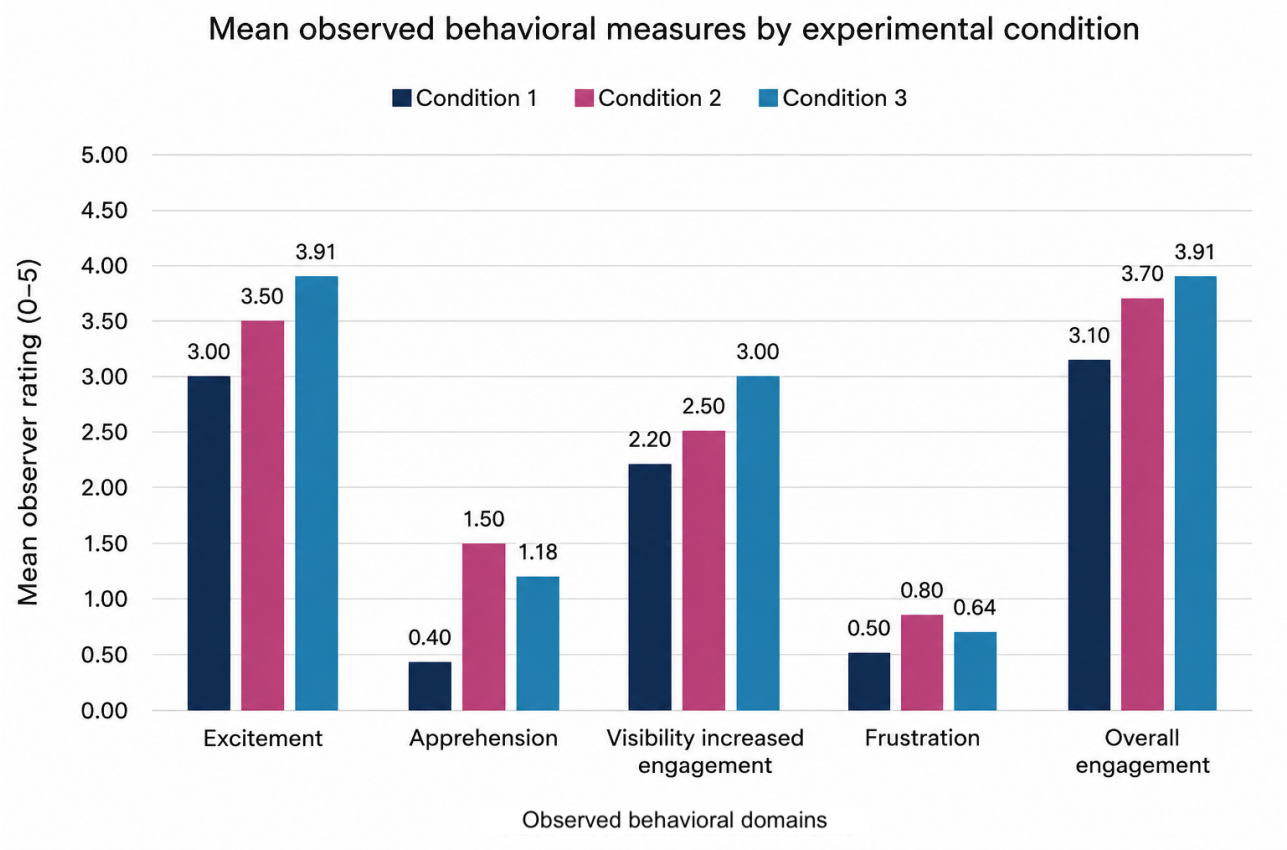
Self-efficacy and satisfaction followed a similar descriptive pattern, with participants in condition 3 reporting the highest scores (self-efficacy: mean 2.97, SD 0.46; satisfaction: mean 3.59, SD 0.44). Across all conditions, mean self-efficacy was 2.88 (SD 0.46), and mean satisfaction was 3.50 (SD 0.45). These measures were not formally compared across conditions and should be interpreted descriptively.

Observed Behavioral Results

Observed behavioral measures varied across experimental conditions (Figure 5). Participants in condition 3 (provider-led support) demonstrated higher levels of excitement and overall engagement, along with lower levels of frustration relative to other conditions. In contrast, participants in condition 2 (self-directed support) exhibited relatively higher levels of apprehension and frustration despite receiving supplemental implementation materials. Participants in condition 1 (unguided use) showed lower excitement and overall engagement in comparison to the other conditions.

XSL•FO
RenderX

Figure 5. Observed behavioral measures across experimental conditions. Mean observer ratings for excitement, apprehension, visible increased engagement, frustration, and overall engagement are shown for each condition. Higher engagement and excitement, along with lower frustration, were observed among participants receiving provider-led support (condition 3).



These observational patterns suggest that provider-led implementation support may be associated with more positive behavioral responses during VR use, including increased engagement and reduced frustration. Notably, self-directed support materials (condition 2) did not appear to reduce apprehension or frustration, suggesting that the presence of materials alone may be insufficient without structured guidance. As these measures were observational and not statistically tested, findings should be interpreted as descriptive.

Usability Outcomes

Objective Task Performance Metrics Results

Usability metrics by condition are summarized in Table 9. Overall task completion rates were high (mean 94.7%, SD 4.3%), with similar performance across conditions. The lowest mean number of errors was observed in condition 3 (provider-led support; mean 2.09, SD 1.76), while condition 2 (self-directed support) had the fewest help requests (mean 1.60, SD 2.12). Mean usability scores were highest in condition 3 (mean 61.64, SD 12.86; mean 77.4%, SD 16.1%), followed by condition 2 (mean 58.90, SD 8.26; mean 73.8%, SD 10.3%) and condition 1 (mean 56.50, SD 13.55; mean 70.8%, SD 16.9%).

Table . Usability measures (task completion rate, errors, help requests, and scores) by experimental condition in a 3-arm randomized feasibility pilot study of virtual reality therapeutics conducted from December 2024 to February 2025.

Experimental condition	Mean task completion rate (SD) (%)	Mean number of errors (SD)	Mean help requests (SD)	Mean observed usability score (SD)	Mean usability score (SD) (%)
Condition 1 (n=10)	94.6 (4.5)	3.30 (2.11)	2.22 (1.39)	56.50 (13.55)	70.8 (16.9)
Condition 2 (n=10)	93.7 (4.7)	3.50 (1.96)	1.60 (2.12)	58.90 (8.26)	73.8 (10.3)
Condition 3 (n=11)	95.8 (3.7)	2.09 (1.38)	2.09 (1.76)	61.64 (12.86)	77.4 (16.1)
Overall (N=31)	94.7 (4.3)	2.94 (1.88)	1.97 (1.75)	59.10 (11.63)	74.1 (14.5)

XSL•FO
RenderX

Help requests were examined to characterize usability challenges across conditions. The most common requests were related to navigating the VR system, selecting menu options, and exiting the application. Task 1 generated the highest number of help requests, particularly for powering on the device, setting up boundaries, and using the controllers. These findings suggest that initial onboarding and system navigation represent key usability challenges in VR-based interventions.

SUS Results

Overall SUS scores summarized in Table 10 indicated relatively high usability scores across the full sample (mean 83.74, SD 12.52), with condition 3 demonstrating the highest mean score (mean 88.26, SD 10.34), followed by condition 2 (mean 82.50, SD 14.00) and condition 1 (mean 80.00, SD 12.85). No statistically significant between-group differences were observed (Kruskal-Wallis χ^2_2=2.3; *P*=.31). Given the modified scale used in this study, scores are interpreted as a relative measure of usability across conditions rather than against standard SUS benchmarks.

Table . System Usability Scale (SUS) scores by experimental condition in a 3-arm randomized feasibility pilot study of virtual reality therapeutics conducted from December 2024 to February 2025[a].

Experimental condition	Values, n	Mean SUS score (SD)
Condition 1	10	80.00 (12.85)
Condition 2	10	82.50 (14.00)
Condition 3	11	88.26 (10.34)
Overall	31	83.74 (12.52)

[a]Between-group differences were assessed using Kruskal-Wallis tests. No statistically significant between-group differences were observed for SUS scores (χ^2_2=2.3; *P*=.31).

Exploratory item-level findings are presented in Multimedia Appendix 6 and interpreted descriptively. Across conditions, higher ratings were observed for intention to use and PEU, while lower ratings were observed for learnability and need for assistance, suggesting that onboarding and independent use represent key areas for improvement.

Fidelity Outcomes

Fidelity outcomes by experimental condition are summarized in Table 11, including overall protocol adherence, time-based adherence, and task-based adherence. Mean fidelity scores were highest in condition 3 (mean 70.9, SD 11.4) relative to conditions 1 (mean 58.0, SD 16.5) and 2 (mean 57.5, SD 14.0). Time-based adherence differed across conditions (Kruskal-Wallis χ^2_2=6.7; *P*=.04), with post-hoc comparisons indicating higher adherence in condition 3 relative to condition 1 (*P*.adj=.046). No other pairwise differences were observed.

Table . Fidelity outcomes by experimental condition, including overall fidelity score, time and task-based adherence, and task completion prior to engaging in exploratory behaviors in a 3-arm randomized feasibility pilot study of virtual reality therapeutics conducted from December 2024 to February 2025[a].

Experimental condition	Values, n	Mean fidelity score (SD)	Time-based adherence, mean (SD) (%)	Task-based adherence, mean (SD) (%)	Completed tasks prior to exploratory behavior, n (%)
Condition 1	10	58.0 (16.5)	69.0 (33.3)	94.6 (4.5)	6 (60)
Condition 2	10	57.5 (14.0)	85.5 (24.1)	93.5 (4.8)	6 (60)
Condition 3	11	70.9 (11.4)	99.1 (15.1)	96.0 (3.7)	1 (9)
Overall	31	62.1 (14.0)	85.0 (27.2)	94.7 (4.3)	13 (41.9)

[a]Between-group differences were assessed using Kruskal-Wallis tests. Time-based adherence differed significantly across conditions (χ^2_2=6.7; *P*=.04; ηH^2=0.222). No significant differences were observed for overall fidelity scores (χ^2_2=5.4; *P*=.07) or task-based adherence (χ^2_2=1.3; *P*=.52).

Task-based adherence remained high across all conditions (mean 94.7%, SD 4.3%). No statistically significant between-group differences were observed for overall fidelity scores (Kruskal-Wallis χ^2_2=5.4; *P*=.07) or task-based adherence (Kruskal-Wallis χ^2_2=1.31; *P*=.52). Although the difference in overall fidelity scores did not reach statistical significance, the *P* value approached the conventional threshold (*P*=.07), suggesting a potential trend toward higher protocol adherence in condition 3 that warrants examination in future adequately powered studies.

Participants in condition 3 were less likely to explore the environment before completing assigned tasks when compared to conditions 1 and 2, suggesting greater adherence to the prescribed task sequence. No significant differences were observed for overall fidelity scores between conditions. Given the exploratory nature of these analyses and small sample size, findings should be interpreted as descriptive.

XSL•FO
RenderX

Tolerability Outcomes

Participants completed the CSQ-VR prior to and following intervention exposure. Conditions 1 and 2 completed the assessment independently, whereas participants in condition 3 were administered the assessment as part of the VR Fitness Assessment conducted by the mock provider. Tolerability outcomes, measured using the CSQ-VR before and after the intervention, are presented by experimental condition in Table 12.

Table . Tolerability outcomes by experimental condition measured using the Cybersickness in Virtual Reality Questionnaire (CSQ-VR) administered pre- and postintervention in a 3-arm randomized feasibility pilot study of virtual reality therapeutics conducted from December 2024 to February 2025[a].

Experimental condition	Values, n	Mean preintervention score (SD)	Mean postintervention score (SD)
Condition 1	10	6.10 (0.32)	6.90 (1.20)
Condition 2	10	6.00 (0.00)	6.20 (0.63)
Condition 3	11	6.73 (1.01)	6.27 (0.90)
Overall CSQ-VR score	31	6.29 (0.69)	6.45 (0.96)

[a]Scores can range from 6 to 42.

No meaningful increases in cybersickness symptoms were observed following VRx exposure. Mean preintervention symptom scores were low across conditions (mean 6.29, SD 0.69), corresponding to absent to very mild symptoms. Postintervention scores remained similar (mean 6.45, SD 0.96), with no evidence of symptom worsening reported across conditions.

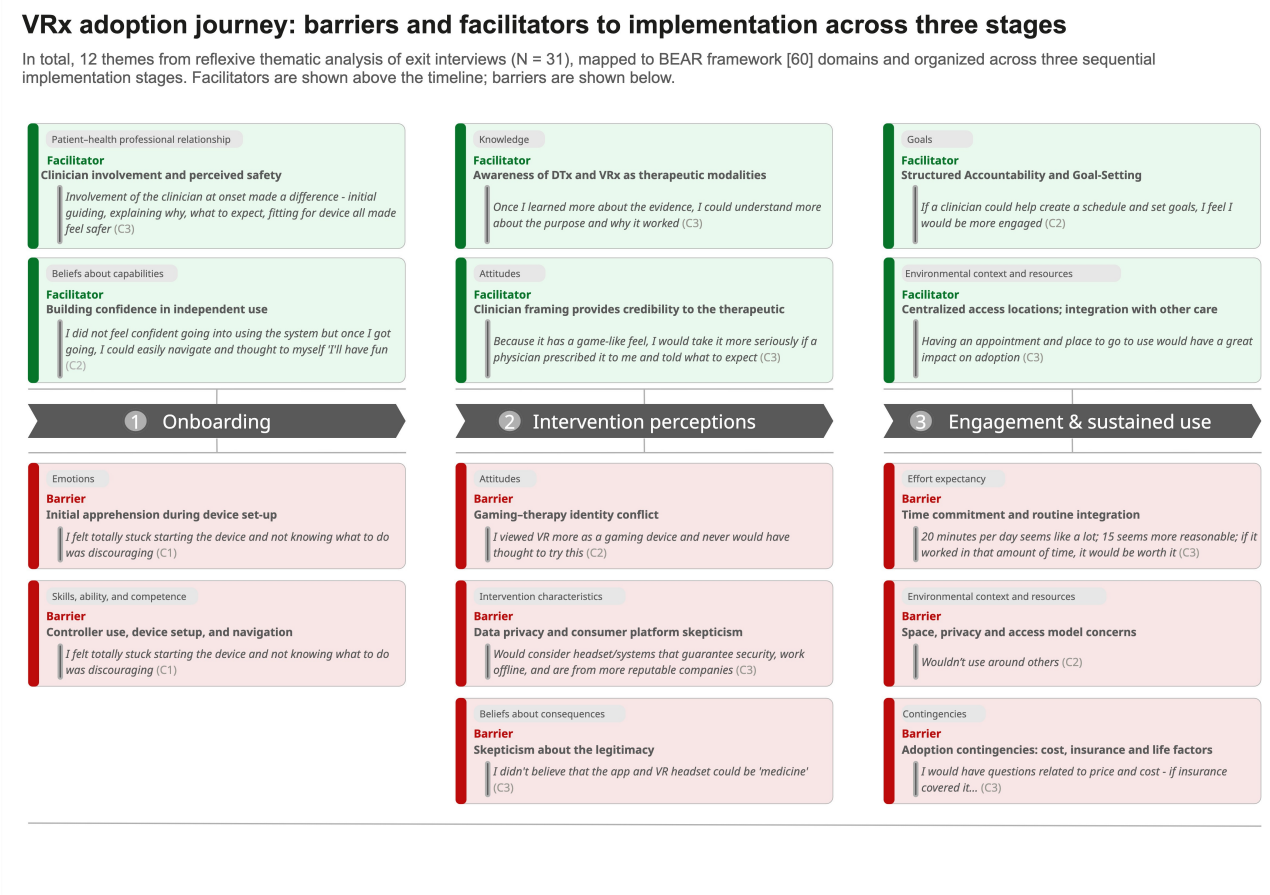
Overall, 74% (23/31) of participants reported no symptoms following the intervention. Among those who reported symptoms, the most common were visual discomfort, dizziness, eye strain, and imbalance or instability. These findings suggest that the VR-based intervention was well tolerated in this sample under the implemented safety and screening procedures.

Qualitative Insights

Qualitative insights are presented according to 14 themes derived from reflexive thematic analysis and mapped to BEAR [60] domains to contextualize implementation-relevant barriers and facilitators. Figure 6 presents key themes organized into 3 stages: onboarding, intervention perceptions, and engagement and sustained use, with representative participant quotes from exit interviews. The full BEAR thematic analysis mapping is provided in Multimedia Appendix 6.

XSL•FO
RenderX

Figure 6. Barriers and facilitators to VRx adoption mapped to BEAR domains across 3 implementation stages. Themes derived from reflexive thematic analysis of exit interviews (N=31). Representative participant quotes are included within each theme card. BEAR: Behavior and Acceptance Framework; C: condition; C1: condition 1 (unguided); C2: condition 2 (self-directed support); C3: condition 3 (provider-supported); DTx: digital therapeutics; VRx: virtual reality therapeutics.



Findings highlight several key factors influencing acceptability, usability, and fidelity. First, emotional responses during onboarding shaped early engagement trajectories. Participants across conditions described initial apprehension, particularly around device setup and controller use, that transitioned into enjoyment and relaxation once interaction with the application began. Participants who engaged in a provider-led consultation (condition 3) expressed higher levels of confidence and reduced uncertainty. Participants who received no support (condition 1) were more likely to describe frustration and discouragement during initial setup. Participants emphasized that the first interaction was a critical adoption threshold, noting that “first touch needs to be positive or people will walk away from using it, especially older people” (condition 1). Provider involvement was consistently identified as the most influential factor in shaping perceived safety, confidence, and motivation during onboarding, with participants describing cybersickness screening, device fitting, and expectation-setting as contributing to a greater sense of comfort and legitimacy.

Perceptions of VRx as a therapeutic modality were strongly influenced by how the intervention was framed and presented. Several participants initially viewed VR as a gaming platform, expressing skepticism about its legitimacy as a health intervention. Provider endorsement as well as evidence-based communication materials shifted these perceptions, with participants noting that professional framing made them take the intervention more seriously. A distinct barrier emerged around identity conflict, in which participants struggled to categorize VRx as medicine rather than entertainment, describing engagement with a game-like interface as therapeutic felt like self-deception without clinical framing.

Concerns related to data privacy and the use of consumer technology platforms for health interventions also emerged, with participants expressing preferences for offline functionality, local data storage, and health-dedicated devices. Limited prior awareness of DTx as a therapeutic modality was common across conditions.

Environmental constraints, such as limited home space, concerns about safety, and apprehension about the immersive properties of VR when others are in the home, shaped preferences for more structured settings and centralized locations where patients can go to use VRx outside of the home. For sustained use, participants emphasized the importance of flexible session length, integration into daily routines, and structured accountability and goal setting. Many expressed that without some form of structure, such as goal setting with providers or progress tracking, they would have difficulty maintaining regular engagement. Participants across conditions raised concerns about the affordability of VR hardware and suggested centralized access models such as clinic-based or community “studios” as alternatives for individual device ownership.

XSL•FO
RenderX

Collectively, these findings highlight the central role of provider involvement, structured onboarding, and contextual fit in supporting adoption and sustained use of VR-based DTx.

Discussion

Principal Findings

This pilot study evaluated 3 implementation strategies for VR-based DTx, comparing unguided use, self-directed use with support materials, and provider-supported use. A single brief exposure to VRx was associated with statistically significant improvements in acceptability across all conditions. However, postintervention acceptance scores remained within the neutral range, suggesting that while perceptions shifted favorably, a single exposure may be insufficient to produce strong endorsement or readiness for sustained adoption.

Differences across intervention conditions were observed in acceptability, usability, and fidelity-related outcomes. Participants receiving provider-led support demonstrated the most favorable and consistent pattern of favorable outcomes, including longer engagement times, higher protocol adherence, and more positive usability ratings. In contrast, self-directed materials produced more variable engagement and, in some cases, appeared to increase cognitive burden during onboarding. Tolerability was acceptable across all conditions, with no meaningful increases in cybersickness symptoms observed. Qualitative findings reinforced the importance of the initial interaction, clinician endorsement, and structured onboarding in shaping early adoption trajectories.

Interpretation of Findings

The findings suggest that early user responses to VR-based DTx are shaped not only by exposure to technology itself but also by the implementation strategy through which that exposure is delivered. While all participants demonstrated improved perceptions of VRx following a single exposure, the differences across conditions indicate that acceptability, usability, and fidelity are strongly influenced by the surrounding implementation approach rather than inherent properties of the intervention alone.

Participants in the provider-supported condition demonstrated the most consistent pattern of favorable outcomes, suggesting that structured onboarding and perceived clinical endorsement reduce uncertainty and increase confidence in use. In contrast, reliance on self-directed materials alone appeared insufficient to support early engagement and may introduce cognitive burden during initial use.

Notably, the disconnect observed between improvements in attitudinal acceptance and more modest shifts in behavioral intention suggests that early exposure may not be sufficient to drive adoption. This reinforces established technology acceptance theories, which suggest that early perceptions may shift quickly, while sustained intention and behavior require continued experience and reinforcement [62]. However, in contrast to prior VR studies that emphasize PU as the primary driver of adoption, our findings suggest that ease of use and structured onboarding may play a more prominent role during initial exposure, particularly for first-time users [79].

Findings in this study align closely with insights from other DTx studies where making DTx feel “ordinary” and familiar accelerated patient adoption and acceptance [42]. In this context, provider involvement and structured implementation strategies may serve as mechanisms that help bridge the gap between initial exposure and sustained engagement.

Role of Provider Support and Human-Centered Implementation

Findings consistently demonstrate that provider involvement plays a critical role in shaping early adoption of VRx. Participants receiving provider-supported onboarding showed greater improvements across acceptability, usability, and adherence outcomes, suggesting that structured human interaction enhances both perceived legitimacy and user confidence. This aligns with broader DTx literature identifying provider endorsement as a critical driver of patient uptake [30,42,80-82]. Our results extend this work by demonstrating that provider involvement influences not only adoption intent but also early usability, engagement, and adherence behaviors during initial exposure.

Provider-led activities, including safety screening, expectation setting, and device setup, appeared to reduce uncertainty and reinforce the therapeutic credibility of VRx. Participants reported greater willingness to engage when the intervention was introduced by a provider, consistent with prior research that provider endorsement enhances perceptions of safety, credibility, and ease of use [28,42,83].

The brief structured nature of the mock consultation suggests that effective implementation may not require a substantial time burden, but rather targeted and well-designed onboarding as part of a familiarization stage. This extends prior research in VR, which found that a familiarization phase before VR exposure reduced surprise effects [84,85]. Future research should evaluate scalable models, including hybrid approaches that combine brief human interaction with digital guidance.

In contrast, reliance on static materials alone was insufficient to support initial engagement. Participants in the self-directed condition relied solely on their own interpretation or ability to digest the information in a timely fashion. The lack of real-time support or personalized guidance may have led to uncertainty or cognitive overload, preventing participants from fully grasping the benefits and ease of using the VR independently. These findings are consistent with other implementation research that found the complexity of the implementation process and the amount of work involved to learn the technologies as barriers for successful adoption [25,86,87]. Future research should evaluate how different types of training delivery (eg, interactive onboarding, human-led instruction, or personalized digital guidance) influence the acceptability of VRx.

Taken together, these findings suggest that structured, human-supported onboarding may reduce early apprehension and improve engagement with VR therapeutics, particularly among first-time users. Provider endorsement remains a powerful lever for uptake, beyond reimbursement or access [88,89]. Notably, the provider interaction observed in this study was brief and structured, suggesting that effective

XSL•FO
RenderX

implementation may not require a significant time burden but rather targeted, well-designed interactions. Future research should evaluate scalable models of clinician-supported onboarding, including hybrid approaches that combine brief human interaction with digital guidance.

Usability Insights

Item-level usability analysis provided granular insights on important barriers to sustained adoption. Participants reported strong intention to use, PEU, and a low perception of system complexity. However, lower confidence, learnability, and a high perceived need for assistance indicate that early usability challenges persist even with implementation support [90].

Compared to other single-exposure VR studies reporting higher usability scores, the slightly lower usability observed in this study may reflect the broader scope of the user experience assessed, from initial device interaction through full session completion, rather than isolated in-app tasks [91]. This more naturalistic approach likely introduced variability reflective of real-world use conditions.

Fidelity and Adherence

A notable qualitative finding was the distortion of time perception during VRx use. Several participants reported that the 20-minute session felt substantially shorter than its actual duration, suggesting a high degree of immersion and cognitive absorption. This aligns with prior research demonstrating the ability of VR to create immersive, attention-capturing experiences that support therapeutic mechanisms such as distraction, presence, and engagement [6,8,92,93]. The relationship between perceived time, immersion, prescription, and clinical outcomes in VRx should be further examined in future studies.

This finding also has important implications for fidelity, as users may be more likely to complete prescribed session durations when the experience feels shorter and more engaging. Related studies on VR use show that the level of immersion may support positive psychological responses, including increased intrinsic motivation, improved mood, and enhanced self-efficacy, which are known contributors to sustained engagement and adherence in rehabilitation and behavioral interventions over time [7,94,95]. Adherence improvements correlate with better clinical outcomes in many DTx studies [30,82]. Further research is needed to show the longitudinal impact of immersion on fidelity and adherence.

Preventing Cybersickness

Cybersickness was minimal across participants and did not appear to meaningfully impact engagement or completion of the VRx session. While a small number of participants reported mild symptoms (eg, slight dizziness or discomfort), these were transient and did not lead to early discontinuation [96-98]. The low incidence of cybersickness in this study may also reflect awareness of the side effect from providing the cybersickness questionnaire prior to use. Cybersickness remains an important consideration for broader implementation, particularly in patient populations who may be more susceptible. Future research should evaluate tolerability over repeated exposures and across diverse clinical populations, as well as identify design and implementation strategies that further mitigate risk.

Scalability and Resource Considerations

The 3 implementation strategies evaluated in this study represent different resource models with important implications for scalability. While provider-supported use produced the strongest outcomes, it introduces additional resource requirements that must be considered in real-world deployment.

The provider interaction in this study was brief (~5 minutes) and delivered by trained personnel, suggesting that effective support may be achievable without substantial burden. In many systems, physicians and psychotherapists act as the main prescribers and gatekeepers, deciding whether DTx are integrated into the treatment pathway [40,43,99]. There is an opportunity for future research to explore alternative delivery models, including nurse-led onboarding, digital-first guidance, or hybrid approaches combining asynchronous support with targeted human interaction.

Participants also identified cost and access as key barriers. Concerns about the affordability of VR hardware and uncertainty around insurance coverage were consistently raised. Suggested solutions included shared-device models, such as clinic-based or community access points, and integration into existing care pathways to reduce both cost and access barriers. Research supports shared-device models, shared digital access, and community or mobile access points as practical, relatively low-cost ways to reduce cost and access barriers, especially when tightly integrated into existing care pathways and supported by appropriate policy and management infrastructure [100-104].

Notably, previous research examining key barriers to DTx adoption showed that improving access to prescriptions and removing cost barriers did not accelerate adoption [34,40,43]. Further investigation found that provider endorsement of DTx significantly impacted patient adoption [43,47]. Adoption of DTx is shaped by intertwined patient, provider, organizational, and regulatory factors [28,42,80,105,106]. Lowering cost or easing prescription alone is insufficient; provider knowledge, trust, workflow fit, and institutional support are central levers for accelerating patient uptake.

While a formal cost analysis was beyond the scope of this study, the relative resource differences across conditions highlight the need for future work evaluating the cost-effectiveness of implementation strategies. Identifying the minimum effective level of provider involvement, as well as opportunities for digital or hybrid onboarding, will be critical for developing scalable models of VRx delivery.

Provider Needs for Implementation

Although provider training was not a primary focus of this study, trained study personnel acting as “mock providers” were able to deliver a brief, structured onboarding interaction that improved perceived safety, usability, and engagement. These materials were designed to be efficient and minimize time burden, addressing a known barrier for provider adoption: limited time during patient encounters [89]. This suggests that

XSL•FO
RenderX

a short "introduce and demonstrate" interaction may be sufficient to support initial adoption without adding substantial time burden to clinical workflows.

Research related to prescribing methods for DTx is limited; methods for prescribing and ensuring patient safety are substantial gaps in how DTx and VRx are delivered today [88,89,107]. There is a substantial opportunity for this research to be expanded on to design simplified tools for patient screening and streamlining the clinical workflow to mitigate adoption barriers experienced by providers. Finally, lightweight screening, dose guidance, and simplified demo devices may further reduce provider friction and support safe, repeatable prescribing.

Theory Integration

This study integrates complementary frameworks to demonstrate that VRx adoption is shaped by the interaction of user perceptions (TAM), usability experience (SUS), and implementation context (BEAR). While brief exposure improved PU and PEU, these shifts did not translate into strong acceptability, indicating that perception alone is insufficient for adoption [61].

Usability findings further highlight a gap between PEU and user confidence, suggesting that initial support may be required even for well-designed systems [64]. The BEAR contextualizes these findings by demonstrating that implementation strategies, particularly through provider involvement, reduce uncertainty and reinforce legitimacy [60]. This reinforces the role of implementation strategies as active drivers of engagement, not simply delivery mechanisms. These findings suggest that successful VRx deployment requires alignment across perception, usability, and implementation strategy to support both initial engagement and sustained use.

Implications

Taken together, these findings suggest that implementation strategy is a critical determinant of early adoption of VR-based DTx. By isolating onboarding and provider support components, this study extends prior VRx research that has largely focused on clinical efficacy to address how these interventions are introduced and supported. The results highlight the importance of integrating structured onboarding and provider touchpoints to enhance usability, engagement, and adherence. These findings have practical implications for the design and deployment of scalable VRx solutions in real-world settings.

Limitations

This study has several limitations that affect the interpretation and generalizability of the findings. First, as a pilot study with approximately 10 - 11 participants per condition, it was not powered to detect between-group differences. As a result, both null and statistically significant findings should be interpreted cautiously, and effect size estimates should be considered exploratory rather than definitive.

Second, multiple comparisons were conducted across TAM and SUS constructs, and exploratory subgroup analyses were performed without correction for multiplicity. While consistent with the exploratory nature of pilot studies, this increases the risk of type I error [69,77]. Accordingly, findings should be interpreted based on overall patterns rather than individual comparisons.

Third, the study was conducted in a controlled, nonclinical sample of healthy adults, which may limit generalizability to intended patient populations. Engagement, motivation, and usability may differ in clinical populations, where symptom burden and provider recommendation play a more central role. There was an imbalance in sex distribution across conditions, with a higher proportion of female participants in condition 3, which also demonstrated the most favorable outcomes. Prior research suggests differences in technology acceptance by sex, and this imbalance may represent a potential confounding variable. Simple randomization was used rather than stratified or block randomization. While baseline characteristics including technology comfort level were monitored across conditions, the absence of formal stratification may have contributed to residual imbalance, including a higher proportion of female participants in condition 3 and uneven distribution of technology comfort level across conditions. Allocation concealment from study personnel was also not implemented, and full blinding of participants and study personnel was not feasible, given the visible differences across study arms. Future adequately powered studies should incorporate formal stratification, allocation concealment, and, where feasible, blinded outcome assessment.

Fourth, provider involvement was simulated and may not reflect real-world prescribing, symptom-driven motivation, or clinical care pathways. In clinical populations, the presence of a diagnosed condition and provider recommendation may further influence engagement, adherence, and perceived value of VR therapeutics.

Fifth, differences between the laboratory and community-based study settings may have influenced usability and engagement. Future studies should evaluate these strategies in fully naturalistic, real-world settings to better understand their impact on adoption and sustained use.

Finally, the study evaluated a single, short-term exposure to VRx, precluding assessment of sustained engagement, long-term adherence, and clinical outcomes. Future research should examine longitudinal use in clinical populations and real-world settings to better understand adoption, adherence, and implementation at scale.

Conclusions

This study provides early evidence that the implementation strategy plays a critical role in shaping the adoption of VR-based DTx. This work is innovative in experimentally isolating implementation components, such as onboarding support and provider involvement, rather than evaluating the therapeutic content alone, which distinguishes it from prior VRx studies that have primarily focused on clinical efficacy. By directly comparing unguided, self-directed, and clinician-supported approaches, this study offers new insight into how implementation design influences early user adoption.

Findings demonstrate that while brief exposure to VRx improved user perceptions across all conditions, structured onboarding and provider involvement produced more consistent

XSL•FO
RenderX

improvements in usability, engagement, and adherence. Notably, even brief provider interaction appeared to increase perceived legitimacy, reduce uncertainty, and support early engagement, suggesting a potential pathway for future scalable implementation through low-burden, hybrid onboarding models, warranting validation in larger studies.

These results contribute to the growing field of DTx by emphasizing that successful deployment depends not only on the technology itself but also on how it is introduced and supported. In real-world settings, integrating lightweight provider touchpoints and structured onboarding may improve uptake and sustained use of VRx interventions. Future research should evaluate these strategies in longitudinal and clinical populations to determine their impact on sustained adherence and clinical outcomes.

Acknowledgments

The authors thank all study participants for their time and engagement. This work was conducted as part of the author's doctoral studies with institutional support from the Carolina Health Informatics Program and the Department of Healthcare Engineering at the University of North Carolina at Chapel Hill. No funding agency had a role in the study design; data collection, analysis, or interpretation; manuscript preparation; or the decision to submit the manuscript for publication. During manuscript preparation, ChatGPT (OpenAI) was used solely for grammar, spelling, and clarity editing of author-written text. All outputs were reviewed and edited by the author, who takes full responsibility for the content. No generative artificial intelligence was used for study design, data collection, data analysis, interpretation, reference selection, or figure creation, and no confidential or identifiable data were provided to the tool.

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

AZ, LM, AWK, RK, GM, CM, and LS conceived the study. The study design was a collaboration between all authors. The protocol was written by AZ with input from LM, RK, GM, AWK, CM, and LS. AZ, JF, and CP executed the study and wrote the manuscript. All authors approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preintervention screening instruments and consent forms.
[PDF File, 98 KB - xr_v3i1e90626_app1.pdf]

Multimedia Appendix 2
Description and rationale for selection of OpenBrush.
[PDF File, 341 KB - xr_v3i1e90626_app2.pdf]

Multimedia Appendix 3
Complete implementation package documentation.
[PDF File, 385 KB - xr_v3i1e90626_app3.pdf]

Multimedia Appendix 4
Pre- and postintervention questionnaires.
[PDF File, 211 KB - xr_v3i1e90626_app4.pdf]

Multimedia Appendix 5
Structured participant observation guide and checklist.
[PDF File, 354 KB - xr_v3i1e90626_app5.pdf]

Multimedia Appendix 6

XSL•FO
RenderX

Statistical methods and exploratory analyses.
[PDF File, 258 KB - xr_v3i1e90626_app6.pdf]

Checklist 1
CONSORT 2010 extension for randomized pilot and feasibility trials.
[DOCX File, 29 KB - xr_v3i1e90626_app7.docx]

Checklist 2
CONSORT-eHEALTH checklist (V1.6).
[DOCX File, 25 KB - xr_v3i1e90626_app8.docx]

References

1. Cancer pain. International Association for the Study of Pain (IASP). URL: https://www.iasp-pain.org/advocacy/global-year/cancer-pain/ [accessed 2022-09-12]
2. Preventing pain: an introduction. International Association for the Study of Pain (IASP). URL: https://www.iasp-pain.org/resources/fact-sheets/preventing-pain-an-introduction/ [accessed 2025-03-31]
3. Rikard SM, Strahan AE, Schmit KM, Guy GP Jr. Chronic pain among adults—United States, 2019-2021. MMWR Morb Mortal Wkly Rep 2023 Apr 14;72(15):379-385. [doi: 10.15585/mmwr.mm7215a1] [Medline: 37053114]
4. Nahin RL, Feinberg T, Kapos FP, Terman GW. Estimated rates of incident and persistent chronic pain among US adults, 2019-2020. JAMA Netw Open 2023 May 1;6(5):e2313563. [doi: 10.1001/jamanetworkopen.2023.13563] [Medline: 37191961]
5. NIH study finds high rates of persistent chronic pain among U.S. adults. National Center for Complementary and Integrative Health. URL: https://www.nccih.nih.gov/news/press-releases/nih-study-finds-high-rates-of-persistent-chronic-pain-among-us-adults [accessed 2025-03-31]
6. Spiegel B. VRx: How Virtual Therapeutics Will Revolutionize Medicine: Basic Books; 2020.
7. Garcia LM, Birckhead 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] [Medline: 33484240]
8. Micheluzzi V, Casu G, Burrai F, et al. The experience of immersive virtual reality in patients with heart failure during cardiac rehabilitation: a qualitative study. Front Med 2025 May 14;12. [doi: 10.3389/fmed.2025.1578399]
9. Ioannou A, Papastavrou E, Avraamides MN, Charalambous A. Virtual reality and symptoms management of anxiety, depression, fatigue, and pain: a systematic review. SAGE Open Nurs 2020;6:2377960820936163. [doi: 10.1177/2377960820936163] [Medline: 33415290]
10. Felnhofer A, Pfannerstill F, Gänsler L, et al. Barriers to adopting therapeutic virtual reality: the perspective of clinical psychologists and psychotherapists. Front Psychiatry 2025;16:1549090. [doi: 10.3389/fpsyt.2025.1549090] [Medline: 40171310]
11. McGirt MJ, Holland CM, Farber SH, et al. Remote cognitive behavioral therapy utilizing an in-home virtual reality toolkit (Vx Therapy) reduces pain, anxiety, and depression in patients with chronic cervical and lumbar spondylytic pain: a potential alternative to opioids in multimodal pain management. North Am Spine Soc J (NASSJ) 2023 Dec;16:100287. [doi: 10.1016/j.xnsj.2023.100287]
12. Torous J, Linardon J, Goldberg SB, et al. The evolving field of digital mental health: current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. World Psychiatry 2025 Jun;24(2):156-174. [doi: 10.1002/wps.21299] [Medline: 40371757]
13. Chen Y, Chen X, Li L, Li Y, Yan Q, Hu X. The efficacy of virtual reality-based interventions on pain, anxiety, depression, and quality of life among patients with cancer: a meta-analysis of randomized controlled trials. Cancer Nurs 2024 Nov 29. [doi: 10.1097/NCC.0000000000001430] [Medline: 39625880]
14. Jingili N, Oyelere SS, Nyström MBT, Anyshchenko L. A systematic review on the efficacy of virtual reality and gamification interventions for managing anxiety and depression. Front Digit Health 2023;5:1239435. [doi: 10.3389/fdgth.2023.1239435] [Medline: 38026832]
15. Su JJ, Chan CK, Batalik L, Chung WC, Lei C, Kwan RYC. Immersive virtual reality-based intervention for psychological wellbeing among older adults: a systematic review and meta-analysis. PLoS Digit Health 2026 Jan;5(1):e0001110. [doi: 10.1371/journal.pdig.0001110] [Medline: 41538389]
16. Baker NA, Polhemus AH, Baird JM, Kenney M. Embodied fully immersive virtual reality as a therapeutic modality to treat chronic pain: a scoping review. Virtual Worlds 2026 Mar;5(1):3. [doi: 10.3390/virtualworlds5010003]
17. Park MJ, Kim DJ, Lee U, Na EJ, Jeon HJ. A literature overview of virtual reality (VR) in treatment of psychiatric disorders: recent advances and limitations. Front Psychiatry 2019;10:505. [doi: 10.3389/fpsyt.2019.00505] [Medline: 31379623]
18. Baghaei N, Chitale V, Hlasnik A, Stemmet L, Liang HN, Porter R. Virtual reality for supporting the treatment of depression and anxiety: scoping review. JMIR Ment Health 2021 Sep 23;8(9):e29681. [doi: 10.2196/29681] [Medline: 34554097]

XSL•FO
RenderX

19. Wu Y, Wang N, Zhang H, Sun X, Wang Y, Zhang Y. Effectiveness of virtual reality in symptom management of cancer patients: a systematic review and meta-analysis. J Pain Symptom Manage 2023 May;65(5):e467-e482. [doi: 10.1016/j.jpainsymman.2023.01.023] [Medline: 36758907]
20. Lee D, Kim YJ. Current trends and future perspectives on digital therapeutics. J Korean Med Assoc 2025 Apr;68(4):235-239. [doi: 10.5124/jkma.25.0033]
21. Bin S, Alrashdi DH, Whitehead T, Riches S, Drini E. Mental health professionals' attitudes towards virtual reality therapies: a systematic review. J Technol Behav Sci 2025 Mar 10;11(1):146-169. [doi: 10.1007/s41347-025-00499-x]
22. Eijlers R, Utens E, Staals LM, et al. Systematic review and meta-analysis of virtual reality in pediatrics: effects on pain and anxiety. Anesth Analg 2019 Nov;129(5):1344-1353. [doi: 10.1213/ANE.0000000000004165] [Medline: 31136330]
23. Ibrahim FE, Elsayed NAM, Zayed HH. A survey on the effectiveness of virtual reality-based therapy and pain management. IJACSA 2021;12(7):15. [doi: 10.14569/IJACSA.2021.0120769]
24. Fodor LA, Cote CD, Cuijpers P, Szamoskozi , David D, Cristea IA. The effectiveness of virtual reality based interventions for symptoms of anxiety and depression: a meta-analysis. Sci Rep 2018 Jul 9;8(1):10323. [doi: 10.1038/s41598-018-28113-6] [Medline: 29985400]
25. Ferrer Costa J, Fernandez Brusco A, Torrecillas Camacho C, et al. Integrating extended reality into primary care chronic pain programs via the REDOCVR intervention: real-world implementation feasibility and usability study. JMIR XR Spatial Comput 2025 Oct 31;2:e82858. [doi: 10.2196/82858]
26. Lau CKY, Saad A, Camara B, Rahman D, Bolea-Alamanac B. Acceptability of digital mental health interventions for depression and anxiety: systematic review. J Med Internet Res 2024 Oct 28;26:e52609. [doi: 10.2196/52609] [Medline: 39466300]
27. Wong KP, Tse MMY, Qin J. Effectiveness of virtual reality-based interventions for managing chronic pain on pain reduction, anxiety, depression and mood: a systematic review. Healthcare (Basel) 2022 Oct 17;10(10):2047. [doi: 10.3390/healthcare10102047] [Medline: 36292493]
28. Prodan A, Deimel L, Ahlqvist J, et al. Success factors for scaling up the adoption of digital therapeutics towards the realization of P5 medicine. Front Med 2022;9. [doi: 10.3389/fmed.2022.854665]
29. Cox D. From sleep apps to chatbots: how a new generation of health tech is widening access to treatment. The Guardian. 2025 Jan 3. URL: https://kemindogroup.com/news/the-dawn-of-digital-therapeutics-2025-jan-03-from-sleep-apps-to-chatbots-how-a-new-generation-of-health-tech-is-widening-access-to-treatment [accessed 2025-02-27]
30. Riley N. Clinical evidence must be at the heart of DTx solutions. Here's why. The Sidebar. 2022 Jun 28. URL: https://www.sidekickhealth.com/news/clinical-evidence-at-the-heart-of-dtx [accessed 2025-02-27]
31. Giebel GD, Speckemeier C, Abels C, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023 May 12;25:e43808. [doi: 10.2196/43808] [Medline: 37171838]
32. Williams MG, Stott R, Bromwich N, Oblak SK, Espie CA, Rose JB. Determinants of and barriers to adoption of digital therapeutics for mental health at scale in the NHS. BMJ Innov 2020 Jul;6(3):92-98. [doi: 10.1136/bmjinnov-2019-000384]
33. Montesano A, Seinfeld S. Virtual reality in psychotherapy: a three-dimensional framework to navigate immersive clinical applications. J Clin Psychol 2025 Sep;81(9):783-789. [doi: 10.1002/jclp.70004] [Medline: 40483725]
34. Fassbender A, Donde S, Silva M, et al. Adoption of digital therapeutics in Europe. Ther Clin Risk Manag 2024;20(20):939-954. [doi: 10.2147/TCRM.S489873] [Medline: 39741688]
35. Wray TB, Kemp JJ, Larsen MA. Virtual reality (VR) treatments for anxiety disorders are unambiguously successful, so why are so few therapists using it? Barriers to adoption and potential solutions. Cogn Behav Ther 2023 Nov;52(6):603-624. [doi: 10.1080/16506073.2023.2229017] [Medline: 37376984]
36. Schwartz DG, Spitzer S, Khalemsky M, et al. Apps don't work for patients who don't use them: towards frameworks for digital therapeutics adherence. Health Policy Technol 2024 Jun;13(2):100848. [doi: 10.1016/j.hlpt.2024.100848]
37. van Kessel R, Roman-Urrestarazu A, Anderson M, et al. Mapping factors that affect the uptake of digital therapeutics within health systems: scoping review. J Med Internet Res 2023 Jul 25;25:e48000. [doi: 10.2196/48000] [Medline: 37490322]
38. Mumtaz H, Riaz MH, Wajid H, et al. Current challenges and potential solutions to the use of digital health technologies in evidence generation: a narrative review. Front Digit Health 2023;5:1203945. [doi: 10.3389/fdgth.2023.1203945] [Medline: 37840685]
39. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]
40. Kendziorra J, Seerig KH, Winkler TJ, Gewald H. From awareness to integration: a qualitative interview study on the impact of digital therapeutics on physicians' practices in Germany. BMC Health Serv Res 2025 Apr 18;25(1):568. [doi: 10.1186/s12913-025-12656-2] [Medline: 40247268]
41. Whitelaw S, Pellegrini DM, Mamas MA, Cowie M, Van Spall HGC. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. Eur Heart J Digit Health 2021 Mar;2(1):62-74. [doi: 10.1093/ehjdh/ztab005] [Medline: 34048508]

XSL•FO
RenderX

42. Carrera A, Zoccarato F, Mazzeo M, et al. What drives patients' acceptance of Digital Therapeutics? Establishing a new framework to measure the interplay between rational and institutional factors. BMC Health Serv Res 2023 Feb 10;23(1):145. [doi: 10.1186/s12913-023-09090-7] [Medline: 36765410]
43. Dahlhausen F, Zinner M, Bieske L, Ehlers JP, Boehme P, Fehring L. There's an app for that, but nobody's using it: insights on improving patient access and adherence to digital therapeutics in Germany. Digit Health 2022;8:20552076221104672. [doi: 10.1177/20552076221104672] [Medline: 35811758]
44. Pong C, Tseng R, Tham YC, Lum E. Current implementation of digital health in chronic disease management: scoping review. J Med Internet Res 2024 Dec 12;26:e53576. [doi: 10.2196/53576] [Medline: 39666972]
45. de la Vega R, Bartels SL, Wicksell RW. Implementation frameworks guiding digital self-management intervention in chronic pain: a scoping review. Eur J Pain 2024 Sep;28(8):1257-1275. [doi: 10.1002/ejp.2262] [Medline: 38483014]
46. Tennant M, McGillivray J, Youssef GJ, McCarthy MC, Clark TJ. Feasibility, acceptability, and clinical implementation of an immersive virtual reality intervention to address psychological well-being in children and adolescents with cancer. J Pediatr Oncol Nurs 2020;37(4):265-277. [doi: 10.1177/1043454220917859] [Medline: 32536320]
47. Chung OS, Robinson T, Johnson AM, et al. Implementation of therapeutic virtual reality into psychiatric care: clinicians' and service managers' perspectives. Front Psychiatry 2021;12:791123. [doi: 10.3389/fpsyt.2021.791123] [Medline: 35058823]
48. Kouijzer M, Kip H, Bouman YHA, Kelders SM. Implementation of virtual reality in healthcare: a scoping review on the implementation process of virtual reality in various healthcare settings. Implement Sci Commun 2023 Jun 16;4(1):67. [doi: 10.1186/s43058-023-00442-2] [Medline: 37328858]
49. Pearce L, Costa N, Sherrington C, Hassett L. Implementation of digital health interventions in rehabilitation: a scoping review. Clin Rehabil 2023 Nov;37(11):1533-1551. [doi: 10.1177/02692155231172299] [Medline: 37132030]
50. Lakhan SE. Prescription digital therapeutics research across clinical, engagement, regulatory, and implementation domains: a bibliometric and thematic study. Cureus 2025 May;17(5):e84528. [doi: 10.7759/cureus.84528] [Medline: 40400654]
51. Phelan I, Carrion-Plaza A, Furness PJ, Dimitri P. Home-based immersive virtual reality physical rehabilitation in paediatric patients for upper limb motor impairment: a feasibility study. Virtual Real 2023;27(4):3505-3520. [doi: 10.1007/s10055-023-00747-6] [Medline: 36686613]
52. Maheta B, Kraft A, Interrante N, 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 Feb 26;27:e54452. [doi: 10.2196/54452] [Medline: 40009834]
53. Garrett B, Taverner T, Gromala D, Tao G, Cordingley E, Sun C. Virtual reality clinical research: promises and challenges. JMIR Serious Games 2018 Oct 17;6(4):e10839. [doi: 10.2196/10839] [Medline: 30333096]
54. Ummels D, Cnockaert E, Timmers I, den Hollander M, Smeets R. Use of virtual reality in interdisciplinary multimodal pain treatment with insights from health care professionals and patients: action research study. JMIR Rehabil Assist Technol 2023 Nov 10;10:e47541. [doi: 10.2196/47541] [Medline: 37948109]
55. Elkefi S, Choudhury A. Role of virtual reality in improving home cancer care: a systematic literature review. Electronics (Basel) 2025 Jan 20;14(2):385. [doi: 10.3390/electronics14020385]
56. Phelan S, Peruvemba S, Levinson D, et al. Feasibility of a virtual reality-based approach to improve behavioral weight management outcomes. Pilot Feasibility Stud 2021 Dec;7(1):129. [doi: 10.1186/s40814-021-00865-5]
57. Denecke K, May R, Borycki EM, Kushniruk AW. Digital health as an enabler for hospital@home: a rising trend or just a vision? Front Public Health 2023;11:1137798. [doi: 10.3389/fpubh.2023.1137798] [Medline: 36875371]
58. Sarkar U, Lee JE, Nguyen KH, Lisker S, Lyles CR. Barriers and facilitators to the implementation of virtual reality as a pain management modality in academic, community, and safety-net settings: qualitative analysis. J Med Internet Res 2021 Sep 22;23(9):e26623. [doi: 10.2196/26623] [Medline: 34550074]
59. Glegg SMN, Levac DE. Barriers, facilitators and interventions to support virtual reality implementation in rehabilitation: a scoping review. PM R 2018 Nov;10(11):1237-1251. [doi: 10.1016/j.pmrj.2018.07.004] [Medline: 30503231]
60. Camacho J, Zanoletti-Mannello M, Landis-Lewis Z, Kane-Gill SL, Boyce RD. A conceptual framework to study the implementation of clinical decision support systems (BEAR): literature review and concept mapping. J Med Internet Res 2020 Aug 6;22(8):e18388. [doi: 10.2196/18388] [Medline: 32759098]
61. 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]
62. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
63. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
64. System Usability Scale (SUS). Usability Body of Knowledge. 2013. URL: https://www.usabilitybok.org/sus [accessed 2026-06-08]
65. Birckhead B, Khalil C, Liu X, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973. [doi: 10.2196/11973] [Medline: 30702436]

XSL•FO
RenderX

66. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239. [doi: 10.1136/bmj.i5239] [Medline: 27777223]
67. 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] [Medline: 22209829]
68. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022 Jan 1;60(1):95-103. [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
69. 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] [Medline: 21035130]
70. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
71. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
72. Golding JF. Predicting individual differences in motion sickness susceptibility by questionnaire. Pers Individ Dif 2006 Jul;41(2):237-248. [doi: 10.1016/j.paid.2006.01.012]
73. Kourtesis P, Linnell J, Amir R, Argelaguet F, MacPherson SE. Cybersickness in Virtual Reality Questionnaire (CSQ-VR): a validation and comparison against SSQ and VRSQ. Virtual Worlds 2023;2(1):16-35. [doi: 10.3390/virtualworlds2010002]
74. Lewis J, Sauro J. Can I leave this one out? The effect of dropping an item from the SUS. J User Exp 2017 Nov 28(1):38-46. [doi: 10.5555/3173069.3173073]
75. Lewis J, Sauro J. Item benchmarks for the System Usability Scale. J User Exp 2018 May 25(3):158-167. [doi: 10.5555/3294033.3294037]
76. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
77. Sullivan GM, Feinn RS. Facts and fictions about handling multiple comparisons. J Grad Med Educ 2021 Aug;13(4):457-460. [doi: 10.4300/JGME-D-21-00599.1] [Medline: 34434505]
78. 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] [Medline: 40228833]
79. Bertrand M, Bouchard S. Applying the technology acceptance model to VR with people who are favorable to its use. Journal of CyberTherapy & Rehabilitation 2008 Jan 1;1:200-210.
80. Armeni P, Polat I, De Rossi LM, Diaferia L, Meregalli S, Gatti A. Exploring the potential of digital therapeutics: an assessment of progress and promise. Digit Health 2024;10:20552076241277441. [doi: 10.1177/20552076241277441] [Medline: 39291152]
81. Rohaj A, Bulaj G. Digital therapeutics (DTx) expand multimodal treatment options for chronic low back pain: the nexus of precision medicine, patient education, and public health. Healthcare (Basel) 2023 May 18;11(10):1469. [doi: 10.3390/healthcare11101469] [Medline: 37239755]
82. Roberts MJ, Wilson WJ, Recugnat M, Bardy F. Barriers to adoption of digital therapeutics (DTx) into audiology clinical practice: acknowledging the challenges, adapting to the future. Semin Hear 2023 Aug;44(3):319-327. [doi: 10.1055/s-0043-1769623] [Medline: 37484989]
83. Kim HS. Apprehensions about excessive belief in digital therapeutics: points of concern excluding merits. J Korean Med Sci 2020 Nov 23;35(45):e373. [doi: 10.3346/jkms.2020.35.e373] [Medline: 33230984]
84. Buche H, Michel A, Piccoli C, Blanc N. Contemplating or acting? Which immersive modes should be favored in virtual reality during physiotherapy for breast cancer rehabilitation. Front Psychol 2021;12:631186. [doi: 10.3389/fpsyg.2021.631186] [Medline: 33897539]
85. Chirico A, Maiorano P, Indovina P, et al. Virtual reality and music therapy as distraction interventions to alleviate anxiety and improve mood states in breast cancer patients during chemotherapy. J Cell Physiol 2020 Jun;235(6):5353-5362. [doi: 10.1002/jcp.29422] [Medline: 31957873]
86. Varsi C, Ekstedt M, Gammon D, Ruland CM. Using the consolidated framework for implementation research to identify barriers and facilitators for the implementation of an internet-based patient-provider communication service in five settings: a qualitative study. J Med Internet Res 2015 Nov 18;17(11):e262. [doi: 10.2196/jmir.5091] [Medline: 26582138]
87. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implementation Sci 2016 Dec;11(1). [doi: 10.1186/s13012-016-0510-7]
88. Ruetsch C, Velligan D, Robinson D, et al. Psychiatric prescriber attitudes, experiences, and proclivities toward digital medicine and how they influence adoption of digital medicine platforms. CNS Spectr 2021 Apr;26(2):144-145. [doi: 10.1017/S109285292000228X]
89. Mogk J, Idu AE, Bobb JF, et al. Prescription digital therapeutics for substance use disorder in primary care: mixed methods evaluation of a pilot implementation study. JMIR Form Res 2024 Sep 2;8:e59088. [doi: 10.2196/59088] [Medline: 39222348]
90. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry: CRC Press; 1995. [doi: 10.1201/9781498710411-35]

XSL•FO
RenderX

91. Martino Cinnera A, Verna V, Marucci M, et al. Immersive virtual reality for treatment of unilateral spatial neglect via eye-tracking biofeedback: RCT protocol and usability testing. Brain Sci 2024 Mar 15;14(3):283. [doi: 10.3390/brainsci14030283] [Medline: 38539671]
92. Fusco A, Tieri G. Challenges and perspectives for clinical applications of immersive and non-immersive virtual reality. J Clin Med 2022 Aug 4;11(15):4540. [doi: 10.3390/jcm11154540] [Medline: 35956157]
93. Maddox T, Sparks CY, Oldstone L, et al. Perspective: the promise of virtual reality as an immersive therapeutic. J Med Extend Real 2024 Jan 1;1(1):13-20. [doi: 10.1089/jmxr.2023.0003]
94. Micheluzzi V, Vellone E, Iovino P. A situation-specific theory on the use of immersive virtual reality in rehabilitation for patients with disabilities. Holist Nurs Pract 2025;39(2):71-79. [doi: 10.1097/HNP.0000000000000718] [Medline: 39666877]
95. Micheluzzi V, Casu G, Sanna GD, et al. Improving adherence to rehabilitation for heart failure patients through immersive VIRTUAL reality (VIRTUAL-HF): a protocol for a randomized controlled trial. Contemp Clin Trials 2024 Mar;138:107463. [doi: 10.1016/j.cct.2024.107463] [Medline: 38302011]
96. Brown P, Powell W. Pre-exposure cybersickness assessment within a chronic pain population in virtual reality. Front Virtual Real 2021 Jun;2. [doi: 10.3389/frvir.2021.672245]
97. Kourtesis P, Papadopoulou A, Roussos P. Cybersickness in virtual reality: the role of individual differences, its effects on cognitive functions and motor skills, and intensity differences during and after immersion. Virtual Worlds 2024;3(1):62-93. [doi: 10.3390/virtualworlds3010004]
98. Eudave L, Martínez M. To VR or not to VR: assessing cybersickness in navigational tasks at different levels of immersion. Front Virtual Real 2025 Jan 29;6. [doi: 10.3389/frvir.2025.1518735]
99. Wilson RL. The right way for nurses to prescribe, administer and critique digital therapies. Contemp Nurse 2018 Jul 4;54(4-5):543-545. [doi: 10.1080/10376178.2018.1507679]
100. The integration of health disciplines and advanced technology: a sustainable approach to improving the quality of healthcare services. Power Syst Technol 2024;48(4):5966-5984. [doi: 10.52783/pst.1391]
101. Anghel I, Cioara T, Bevilacqua R, et al. New care pathways for supporting transitional care from hospitals to home using AI and personalized digital assistance. Sci Rep 2025;15(1). [doi: 10.1038/s41598-025-03332-w]
102. Dukhanin V, Wolff JL, Gleason K, et al. Users' perspectives on a demonstration to increase shared access to older adults' patient portals. BMC Health Serv Res 2025 Apr 23;25(1):586. [doi: 10.1186/s12913-025-12755-0] [Medline: 40269800]
103. Gellert GA, Borgansano D, Pickering R, Kelly SP. Shared-use mobile device management practices, capabilities and needs in healthcare delivery organizations: an international survey. J Hosp Admin 2026 Jan 12;15(1):23-37. [doi: 10.63564/jha.v15n1p23]
104. Harvey J, Powell J. Harnessing mobile devices to support the delivery of community-based clinical care: a participatory evaluation. BMC Med Inform Decis Mak 2019 Dec;19(1). [doi: 10.1186/s12911-019-0869-x]
105. Crisafulli S, Santoro E, Recchia G, Trifirò G. Digital therapeutics in perspective: from regulatory challenges to post-marketing surveillance. Front Drug Saf Regul 2022;2. [doi: 10.3389/fdsfr.2022.900946]
106. Carrera A, Lettieri E, Lietti G, Martignoni S, Sgarbossa C, Cafazzo J. Therapies go digital. What drives physicians' acceptance? PLoS One 2024;19(5):e0303302. [doi: 10.1371/journal.pone.0303302] [Medline: 38728346]
107. Digital Therapeutics: Redefining Prescription Medication in the Information Era. URL: https://www.pharmasalmanac.com/articles/digital-therapeutics-redefining-prescription-medication-in-the-information-era [accessed 2024-11-24]

Abbreviations

ATT: attitude toward technology
BEAR: Behavior and Acceptance Framework
BI: Behavioral Intention to Use
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-eHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
CSQ-VR: Cybersickness in Virtual Reality Questionnaire
DTx: digital therapeutics
IRB: institutional review board
PEU: perceived ease of use
PU: perceived usefulness
REDCap: Research Electronic Data Capture
SUS: System Usability Scale
TAM: technology acceptance model
VR: virtual reality
VRx: virtual reality therapeutics

XSL•FO
RenderX

Edited by I Steenstra, S Brini; submitted 31.Dec.2025; peer-reviewed by I Tosun, K Fuji; revised version received 06.May.2026; accepted 07.May.2026; published 30.Jun.2026.

Please cite as:
Zebrowski A, Fernandez RJ, Pinkelton C, Kitzmiller R, Kiefer AW, Stanley L, Mody G, Moore C, Mazur L
Evaluating the Effects of Clinician Prescribing and Implementation Materials on Adoption of Virtual Reality Therapeutics: Randomized Feasibility Pilot Study
JMIR XR Spatial Comput 2026;3:e90626
URL: https://xr.jmir.org/2026/1/e90626
doi:10.2196/90626

© Ashlyn Zebrowski, R Jackson Fernandez, Cameron Pinkelton, Rebecca Kitzmiller, Adam W Kiefer, Laura Stanley, Gita Mody, Carlton Moore, Lukasz Mazur. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 30.Jun.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patient Perception of the Use of the Apple Vision Pro Virtual Reality Headset in an Anesthesiology Informed Consent Process: Prospective Randomized Controlled Crossover Trial

Sabrina Chriqui[1*], BS; John Francis Ryan[2*], MD; Bao Kien Luu[1,3], MD; Minh Tran[4], MBBS; Byron Fergerson[4], MD; Alan You[5,6], MD; Christian Dameff[5,7,8], MD; Preetham John Suresh[4], MD; Jeffrey Logan Tully[9], MD

[1]School of Medicine, University of California San Diego, 9500 Gilman Dr, La Jolla, CA, United States

[2]Department of Anesthesiology and Pain Medicine, University of California Davis Health, Sacramento, CA, United States

[3]Department of Surgery, University of California San Diego Health, La Jolla, CA, United States

[4]Department of Anesthesiology, University of California San Diego Health, La Jolla, CA, United States

[5]Division of Clinical Informatics and Telehealth, Department of Emergency Medicine, University of California San Diego Health, La Jolla, CA, United States

[6]Department of Medicine, University of California San Diego Health, La Jolla, CA, United States

[7]Department of Biomedical Informatics, Department of Medicine, University of California San Diego Health, La Jolla, CA, United States

[8]Department of Computer Science and Engineering, Jacobs School of Engineering, University of California San Diego, La Jolla, CA, United States

[9]Division of Perioperative Informatics, Department of Anesthesiology, University of California San Diego Medical Center, La Jolla, CA, United States

[*]these authors contributed equally

Corresponding Author:
Sabrina Chriqui, BS
School of Medicine, University of California San Diego, 9500 Gilman Dr, La Jolla, CA, United States

Abstract

Background: Augmented reality (AR) and virtual reality (VR) have become mainstream commercial technologies, with AR integrating virtual elements into displays of reality and VR offering fully immersive computer-generated visual experiences. The Apple Vision Pro is a VR headset with an outward-facing display that depicts the user's eyes to the outside world. This headset is actively incorporated into patient care, but questions remain about the impact of such a device on the patient-physician relationship.

Objective: This study surveyed patients' responses to the use of an Apple Vision Pro VR headset by anesthesiologists during an informed consent discussion for sedation for endoscopic gastrointestinal procedures. We sought to qualify the impact, if any, of the Apple Vision Pro on patients' perceptions of connectedness with their anesthesiologist, including compassion and empathy felt from their physician.

Methods: This prospective randomized controlled crossover study involved 60 patients undergoing sedation for gastrointestinal procedures at the University of California, San Diego Health's Hillcrest Medical Center and Jacobs Medical Center. An anesthesiologist had 2 encounters with each patient prior to their procedure, 1 while the anesthesiologist wore an Apple Vision Pro and 1 without wearing any headset. After both encounters, patients completed a standardized, validated survey regarding the compassion and empathy they felt from the anesthesiologist during each encounter. They also completed a custom survey with additional questions related to their perception of technology in health care and their own personal prior experience with AR and VR.

Results: While 16.7% (10/60) of the participants owned an AR or VR headset of their own, 73% (44/60) had no prior experience with AR or VR. Participants reported consistently high ratings across all empathy and compassion measures for both the with-headset and without-headset conditions; however, mean Sinclair Compassion Questionnaire scores were statistically significantly higher for the no-headset condition (4.69, SD 0.48) when compared to the with-headset condition (4.42, SD 0.87; P=.003). Participants reported significantly reduced ability to focus on their conversation with the anesthesiologist when they were wearing the Apple Vision Pro (P<.001), with a large effect size (r=0.614). The most pronounced difference was observed with regard to patients' satisfaction with eye contact made with their physician, which decreased significantly when they wore the Apple Vision Pro (P<.001; large effect size: r=0.736).

Conclusions: This study examined the impact of physicians wearing the Apple Vision Pro during patient interactions, revealing a statistically significant reduction in patients' perceived compassion from and connectedness with their anesthesiologist. This highlights the need for caution when integrating AR or VR technology into clinical practice as it may interfere with the

XSL•FO
RenderX

patient-physician relationship. Careful evaluation of the potential effects on patients' trust in their physician is essential before adopting AR or VR in health care settings.

(JMIR XR Spatial Comput 2026;3:e95806) doi:10.2196/95806

KEYWORDS

Apple Vision Pro; augmented reality; virtual reality; compassion; headset; Sinclair Compassion Questionnaire

Introduction

The integration of augmented reality (AR) and virtual reality (VR) headsets into clinical practice is a rapidly developing objective across research and industry [1-3]. AR refers to the use of devices that display views of the real world that are modified or enhanced with superimposed computer-generated graphics [4]. VR refers to the use of devices in which the visual field is composed solely of computer-generated graphics [5]. In the last decade, technology companies have made relatively affordable head-mounted displays available to the public, thus making AR and VR increasingly accessible to the general consumer. As such, these technologies are being rapidly integrated into medical training and clinical practice. For example, AR has been used to improve surgical education and intraoperative guidance for surgeons, with multiple studies using the technology to improve operative efficiency and outcome measures in neurosurgery, orthopedics, oncology, ophthalmology, and otolaryngology [6-11]. Additionally, AR and VR have been studied in the context of patients wearing a head-mounted device for the purposes of rehabilitation, patient education, and even distraction [12]. However, the use of AR and VR in patient care by clinicians outside of the operative or procedural setting remains largely unexplored.

In June 2023, the release of the Apple Vision Pro by Apple Inc appeared to represent a paradigm shift in the intersection of AR and VR and medicine. Clinicians and researchers alike rushed to take advantage of the device's increased processing power, futuristic hand-eye tracking, and realistic "pass-through" vision [13-19]. Additionally, major health care players began exploring the use of the Apple Vision Pro in more everyday clinical settings, such as clinicians using the headset to augment documentation in the electronic health record during face-to-face patient interactions. Epic Systems Corporation introduced its Epic Spatial Computing Concept application for the Apple Vision Pro and announced a partnership with Sharp HealthCare to create the Spatial Computing Center of Excellence [20-22]. Ultimately, the Apple Vision Pro has inspired a departure from the more well-studied use of AR and VR during non–patient-facing encounters. Now, the prospect of clinicians entering face-to-face patient interactions wearing an eye-occluding head-mounted display is becoming increasingly realistic, raising an important question: How would patients perceive their physician wearing an AR or VR headset during an encounter?

A strong physician-patient relationship is the cornerstone of effective health care as this impacts patient satisfaction, adherence to treatment plans, and overall health outcomes [23]. At the core of this relationship is the physician's ability to express empathy, establish trust, and communicate effectively. Eye contact is one of the most important nonverbal factors affecting perceived physician empathy, even more so than social touch or encounter length [24,25]. As such, it is reasonable to question whether a physician wearing an AR or VR headset would hinder patients' perception of empathy and/or compassion from that physician.

Previous studies have demonstrated that certain AR headsets with clear lenses, which still allow for natural eye contact to be made between the physician and patient, do not necessarily negatively impact the patient-physician relationship. For example, one study taking place in an outpatient dermatology clinic found that patients gave physicians wearing the Google Glass smart glasses high scores on questions evaluating their relationship with their physician (9.4/10, SD 0.93) and their physician's communication with them (9.5/10, SD 1.10). A total of 87.1% of these patients were somewhat or extremely comfortable with their physician wearing the smart glasses, and 94.7% said that they were equally or more likely to trust their physician while they were wearing the Google Glass [26]. However, the impact of completely occlusive VR headsets on eye contact and patients' perception of compassion and empathy has not been studied. Although the Apple Vision Pro is technically a completely occlusive VR headset, it does feature an external display that projects a virtual representation of the wearer's eyes, which Apple calls EyeSight. This feature aims to keep users connected with those around them by simulating real-world eye contact [27].

In this prospective crossover trial, we investigated how a physician wearing the Apple Vision Pro headset with EyeSight impacted the patient-physician relationship by assessing patients' perceptions of empathy, compassion, and communication when speaking to a physician wearing the Apple Vision Pro. Understanding these perceptions is essential for ensuring that emerging technologies enhance rather than hinder patient-centered care.

Methods

Study Design and Setting

This prospective randomized controlled crossover study took place in the gastrointestinal (GI) endoscopy unit's preprocedural areas at the University of California, San Diego Health's Jacobs Medical Center and Hillcrest Medical Center between September 2024 and December 2024. Three attending anesthesiologists (authors BF, MT, and JLT) approached patients scheduled for endoscopic GI procedures and obtained informed consent for monitored anesthesia care 2 times, once without wearing a headset and once while wearing the Apple Vision Pro. A survey was administered to the patient after each encounter to evaluate

XSL•FO
RenderX

their perception of their anesthesiologist's compassion and empathy, connectedness, communication, and eye contact.

Participants

This study used a convenience sample of patients scheduled for nonemergent endoscopic GI procedures at the University of California, San Diego Health. Inclusion criteria required that patients be at least 18 years of age and able to converse, read, and write in English. Exclusion criteria were patients belonging to the following vulnerable populations: children under 18 years of age, pregnant individuals, incarcerated individuals, and those with an unhoused status. Also excluded were patients undergoing an emergent procedure or who, based on the judgment of the anesthesiologist, lacked capacity to provide informed consent to participate in the study. Participants who completed the study protocol were compensated for their time and effort with a US $5 gift card.

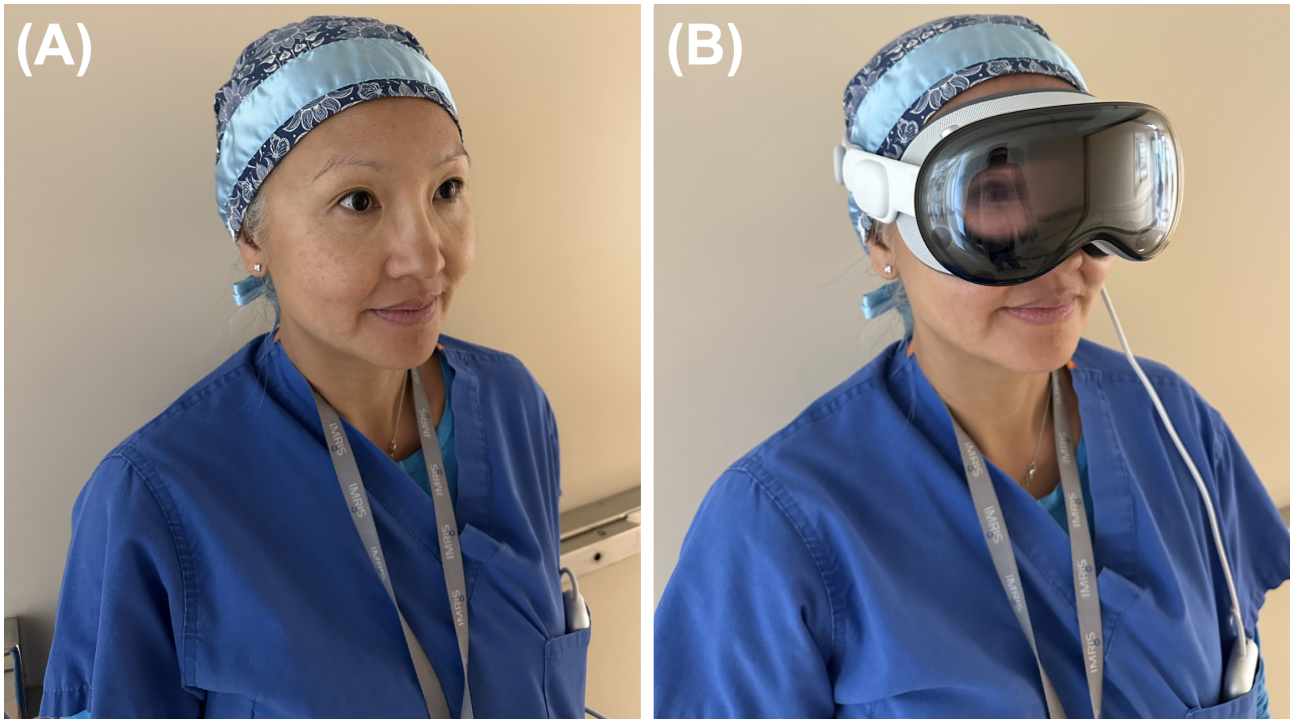
Randomization and Study Protocol

Each participant completed 2 consecutive informed consent encounters: 1 in which the physician did not wear the Apple Vision Pro headset and 1 in which the physician wore the headset. Participants were randomly assigned to 1 of 2 encounter sequences (with the headset first or without the headset first). In both sequences, an anesthesiologist conducted a standard preprocedure evaluation including medical history, physical examination, and informed consent discussion for monitored anesthesia care. Every participant then underwent a repeat informed consent discussion under the alternate condition. Both encounters were conducted sequentially, with only a 30-second interval between them; the physician briefly exited and then re-entered the patient area to either put on the headset or remove it. When the Apple Vision Pro headset was used first, it was removed during the physical examination section to allow for an adequate clinical assessment. Although sequence order was randomized to minimize order effects, sequence assignment was not recorded for individual participants. Participants were informed before the encounters that the study aimed to assess patient perceptions of physician communication with and without the use of an AR or VR headset. They were not told that the headset was expected to improve clinical care but were made aware that it is an established technology currently used in health care settings and that the study sought to understand their perspectives on its use.

Technology Implementation

Prior to patient encounters, the Apple Vision Pro headset was configured to display a realistic representation of the physician's eyes using the device's built-in face scanning technology (Figure 1). This feature was implemented to maintain a simulacrum of eye contact during patient interactions.

Figure 1. EyeSight displays a representation of the wearer's eyes on the Apple Vision Pro's outward-facing display. This figure depicts an anesthesiologist (author MT) before (A) and after (B) donning the Apple Vision Pro headset. The wearer must undergo a brief facial scan to activate EyeSight. The representation of the wearer's eyes may vary depending on the location of the viewer and the angle at which they are viewing the outward-facing display.



Data Collection

Nonidentifying demographic data were obtained from each participant, including their age, gender, and race and ethnicity, as well as basic questions regarding their prior experience with AR and VR, if any. Immediately following both encounters, patients completed the Sinclair Compassion Questionnaire (SCQ), a validated 15-item instrument created by Sinclair et al [28,29] that measures patient-perceived compassion in health care encounters and includes assessments of clinician attentiveness, empathy, and caring behaviors. Each item was rated on a 5-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5). The SCQ is scored by calculating the mean score of all items, with higher scores indicating greater

XSL•FO
RenderX

reported compassion. Patients completed the SCQ once for each encounter (headset and no headset). In addition, the patients completed a separate questionnaire (Multimedia Appendix 1) more specific to this study that assessed, from the patient perspective, patients' and physicians' ability to focus during each conversation, their satisfaction with eye contact made, how realistic EyeSight felt, and how concerned they were about their privacy during each encounter. They also responded to 5 questions assessing their general beliefs about technology in health care, whether AR or VR can improve health care delivery, and whether informed consent should be required from patients prior to using AR or VR during clinical encounters.

Statistical Analysis

Survey responses were analyzed to compare patient perceptions between the with-headset and without-headset conditions. Due to the paired nature of the crossover design and the nonnormal distribution of the Likert-scale responses, the Wilcoxon signed-rank test was used to compare SCQ scores and individual survey items between conditions [30]. Effect sizes were calculated as $Z/\sqrt{N}$, where Z is the standardized test statistic and N is the number of paired observations, and interpreted according to the criteria by Cohen [31] (small: r=0.1; medium: r=0.3; large r=0.5). All statistical analyses were conducted using R (version 4.3.1; R Foundation for Statistical Computing). All P values were 2 tailed, and a P value of less than .05 was considered significant.

Ethical Considerations

The study protocol was reviewed by the director of the Office of Institutional Review Board (IRB) Administration, IRB chair, or IRB chair's designee at the University of California, San Diego, and was determined to be exempt from institutional review (810467) under Title 45 of the Code of Federal Regulations, part 46.104(d) [32].

Results

Demographic Characteristics

This study included 60 participants with a mean age of 59 (SD 15) years, ranging from 18 to 91 years (Table 1). A total of 53.3% (n=32) of the participants were female. The racial and ethnic breakdown was 80% (n=48) White, 8.3% (n=5) Asian, 6.7% (n=4) Hispanic or Latino, and 5% (n=3) Black. Approximately 73% (n=44) of the participants had no prior experience with AR or VR, and only 16.7% (n=10) owned an AR or VR headset.

Table . Overview of participant demographics. As part of the initial survey, patients voluntarily provided their self-identified gender, age, and race and ethnicity (N=60).

Demographic	Participants, n (%)
Gender	
Women	32 (53.3)
Men	28 (46.7)
Age (y)	
29	1 (1.7)
30-39	6 (10)
40-49	12 (20)
50-59	11 (18.3)
60-69	15 (25)
70-79	13 (21.7)
80-89	1 (1.7)
≥90	1 (1.7)
Race and ethnicity	
Asian	5 (8.3)
Black or African American	3 (5)
Hispanic or Latino	4 (6.7)
Native Hawaiian or other Pacific Islander	0 (0)
White	48 (80)
Other	0 (0)

Physician Empathy and Compassion Satisfaction Scores

Participants reported consistently high ratings across all empathy and compassion measures for both with-headset and without-headset conditions; however, mean SCQ scores were statistically significantly higher for the without-headset condition (4.69, SD 0.48) when compared to the with-headset condition (4.42, SD 0.87) using the Wilcoxon signed-rank test (P=.003 [28,29]).

XSL•FO
RenderX

Paired Comparison of Custom Questions

The difference between responses to custom survey questions in the headset and no-headset conditions are shown in Table 2. Participants reported significantly reduced ability to focus on their conversation with the anesthesiologist when they were wearing the Apple Vision Pro (P<.001), with a large effect size (r=0.614). Patients also perceived that their anesthesiologist was less focused while wearing the Apple Vision Pro (P<.001), with a medium effect size (r=0.454). The most pronounced difference was observed with regard to patients' satisfaction with eye contact made with their physician, which decreased significantly when they wore the Apple Vision Pro (P<.001; large effect size: r=0.74). While statistically significant differences were observed between conditions, overall ratings across both conditions remained generally high.

Table . Summary and statistical analysis of participants' responses to survey statements comparing their interactions with the physician with and without the Apple Vision Pro headset[a].

Survey statement (abbreviated)	No headset, median rating	With headset, median rating	*P* value	Effect size	Effect classification
I was able to focus on the conversation.	Strongly agree	Agree	<.001	0.614	Large
My physician was able to focus on the conversation.	Strongly agree	Strongly agree	<.001	0.454	Medium
I was satisfied with my ability to make eye contact.	Strongly agree	Agree	<.001	0.736	Large
I was concerned about my privacy.	Strongly disagree	Strongly disagree	.002	0.416	Medium
The portrayal of the physician's eyes was realistic.	__[b]	Agree	—	—	—
If the headset was completely occlusive, I would be opposed to my physician wearing it.	—	Neutral	—	—	—

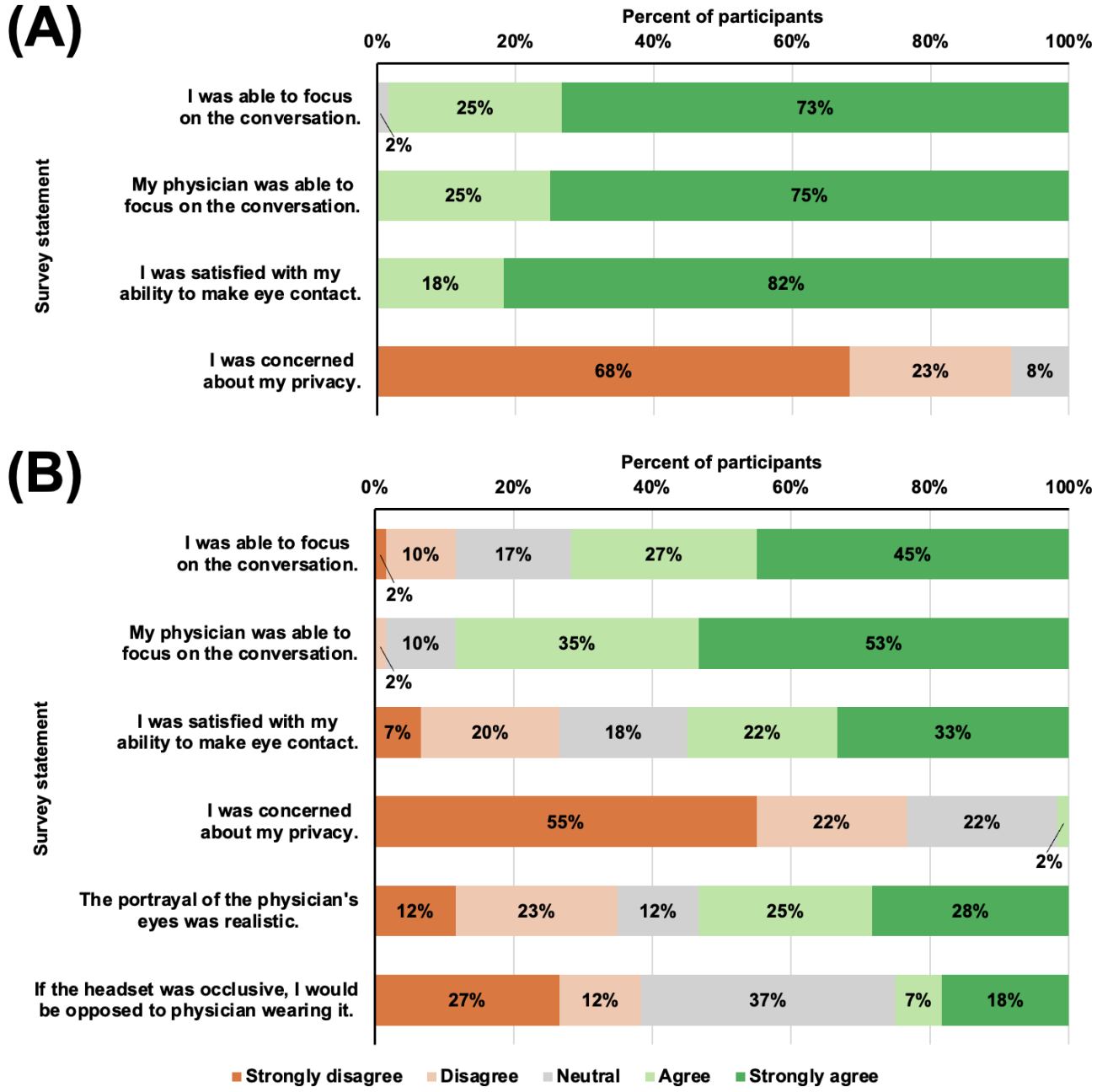
[a]Responses were recorded on a 5-point Likert scale ("strongly disagree" to "strongly agree"). Paired patient responses were evaluated for statistical significance using the Wilcoxon signed-rank test. Effect sizes were interpreted according to the criteria by Cohen [31].

[b]Asked only in the headset condition; no without-headset comparison was available, so these items are reported descriptively.

Regarding the Apple Vision Pro's EyeSight feature, 53.3% (32/60) of the patients agreed or strongly agreed that the representation of the physician's eyes on the outward-facing display was realistic, whereas 35% (21/60) disagreed or strongly disagreed. One-quarter (15/60, 25%) of patients reported that they would be opposed to their physician wearing the headset if it were completely occlusive and did not allow for simulated eye contact (Figure 2).

XSL•FO
RenderX

Figure 2. Patient responses to custom survey questions regarding their experience with an anesthesiologist under traditional circumstances (A) vs while wearing an Apple Vision Pro headset (B). The figure depicts Likert-scale data, with participants rating each statement from "strongly disagree" to "strongly agree." These responses were statistically analyzed, the results of which are presented in Table 2.



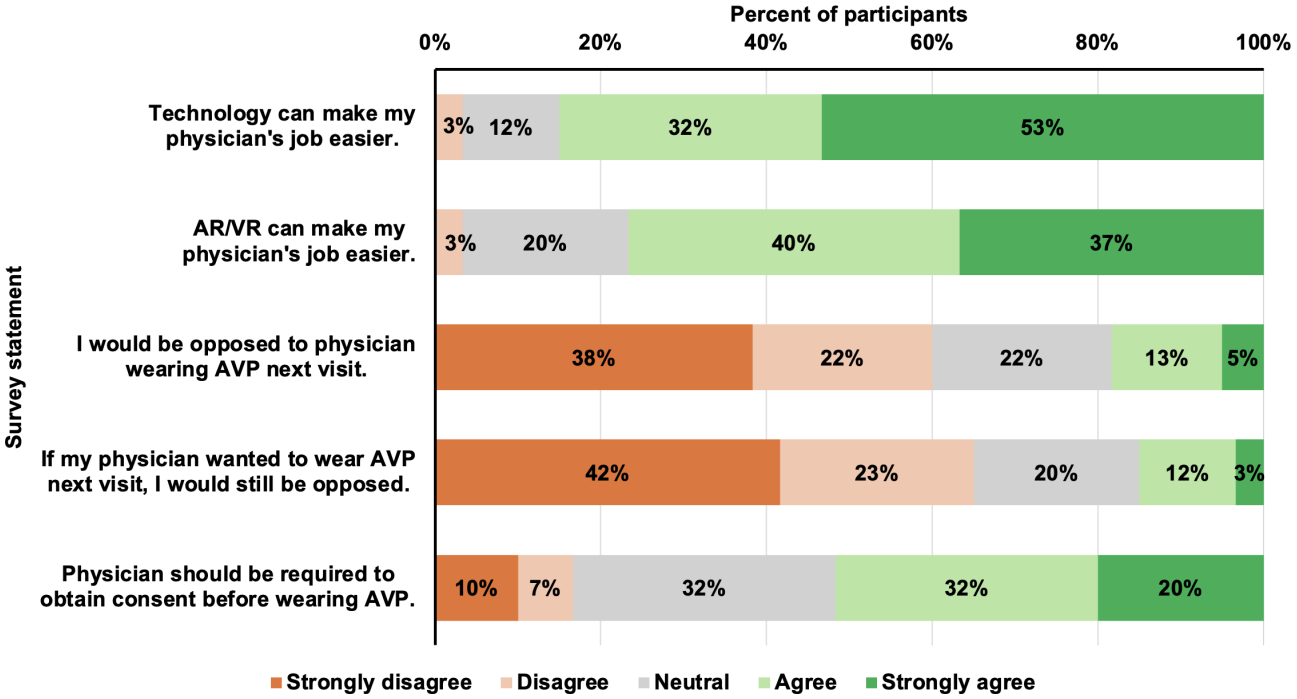
Technology Acceptance and Privacy

Custom questions regarding the patients' overall opinions on the role of technology in health care revealed high levels of technology acceptance (Figure 3). In total, 85% (51/60) of the patients agreed or strongly agreed that technology can make a physician's job easier, with 76.7% (46/60) agreeing that AR or VR can make a physician's job easier during patient encounters. A total of 18.3% (11/60) of the patients were opposed or strongly opposed to their physician wearing an AR or VR headset at their next encounter. When the question was slightly changed to ask whether they would still be opposed if their physician told them they preferred to use it, 13.3% (8/60) of the patients were opposed.

Just over half (31/60, 51.7%) of participants believed that physicians should be required to obtain consent from patients before wearing an AR or VR headset during an encounter. Patients reported slightly increased privacy concerns during the headset encounter compared to no headset (P=.002; r=0.42); however, this was primarily caused by an increased number of patients responding neutrally (3 out of 5) to the following statement: "I was concerned about my privacy." One patient in the headset condition agreed with the statement, and no patients strongly agreed.

XSL•FO
RenderX

Figure 3. Patient perspectives on the role of augmented reality (AR) and virtual reality (VR) in health care and the need for informed consent from patients in these situations. Participants responded to each statement on a Likert scale ranging from "strongly disagree" to "strongly agree." AVP: Apple Vision Pro.



Discussion

Principal Findings

We found that the use of an AR or VR headset negatively impacted perceptions of provider empathy and communication. While participants in both conditions reported relatively high SCQ ratings, the with-headset encounters received statistically significantly lower SCQ scores than the without-headset encounters, suggesting that the technology may create subtle barriers to physician-patient connection. We hypothesize that this decline in perceived compassion could stem from reduced eye contact, altered nonverbal communication, or a sense of detachment caused by the headset's presence.

EyeSight received mixed reviews from participants in this study as just over half of patients found EyeSight to be realistic (32/60, 53%) and felt satisfied with their ability to maintain eye contact with their physician (33/60, 53%). Although we did not evaluate a fully occlusive headset in our study, 25% (15/60) of the participants reported that they would be opposed to their physician wearing one. Our findings are relatively consistent with those of Vergari et al [33], who conducted a pilot study showing that interactions with the Apple Vision Pro were perceived as more engaging and comprehensible compared to those using the Meta Quest 3, a fully occlusive headset, but less than face-to-face interactions.

The findings of the technology acceptance questionnaire suggest that patients remain open to the use of technology in health care workflows, with most patients recognizing its potential to enhance physician efficiency and health care delivery. Specifically, over three-quarters of the participants agreed that technology in general can improve health care delivery (51/60, 85%) and that AR and VR specifically could enhance physician capabilities (46/60, 77%).

Despite this overall acceptance, some reservations remain regarding the direct implementation of AR and VR headsets in patient encounters. A subset of patients opposed physician use of the headset, and over half (31/60, 51.7%) of patients believed that physicians should obtain consent before wearing the headset during an appointment, highlighting the importance of maintaining patient autonomy in technology integration. Notably, privacy concerns were minimal, with only 1.7% (1/60) of the patients expressing concern while the physician used the Apple Vision Pro. These findings underscore the need for clear communication between physicians and patients regarding the purpose and benefits of AR and VR technology while also ensuring that workflows integrating this technology incorporate patient preferences and privacy-preserving frameworks.

This study offers several insights into how AR technology may be more effectively integrated into clinical encounters. As 16.7% (10/60) of the patients would be opposed to their physician wearing the Apple Vision Pro at their next visit and 25% (15/60) would be opposed to their physician wearing a fully occlusive headset, the simulated eye contact provided by the EyeSight feature may be slightly preferable to a fully occlusive headset; however, it likely does not fully preserve the benefits of natural eye contact [33]. Finally, the justification for physician use of these technologies, including measured or perceived benefits, and an acknowledgment of the potential for friction in the patient-physician relationship should be integrated into larger discussions of informed consent. Such transparency may reduce patient confusion or discomfort, allowing the technology to be a supportive element of care rather than a distraction.

While the Apple Vision Pro certainly holds potential to enhance physician workflows and patient care, we believe that its success

XSL•FO
RenderX

may depend on thoughtful design, clinician preparation, and open communication with patients.

Limitations

This was a single-center study at an academic institution, which may limit the generalizability of these results. Participant demographics were skewed relative to the local patient population (eg, 48/60, 80% of the participants identified as White). We studied a convenience sample of patients undergoing endoscopic GI procedures due to the high turnover of patients in one day and the relatively lower acuity of cases; however, this likely skewed the age range of our participants (53/60, 88.3% were over 40 years of age, and 30/60, 50% were over 60 years of age). Generalizability is also limited by our study evaluating brief preprocedural encounters; these results may not be generalizable to longer patient encounters, such as in a primary care clinic. The short interval between the headset and no-headset interactions may have introduced a carryover effect into our data as there was no washout period. Finally, the novelty effect associated with new technologies warrants further longitudinal studies to assess whether reactions may normalize over time as adoption increases.

Future Directions

Future research should build on this study by incorporating a larger and more diverse patient population to better understand how patients of different ages and from different backgrounds perceive the use of AR and VR in health care. Comparing additional headset designs—including AR glasses (eg, Meta Orion), an AR headset with transparent lenses (eg, Microsoft HoloLens), or a fully occlusive VR headset (eg, Meta Quest)—could help determine which type of design best balances physician needs and patient comfort with or trust in their physician. Additionally, further investigation is needed to develop strategies that minimize perceived reductions in physician attention and empathy. Future studies should also more deeply assess AR and VR's influence on the physician-patient relationship over time by including validated scales of therapeutic alliance and communication. Finally, long-term studies are also needed to follow how patient opinions change over time as availability, affordability, and adoption of this technology increases.

Conclusions

As AR and VR technologies are adopted into clinical settings, their impact on the empathy and communication from physicians must be carefully considered. Statistically significantly lower SCQ scores in the headset condition suggest that, without thoughtful design and implementation, AR and VR headsets may introduce unintended barriers to patient focus and engagement. However, patients remain open-minded with regard to the implementation of technology, including AR and VR, into clinical practice, and most would not oppose its use if it helps their physician perform their job better or more easily. Careful implementation and ongoing research will be essential to fully realize the potential of AR and VR in clinical settings while ensuring that patient-centered care and high-quality communication remain a top priority.

Acknowledgments

The authors would like to thank the Simulation Training Center at the University of California, San Diego School of Medicine for lending their Apple Vision Pro to them for use in this study. They would also like to thank the T. Denny Sanford Institute for Empathy and Compassion at the University of California, San Diego, for providing the funding for the US $5 Visa gift cards that were provided to participants who completed the study protocol. The authors attest that no generative AI was used at any point during the conduct of this study or the creation of this manuscript.

Funding

Funding for the US $5 gift cards provided to participants who completed the study protocol was provided by the T. Denny Sanford Institute for Empathy and Compassion at the University of California, San Diego. No other funding was required or used in this study.

Data Availability

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

Authors' Contributions

Conceptualization: SC, JFR, MT, BF, AY, CD, PJS, JLT
Data curation: JFR, BKL, JLT
Formal analysis: JFR, BKL, JLT
Funding acquisition: PJS, SC
Investigation: SC, JFR, MT, BF, JLT
Methodology: SC, JFR, MT, PJS, JLT
Project administration: SC, JFR, PJS, JLT
Resources: PJS, JLT
Supervision: JFR, MT, PJS, JLT
Visualization: SC, JFR, BKL, JLT

XSL•FO
RenderX

Writing—original draft: SC, JFR
Writing—review and editing: SC, JFR, BKL, MT, BF, AY, CD, PJS, JLT

Conflicts of Interest

None declared.

Multimedia Appendix 1
Custom survey questions answered by participants.
[DOCX File, 14 KB - xr_v3i1e95806_app1.docx]

References

1. Solbiati L, Gennaro N, Muglia R. Augmented reality: from video games to medical clinical practice. Cardiovasc Intervent Radiol 2020 Oct;43(10):1427-1429. [doi: 10.1007/s00270-020-02575-6] [Medline: 32632853]
2. Singh B, Vig K, Kaunert C. Modernizing healthcare: application of augmented reality and virtual reality in clinical practice and medical education. In: Hiran KK, Doshi R, Patel M, editors. Modern Technology in Healthcare and Medical Education: Blockchain, IoT, AR, and VR: IGI Global Scientific Publishing; 2024:1-21. [doi: 10.4018/979-8-3693-5493-3.ch001]
3. Yeung AW, Tosevska A, Klager E, et al. Virtual and augmented reality applications in medicine: analysis of the scientific literature. J Med Internet Res 2021 Feb 10;23(2):e25499. [doi: 10.2196/25499] [Medline: 33565986]
4. Carmigniani J, Furht B. Augmented reality: an overview. In: Furht B, editor. Handbook of Augmented Reality: Springer; 2011:3-46. [doi: 10.1007/978-1-4614-0064-6_1]
5. Abbas JR, O'Connor A, Ganapathy E, et al. What is virtual reality? A healthcare-focused systematic review of definitions. Health Policy Technol 2023 Jun;12(2):100741. [doi: 10.1016/j.hlpt.2023.100741]
6. Barcali E, Iadanza E, Manetti L, Francia P, Nardi C, Bocchi L. Augmented reality in surgery: a scoping review. Appl Sci 2022;12(14):6890. [doi: 10.3390/app12146890]
7. Burström G, Persson O, Edström E, Elmi-Terander A. Augmented reality navigation in spine surgery: a systematic review. Acta Neurochir (Wien) 2021 Mar;163(3):843-852. [doi: 10.1007/s00701-021-04708-3] [Medline: 33506289]
8. Tabernée Heijtmeijer S, Glas H, Janssen N, et al. Accuracy of augmented reality navigated surgery for placement of zygomatic implants: a human cadaver study. PeerJ 2024;12:e18468. [doi: 10.7717/peerj.18468] [Medline: 39670105]
9. Lavé A, Meling TR, Schaller K, Corniola MV. Augmented reality in intracranial meningioma surgery: report of a case and systematic review. J Neurosurg Sci 2020 Aug;64(4):369-376. [doi: 10.23736/S0390-5616.20.04945-0] [Medline: 32347678]
10. Tanzer M, Laverdière C, Barimani B, Hart A. Augmented reality in arthroplasty: an overview of clinical applications, benefits, and limitations. J Am Acad Orthop Surg 2022 May 15;30(10):e760-e768. [doi: 10.5435/JAAOS-D-21-00964] [Medline: 35245236]
11. Rai AS, Rai AS, Mavrikakis E, Lam WC. Teaching binocular indirect ophthalmoscopy to novice residents using an augmented reality simulator. Can J Ophthalmol 2017 Oct;52(5):430-434. [doi: 10.1016/j.jcjo.2017.02.015] [Medline: 28985799]
12. Bruno RR, Wolff G, Wernly B, et al. Virtual and augmented reality in critical care medicine: the patient's, clinician's, and researcher's perspective. Crit Care 2022 Oct 25;26(1):326. [doi: 10.1186/s13054-022-04202-x] [Medline: 36284350]
13. Chow AR. Why surgeons are wearing the Apple Vision Pro in operating rooms. TIME. 2024. URL: https://time.com/7093536/surgeons-apple-vision-pro [accessed 2026-02-28]
14. Apple Vision Pro. Apple. URL: https://www.apple.com/apple-vision-pro/specs [accessed 2026-07-02]
15. Shanbhag NM, Bin Sumaida A, Al Shamisi K, Balaraj K. Apple Vision Pro: a paradigm shift in medical technology. Cureus 2024 Sep;16(9):e69608. [doi: 10.7759/cureus.69608] [Medline: 39308843]
16. Armstrong DG, Bazikian S, Armstrong AA, Clerici G, Casini A, Pillai A. An augmented vision of our medical and surgical future, today? J Diabetes Sci Technol 2024 Jul;18(4):968-973. [doi: 10.1177/19322968241236458] [Medline: 38439541]
17. Masalkhi M, Waisberg E, Ong J, et al. Apple Vision Pro for ophthalmology and medicine. Ann Biomed Eng 2023 Dec;51(12):2643-2646. [doi: 10.1007/s10439-023-03283-1] [Medline: 37332003]
18. Kaye AD, Islam RK, Islam KN, et al. Apple Vision Pro and its implications in Mohs micrographic surgery: a narrative review. Cureus 2024 Oct;16(10):e71440. [doi: 10.7759/cureus.71440] [Medline: 39544562]
19. Volin J, Viswanathan V, Krupinski EA, et al. Feasibility of using the Apple Vision Pro for diagnostic radiology: user experience and perceived clinical utility. J Imaging Inform Med 2025 Nov 17. [doi: 10.1007/s10278-025-01745-7] [Medline: 41249669]
20. Diaz N. Epic dives into spatial computing. Becker's Health IT. 2024 Feb 28. URL: https://www.beckershospitalreview.com/healthcare-information-technology/ehrs/epic-dives-into-spatial-computing/ [accessed 2026-07-02]
21. Apple Vision Pro unlocks new opportunities for health app developers. Apple. 2024. URL: https://www.apple.com/newsroom/2024/03/apple-vision-pro-unlocks-new-opportunities-for-health-app-developers/ [accessed 2026-07-02]
22. Sharp and Apple Vision Pro: the future of health care. Sharp HealthCare. 2024. URL: https://www.sharp.com/health-news/sharp-and-apple-vision-pro-the-future-of-health-care [accessed 2026-07-10]

XSL•FO
RenderX

23. Drossman DA, Ruddy J. Improving patient-provider relationships to improve health care. Clin Gastroenterol Hepatol 2020 Jun;18(7):1417-1426. [doi: 10.1016/j.cgh.2019.12.007] [Medline: 31843593]
24. Montague E, Chen PY, Xu J, Chewning B, Barrett B. Nonverbal interpersonal interactions in clinical encounters and patient perceptions of empathy. J Particip Med 2013;5:e33 [FREE Full text]
25. Gorawara-Bhat R, Cook MA. Eye contact in patient-centered communication. Patient Educ Couns 2011 Mar;82(3):442-447. [doi: 10.1016/j.pec.2010.12.002] [Medline: 21211926]
26. Odenheimer S, Goyal D, Jones VG, Rosenblum R, Ho L, Chan AS. Patient acceptance of remote scribing powered by Google Glass in outpatient dermatology: cross-sectional study. J Med Internet Res 2018 Jun 21;20(6):e10762. [doi: 10.2196/10762] [Medline: 29929947]
27. What EyeSight shows on Apple Vision Pro. Apple. URL: https://support.apple.com/en-us/120481 [accessed 2026-07-02]
28. Compassion Measure. URL: https://www.compassionmeasure.com [accessed 2026-07-02]
29. Sinclair S, Hack TF, MacInnis CC, et al. Development and validation of a patient-reported measure of compassion in healthcare: the Sinclair Compassion Questionnaire (SCQ). BMJ Open 2021 Jun 16;11(6):e045988. [doi: 10.1136/bmjopen-2020-045988] [Medline: 34135041]
30. Woolson RF. Wilcoxon signed-rank test. In: Wiley Encyclopedia of Clinical Trials: John Wiley & Sons; 2007. [doi: 10.1002/9780471462422]
31. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Routledge; 1988. [doi: 10.4324/9780203771587]
32. 45 CFR 46.104: exempt research. Electronic Code of Federal Regulations. URL: https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46/subpart-A/section-46.104 [accessed 2026-06-29]
33. Vergari M, Kojic T, Wardah W, et al. Digital eyes: social implications of XR EyeSight. In: VRST '24: Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology: Association for Computing Machinery; 2024. [doi: 10.1145/3641825.3689526]

Abbreviations

AR: augmented reality
GI: gastrointestinal
IRB: Institutional Review Board
SCQ: Sinclair Compassion Questionnaire
VR: virtual reality

Edited by I Steenstra; submitted 30.Mar.2026; peer-reviewed by BS Chaparro, S Slatman; revised version received 21.Jun.2026; accepted 23.Jun.2026; published 23.Jul.2026.

Please cite as:
Chriqui S, Ryan JF, Luu BK, Tran M, Fergerson B, You A, Dameff C, Suresh PJ, Tully JL
Patient Perception of the Use of the Apple Vision Pro Virtual Reality Headset in an Anesthesiology Informed Consent Process: Prospective Randomized Controlled Crossover Trial
JMIR XR Spatial Comput 2026;3:e95806
URL: https://xr.jmir.org/2026/1/e95806
doi:10.2196/95806

© Sabrina Chriqui, John Francis Ryan, Bao Kien Luu, Minh Tran, Byron Fergerson, Alan You, Christian Dameff, Preetham John Suresh, Jeffrey Logan Tully. Originally published in JMIR XR and Spatial Computing (https://xr.jmir.org), 23.Jul.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 XR and Spatial Computing, is properly cited. The complete bibliographic information, a link to the original publication on https://xr.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX