DOTTORATO DI RICERCA IN
SCIENZA APPLICATE A BENESSERE E SOSTENIBILITA'

CICLO XXXVIII

COORDINATORE Prof. Carlo Baldari

TITOLO

The User eXperience questionnaire for Virtual Reality in Healthcare (UXVR-H): Development from Literature Analysis to Psychometric Evaluation in Older Adults

Settore Scientifico Disciplinare M-PSI/03

Dottorando	**Tutor**
Dott.ssa Francesca Bruni	Prof.ssa Elisa Pedroli

ABSTRACT

Within the framework of healthcare, virtual reality has emerged as a transformative tool for delivering therapeutic interventions through increasingly sophisticated and effective applications. The successful implementation of virtual reality-based solutions is linked to the overall perception that users have when interacting with the product - the so-called user experience. Positive user experience yielded several advantages, such as identification of usability barriers, accommodation of patients' specific needs and limitations, enhanced treatment adherence, sustained patient engagement, and long-term adoption of the proposed intervention. Despite this recognized importance, comprehensive instruments for evaluating the experience of users in healthcare settings remain limited. Existing tools often fail to capture the multifaceted nature of novel technological solutions or lack appropriate psychometric validation.

In order to fill these gaps, this doctoral dissertation developed the User eXperience questionnaire for Virtual Reality applications in Healthcare (UXVR-H), a psychometrically sound tool created to evaluate user experience in virtual reality-based applications for healthcare among older adults. The development process employed a systematic, multi-phase approach grounded in comprehensive literature analysis and rigorous validation procedures.

The research comprised four interconnected studies. First, a scientometric analysis of the field, examining the evolution of user experience, thematic shift, and emerging trends to delineate a more precise conceptual profile of this complex and multifaceted construct. Second, the literature analysis was completed by a scoping review specifically focused on the healthcare context, exploring theoretical frameworks of user experience when virtual reality is used for clinical purposes in aging. As a result, we theorized an eight-factor framework for defining user experience, encompassing: usability and functionality, aesthetics and design, engagement, emotional state, presence, environmental realism, side effects, and usage motivation. This model served as the foundation for questionnaire construction.

The UXVR-H development starts with item generation (271 items), formulation of the response scale, and drafting of administration instructions. The design process included the consultation of older adults to apply user-centered design principles. This participatory approach guarantees the understandable, pertinent, and suitability of items for the target population. Additionally, experts in virtual reality technology, methodology, and user experience evaluated the content validity and theoretical alignment. Through integration of user feedback and expert judgment,

the refined instrument includes 24 items for further study. It is aimed at evaluating the psychometric qualities of the questionnaire. It involves older adults who completed the UXVR-H following a virtual reality session featuring cognitive tasks. Data screening, confirmatory factor analysis, reliability analysis, and nomological validity testing were all part of the psychometric evaluation. Results revealed a refined 13-item structure distributed across six empirically supported dimensions: usability and functionality, presence, side effects, aesthetics and design, environmental realism, and usage motivation. The UXVR-H emerges as a practical self-report instrument utilizing a 5-point Likert scale that provides a balanced assessment of essential UX dimensions while minimizing respondent burden through its brevity (approximately 5 minutes administration time).

In conclusion, the present research analyzes the complex framework of user experience, highlights current challenges in evaluating user experience following virtual reality-based applications in the clinical context for aging, and offers a valuable contribution to the limited collection of validated assessment tools.

INDEX

ABSTRACT 3
INTRODUCTION 7
User Experience 7
Pragmatic vs hedonic aspect 10
Which factors influence UX? 11
Time as an important factor in UX 13
Methodology to evaluate UX? 15
Virtual Reality 18
Application of virtual reality 22
Virtual reality in healthcare 22
User Experience in VR Applications for health 24
EXPERIMENTAL STUDIES 27
A Scientometric Analysis of User Experience: Exploring Evolution, Emerging Themes, and Impacts 28
Introduction 28
Material and methods 29
Results 30
Discussion 37
How to evaluate User Experience in Digital Health? A Scoping Review of tools in Virtual Reality Applications 38
Introduction 38
Methods 40
Results 42
Discussion 57
Conclusion 60
User eXperience questionnaire for Virtual Reality applications in Healthcare: development of a new tool 60
Introduction 60
Methods 62

Results 65

Discussion 68

Preliminary Validation of the UXVR-H: A User Experience Questionnaire for Healthcare Virtual Reality 69

Introduction 69

Material and methods 71

Results 74

Discussion 80

DISCUSSION 84

CONCLUSION 88

References 89

INTRODUCTION

User Experience

User Experience (UX) has evolved significantly over recent years, gaining substantial attention from both academia and industry. Nielsen Norman Group, widely acknowledged as the founding authority in the UX domain, defines UX as all aspects of the end-user's interaction with the company, its services, and its products (Nielsen, 1994). While this definition has been perfected over time, it remains predominantly abstract and fails to delineate a clear conceptualization of which is the range of factors that constitute UX. Nevertheless, this approach emphasizes UX as a multifaceted interaction a user has with a product or service, from first contact to continuous use, not limited to the interface or usability. Indeed, UX is fundamentally inseparable from the broader concept of experience per se (Hassenzahl, 2010). To develop a theoretical framework for UX, it is first necessary to examine the ontological nature of experience. Although experience may be conceptualized differently across disciplinary contexts, its essential characteristics remain dynamic and subjective attributes. This indicates that experience is formed in the mind through interaction with the world, involving sensory, cognitive, and social elements - rather than existing as an objective phenomenon independent of the experiencer (Glanznig, 2012; Godovykh & Tasci, 2020; Hassenzahl, 2010; Hassenzahl & Tractinsky, 2006; Matus et al., 2021; Wolf & Jason, 2014).

In the technological context, the term "experience" encompasses all dimensions of human-product interaction such as the aspects of how people interact with a product such as the way it feels in their hands, how well they understand how it works, how they feel about it while they are using it, how well it serves their purposes, how well it fits into the entire context in which they are using it.

The fusion of all these multiple components, including users' internal states (emotions, expectations, and active goals), makes UX challenging (Abro et al., 2015; Basri et al., 2016). Experience is what emerges as the consequence of this integrative process wherein these constituent elements intersect, interact, and reciprocally transform one another (Hassenzahl and Tractinsky, 2006). Even though the theoretical subjectivity emerging in the literature regarding UX conceptualization, there is a common agreement that UX represents a dynamic and complex construct requiring multifaceted analytical approaches.

Historical Foundations: From Ergonomics to User Experience

The emergence of UX represents the culmination of various disciplinary traditions concerned with human-technology interaction. To fully appreciate the contemporary conceptualization of UX, it is essential to trace its theoretical origin through multiple domains that have historically addressed the relationship between humans and technological systems (Triberti and Brivio, 2017).

The field of study emerged in the late 19th and early 20th centuries with the advent of industrial ergonomics, primarily focused on optimizing worker productivity and safety in industrial settings. This approach was exemplified by Taylor's scientific management principles, which establish the foundation for systematic analysis of human performance in technological contexts (Taylor, 1911). Taylor’s model predominantly emphasized objective performance metrics and physiological constraints while largely neglecting the subjective dimensions of human experience and the emotional relationship between workers and their tools. During World War II, with the need to optimize complex military equipment interfaces, the field of ergonomics shifted its focus from adapting humans to machines toward designing machines that accommodate human capabilities and limitations. This paradigm introduced critical principles regarding cognitive workload, attention allocation, perception, and human error that remain fundamental to contemporary technology evaluation (Wickens et al., 2004). In this perspective, rigorous methodologies gained importance for measuring performance, analyzing task demands, and identifying design flaws that compromise safety or efficiency. In the 1980s, the domain of human-computer interaction (HCI) emerged as computing technology transitioned from specialized technical communities to broader public adoption. HCI introduced methodologies specifically tailored to interactive digital systems, emphasizing iterative design processes, user-centered evaluation techniques, and interface design principles (Norman and Draper, 1986). HCI researchers developed sophisticated evaluation techniques, including heuristic evaluation and cognitive walkthroughs, that systematically assessed how users interact with interfaces. Yet early HCI research primarily addressed pragmatic concerns, without consideration of emotional engagement, pleasure, or the broader life context within which technology is experienced. In this context, the concept of usability was formalized. It has been defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (ISO 9241-11, 1998). It represented a significant milestone as usability evaluation established standardized frameworks for assessing whether systems enable users to accomplish their intended tasks without excessive cognitive burden or frustration. However, usability frameworks remained fundamentally constrained by their instrumental focus (Hassenzahl, 2004; Triberti and Brivio, 2017). They treated technology primarily as a tool for achieving predefined goals, inadequately capturing the experiential richness in their human-technology relationships. An in-depth approach is presented by theories analyzing motivational factors that influence technology adoption in organizational and consumer contexts, particularly the Technology Acceptance Model (TAM; Davis et al., 1989) and its successors, the Unified Theory of Acceptance and Use of Technology (UTAUT; Venkatesh et al., 2012). These models identified perceived usefulness and perceived ease of use as key determinants of behavioral intention and system use, establishing theoretical frameworks that bridged cognitive psychology with practical concerns about technology implementation. However, technology acceptance models typically treated experience as something that influences adoption decisions rather than as a multifaceted phenomenon worthy of comprehensive investigation. Against this disciplinary backdrop, the explicit articulation of UX as a construct emerged in the mid-1990s. This conceptual innovation represented a fundamental paradigm shift that transcended previous frameworks by positioning subjective experience—encompassing emotional responses, aesthetic appreciation, meaning-making, identity construction, and temporal dynamics—as central rather than peripheral to technology evaluation. The UX concept recognizes that the relationship between users and technological objects is not merely functional but deeply psychological, involving cognitive, emotional, motivational, and social dimensions that collectively shape how technology is perceived, valued, and integrated into daily life. This expanded perspective acknowledges that people do not merely use technology; they bring to these interactions personal histories, cultural backgrounds, expectations, and aspirations. The UX framework thus integrates insights from diverse traditions—ergonomics' attention to human capabilities and constraints, HCI's iterative design methodologies and interface principles, usability's effectiveness metrics and user-centered evaluation, and technology acceptance's motivational constructs—while fundamentally reorienting evaluation toward holistic experiential quality.

The differentiation between pragmatic and hedonic aspects of technological interaction represents one such theoretical distinction.

Pragmatic vs hedonic aspect

A prevalent methodological approach to conceptualizing UX involves delineating specific dimensions of which it is composed. Although numerous theoretical models have emerged over time, researchers' consensus primarily consolidates around the fundamental distinction between pragmatic and hedonic components (Maia and Furtado, 2016). Traditionally, following Hassenzahl's influential framework, UX was identified through a combination of pragmatic (i.e., instrumental) and hedonic (i.e., non-instrumental) qualities of a system (Hassenzahl, 2004). Pragmatic is associated with the so-called utility and usability aspects, such as perceived usefulness, effectiveness, ease of use, and operational efficiency. This dimension pertains to goal-directed behavior, exemplified by task completion such as locating a specific folder on a computer. Meanwhile, hedonic attributes reflect the joy of use, incorporating elements of stimulation, identification, and evocation derived from system interaction. This dimension relates to the subjective perception of competence, exemplified by the feeling of being competent, such as the perceived satisfaction associated with successfully navigating through the computer.

While pragmatic and hedonic aspects are conceptually distinct, evidence demonstrates their significant positive correlation (Hassenzahl, 2007; Kollmorgen et al., 2022); they function synergistically to generate either positive or negative feelings, which guide the acceptance of the product. Hedonic qualities improve the whole experience by making it pleasurable and emotionally fulfilling, whereas pragmatic qualities guarantee that the product is functional and effectively satisfies user demands (Kujala et al., 2011; Schrepp et al., 2006). Nevertheless, the hedonic component influences long-term user happiness and loyalty in a way that varies more and more depending on the temporal environment, frequently growing in importance with time (Kollmorgen et al., 2022; Kujala et al., 2011). Minge (2008) observed that while pragmatic attributes facilitate users to achieve hedonic goals, it is predominantly the hedonic qualities that determine product adoption. Furthermore, longitudinal research indicates that pragmatic elements represent primary determinants of perceived quality during initial interactions with novel technologies. However, after four weeks of use, their influence diminishes in favor of hedonic qualities, specifically identification, integration in daily rituals and incorporation into habituated activities (Karapanos et al., 2009, 2008). While longitudinal evidence suggests a decline in pragmatic influence in favor of hedonic qualities as interaction progresses, the hierarchical overall architecture shows pragmatics as essential. A hierarchical paradigm for designing good experiences was advanced by Jordan (Jordan, 2000). It proposed that a product must exhibit useful and practical features before hedonic qualities, such as aesthetic appreciation and stimulation, can appear. This cornerstone has remained at the center of several models in recent years. Authors added various layers, core elements, and components of UX, suggesting numerous variables operate within this complex framework, all based on good product functionality (Albert and Tullis, 2022). While instrumental features exert substantial influence on UX quality, the literature suggests that the pragmatic aspect alone is insufficient for a comprehensive understanding of UX; thus, researchers must direct substantial attention to the hedonic dimension while also investigating the broader constellation of factors that collectively shape UX.

Which factors influence UX?

Throughout its conceptual evolution, UX has been persistently confused and intertwined with the term usability, to the extent that some practitioners still employ the term UX as a mere synonym for usability. While usability constitutes an important factor, it provides only a pragmatic perspective limited to objective assessment parameters, failing to address the complex panorama of subjective technological interaction (Schulze et al., 2023). Transcending the constraints of purely functional considerations, numerous non-instrumental features have been introduced into theoretical frameworks over the years, refining the conceptualization of UX and detaching it from the device-related quality characteristic of the usability paradigm. As early as 1996, Alben introduced aesthetics as a critical component of technology evaluation (Alben, 1996), a position reinforced during time by various authors who demonstrated how interface aesthetics function as a powerful determinant of user satisfaction and pleasure (Abbasi et al., 2012; de Souza Lima et al., 2022; Lavie and Tractinsky, 2004; Li et al., 2024). For instance, aesthetic impressions may induce a halo effect, in which positive feelings toward the aesthetics extend toward other attributes, such as usability, content quality, and other attributes (Hartmann et al., 2008). Also, the context of use and the product's intended purpose remain important beyond the hedonic dimensions. Van der Heijden (2004) noted that the most critical variables targeting the UX differ based on the system's goal and purpose in relation to the hedonic dimensions. In some cases, the enjoyment rather than the usefulness of a system may dominate the prediction of a positive user experience. In other cases, the emotional fit of a product, its aesthetics, and its identity may take precedence in determining whether a user's experience is positive or negative. In contrast, more functional products may require higher user engagement, perceived usefulness, and quality of interaction with the system (Van der Heijden, 2004).

Research inquiry has expanded to examine user-specific attributes as critical mediators of technological experience (Augstein et al., 2017). Significant differences have been documented across demographic variables for example (i.e., gender and age). Studies indicate that gender differences substantially influence technological experience (Parsons et al., 2004). These differences involve the technological adoption decision-making process, access, and ease of use to the point that it can be conceptualized as a gender digital divide (Venkatesh and Morris, 2000). However, gender differences were not significant in the context of the product usage life cycle (Yoon et al., 2020). Moreover, age represents another important factor that influences understanding systems (Nimmo et al., 2024). Particularly, older adults frequently report diminished self-efficacy regarding interaction with contemporary technologies such as smartphones or tablets, which subsequently impedes technology adoption and a noticeable performance gap between older and younger users (J. Liao et al., 2018). Systems typically prioritize design consideration for a young public, resulting in interaction paradigms that older users often experience as complex and frustrating (Strada et al., 2013). While younger users prefer modern designs, older adults favor designs that use metaphors from the non-digital world (Backhaus et al., 2018), reporting higher satisfaction with simpler and intuitive interfaces.

Other important aspects era educational background and interface design preferences, which significantly influence UX. Research demonstrates that distinct user demographics exhibit varying preferences for interface types, which subsequently affect both task performance and system evaluations. Is the case of Yi and colleagues, who compare university students with workers. Students typically prefer minimalist interfaces with customization options, while industrial workers often value function-focused interfaces with prominent safety features and clear visual hierarchies (Yi et al., 2021). This evidence underscores how users' responses to technologies are shaped by multiple interconnected factors, where psychological elements also deserve a special mention. For instance, personality traits, values, individual preferences, and prior experiences play a crucial role in this process. Among other users with high openness to experience are more likely to explore novel interface features, while those scoring high on introversion may prefer predictable, consistent designs that minimize uncertainty (Koch et al., 2013; Salimun et al., 2021). The identity and emotional responses of users fundamentally shape their experiences. This indicates the need to think about the experience not as separate, stand-alone events, but continuous relational story. This time dimension takes into account the past personality traits and influences like established aspects of a personality, thoughts and valuable ideals, for example a past negative experience with similar technologies and possibly an emotional resistance to adoption, and active perceptual frameworks like a cultural background that may determine how a user perceives color, and symbols, as well as how users predicted the technological advances that may help in the fulfillment of their needs. The results are that an additional range of non-instrumental elements intervene in the UX scene, such as surprise, distraction, intimacy, need, and value alignment, just to name a few (Gaver and Martin, 2000; Lun et al., 2024). For example, a banking application may be functionally robust yet fail to engage users who have experienced previous negative interactions or who possess naturally cautious dispositions, particularly if it does not establish a profound sense of security and trust through thoughtfully implemented design elements that align with users' fundamental psychological needs for safety and reliability.

Beyond the numerous factors influencing user experience—from product aesthetic and functional characteristics to user-specific attributes like age, gender, and educational background—a fundamental and transversal dimension emerges that deserves particular attention: temporality, or how user experience develops, evolves, and is perceived over time.

Time as an important factor in UX

The dynamic nature of UX necessitates the consideration of temporal variables in UX research (Kujala, 2013; Minge and Fr, 2008). The role of time on experiences can be viewed on the accounts of memory processes (Kujala, 2013; Pohlmeyer et al., 2009). While during immediate interaction may be separated by memory, specific experiences subsequently remain in users' minds, significantly influencing their holistic perception of the product. Pohlmeyer and colleagues emphasize the persistence of experiential beyond single behavioral episodes and momentary evaluations, stating that "a user experience continues as long as memories last." (Pohlmeyer et al., 2009). This raises important methodological questions about when to collect the data. Data collection can occur in three stages in user research: pre, real-time (i.e., during-interaction), and post-interaction.

Pre-interaction data collection occurs when a user has not yet engaged with a product or service. This practice has the researcher record baseline user information, expectations, and most importantly, attitudinal and perceptual frameworks. Surveys and questionnaires predominate these methods; however, they cannot capture real-time data and refer to preliminary vibes. It may be useful when combined with post-experience data to compare user attitudes before and after product usage (Kurosu et al., 2018), thereby highlighting expectation-experience gaps that may influence overall satisfaction.

Real-time evaluation approach captures users' immediate performance and reactions during product interaction, aiming to provide insight into momentary experiences. These approaches measure parameters such as task completion time, error rates, error correction duration, and learning curves, often supplemented by physiological metrics (Hassenzahl and Tractinsky, 2006). While providing high ecological validity and minimizing possible recall bias, real-time data collection required complex tracking instruments that might provide intrusive and disrupt the user's natural interaction experience, potentially compromising the authenticity of the observed behaviors (Inan Nur et al., 2021).

Post-experience reports consisted of the detection of the user perspective after their interaction with the product, commonly through questionnaires and interviews. Thanks to several advantages, this measurement remains the predominant approach in UX research. It provides feedback about overall UX while minimizing interference during product usage (Hinderks et al., 2019). Post-interaction reports offer deeper insights that objective, real-time metrics cannot capture, despite being reliant on users' retrospective memories, which introduces potential recall bias. Participants can reflect on their experience, allowing nuanced emotional responses, unexpected challenges, and comprehensive impressions that might not be immediately apparent during interaction. Moreover, post-interaction requires low technological and logistical complexity compared to real-time tracking methodologies.

UX research increasingly advocates for a combined approach that integrates data across all three temporal stages (pre-, during, and post-interaction) - offering the most holistic understanding of UX - or even real-time performance metrics with post-experience reflective data (Maia and Furtado, 2016). Although few studies collect data across all three temporal stages, concurrent evaluation methodologies help prevent the loss of relevant information that might occur with delayed data collection. Nevertheless, post-interaction measurement remains the predominant approach in the field, largely due to its favorable cost-benefit profile. Furthermore, evidence suggests that UX can or even should be investigated even long after to capture the long-term experience (Law et al., 2009). From initial discovery, learning, and enthusiasm through several stages of competency growth and integration as part of everyday life, this approach offers useful insights on how a user's experience and relationship with a product evolves (Kujala, 2013). This longitudinal approach is also able to provide critical insights that cannot be obtained through immediate post-use evaluation. These insights include patterns of adaptation, shifting mental models, shifting emotional attachments, and the emergence of usage behaviors that were not foreseen during the initial stages of design. The present framework provides more robust predictive validity regarding product sustainability and long-term user satisfaction, lining up with contemporary understanding of technology adoption as a process rather than a discrete event. Despite this widespread acknowledgment of time's dimension, contemporary UX evaluation methodologies predominantly focus on discrete behavioral episodes and momentary assessments (Maia & Furtado, 2016), failing to adequately capture experience evolution throughout users' daily interactions with products. This perspective raises questions about the most appropriate techniques for capturing the full spectrum of UX dimensions.

Methodology to evaluate UX

Researchers continuously debate the most effective approaches to capture the multifaceted nature of user interactions with products and systems. The only consensus in this ongoing debate is the significance of evaluating UX to identify critical elements of creating high-quality interactive products and offering a satisfying experience overall (Kula et al., 2019), which highlights the difficulty of measuring subjective experiences alongside objective performance metrics.

Traditional methods of UX assessment rely on three primary approaches (Inan Nur et al., 2021), as shown in Table 1, each contributing unique perspectives to the discourse while presenting distinct methodological challenges: physiological methods, observational methods, and self-reported methods (Hussain et al., 2018).

Table 1. Traditional method to evaluate UX

Methods	Description	Strengths	Weaknesses
Physiological Measure	Bodily responses (heart rate, skin conductance, EEG) to assess reactions during user experience	Objective data unaffected by recall bias; captures unconscious responses	Specialized equipment/expertise; complex products
Observational Measure	Expert monitoring of user responses in interacting with a product	Reveals natural user behavior and unexpected interaction patterns	Observer bias: risk of missing internal cognitive/emotional processes without visible manifestations

Self-report Measure	User feedback through questionnaires, interviews, and surveys to understand perceived experiences	Subjective perceptions; easily implemented across large samples with standardized metrics	Vulnerable to recall bias, social desirability effects, and limited ability to capture unconscious or momentary responses
Mixed Methods	Combination of multiple approaches	Comprehensive insights through data triangulation of objective and subjective data	Significant resources; complex coordination challenges

Physiological methods involve biometric sensors mounted on the user's body for collecting physical information. These techniques measure the biological information that reflects how the users feel during product interaction, such as emotional arousal, stress, motivation, and visual attention. Multiple biometric sensors serve different purposes (Hussain et al., 2018), for example, eye trackers detect visual attention (Fu et al., 2016), electroencephalography can detect user motivations and emotional responses (Liu et al., 2017), electroencephalography identifies user motivations and emotional responses, and galvanic skin response measures stress and arousal through skin conductivity (Nourbakhsh et al., 2017), and electrocardiogram and electromyogram assess stress levels and muscle-arousing activities (Greene et al., 2016). These data provide valuable insights unachievable through other techniques; moreover, physiological results offer objective data on user experience otherwise unobtainable (Inan Nur et al., 2021). However, researchers acknowledge their limitations regarding specialized equipment and expertise to use specific devices, synchronize data, analyze results, and produce informed decisions relating to UX. This complexity explains why, despite their potential for objective measurement, physiological methods remain the least frequently employed in UX research practice (Inan Nur et al., 2021).

Observational measurements represent an alternative approach consisting of an expert real-time observation of users during product interactions. Among the observational techniques, the most used involve usability testing and the thinking-aloud method. The latter consists of collecting users' verbalization and their thoughts while using a product; this helps researchers understand cognitive processes and investigate possible issues (Lewis, 1982). Conversely, usability testing involves direct observation and problem identification by experts (Ofosu-Asare, 2024). Researchers detect information related to the task or user performance, efficiency, and errors while the users interact with the system, focusing on both voluntary and involuntary user information. This is particularly useful for participants with limited verbal or cognitive abilities. Using this technique, different variables can be detected, such as facial expression through the video-based facial expression analysis, emotion from voice (Siddiqi et al., 2016), users' interaction by tracking actions through page tracking, or event tracking to app/screen tracking (Scherr et al., 2017). While recognizing emotions makes observational measurements a valuable assessment for affective responses and emotional valence (Noroozi et al., 2017; Zaman and Shrimpton-Smith, 2006), this method has some limitations. It cannot identify emotional arousal or the user's psychological state during product usage. Furthermore, observational measurements face a lot of problems with measurement variance, due to both planned and unplanned factors. These include consistent biases in the observation protocol, and random and unpredictable variations that occur naturally during human observation processes (i.e., observer fatigue, momentary distractions, inconsistencies in the interpretation of user behaviors).

Self-reported measurements are the most frequently used method, whereby users report their feelings and thoughts without expert intervention, unlike in observational measurement methods that require expert oversight. Questionnaire and interviews are the most common self-reporting tools. Questionnaires provide standardized metrics for UX. Interviews, either structured or semi-structured, allow for in-depth exploration of users' perspectives. While various tools have been developed in literature for measuring different UX aspects (such as pragmatics, hedonics, and satisfaction) (Bruni et al., 2025), the critical point highlighted in self-reported assessments regards the difficulty in capturing momentary experiences that occur during product interaction. Moreover, they may be limited by possible bias, such as social desirability effects, or participants' capacity for accurate introspection, potentially compromising the validity of self-reported data in capturing authentic user experiences (Kula et al., 2019). Even when items are clear, most participants struggle with honest and accurate responses; consequently, they may not faithfully articulate their true emotions, abilities, and experiences. Despite their inherent subjectivity, this method remains at the center of methodological debates due to its practical advantages, including cost-effectiveness, ease of administration, scalability across large participant samples, and the standardization of data collection that facilitates quantitative analysis (Inan Nur et al., 2021; Schrepp, 2015).

The ongoing debate increasingly advocates for a mixed-method approach that combines observational, self-reported, and physiological measures to address the limitations inherent in each individual methodology (Hussain et al., 2018). This integrative perspective recognizes that while self-reporting introduces a high source of subjectivity, both self-reporting and observation methods may inadequately capture psychological states, and physiological and observational methods demand specialized expertise; triangulation across methods offers more comprehensive insights into the user experience.

Despite its theoretical advantages, the mixed-method approach presents several limitations, and in practice, it is often difficult to overlook. Implementing diverse methodologies requires complex experimental designs with sophisticated research protocols that can be challenging to coordinate effectively (Hussain et al., 2018). Furthermore, this approach necessitates multidisciplinary research teams with specialized expertise across diverse domains, including psychophysiology, qualitative analysis, biometric data interpretation, and psychometric instrument design. Resource implication represents another key question (Peng and Annansingh, 2015). Mixed methods typically lead to higher costs - due to technical instrumentation and specialized personnel - and require considerably more time than single methods for the collection, analysis and interpretation of multiple data. In this framework, experience becomes demanding for users who last extended sessions with numerous modalities. These requests might induce fatigue and increase attrition rates.

In conclusion, the appropriate UX evaluation methodologies depend on a weighted choice, not only based on the theoretical strengths of various methods but also considering a practical feasibility within a contextually sensitive approach with specific environments and target populations. This methodological consideration reflects the fundamental principles established throughout these paragraphs, that UX constitutes a multifaceted and dynamic construct that varies significantly across contexts and user groups. This context-dependent perspective becomes especially critical when investigating novel technological interventions in specialized settings, as in the case of the virtual reality-based applications designed for healthcare.

Virtual Reality

The term virtual reality (VR) refers to a simulation of a three-dimensional environment with which users can engage in a seemingly real or physical reality, despite the absence of physical interaction (Riva and Mantovani, 2014, 2019; Riva et al., 2018; Moreno et al., 2019). The capacity of VR to elicit realistic responses depends fundamentally on two concepts: immersion and presence (Slater, 2018, 2009). Immersion represents the objective property of the VR system, determined by the extent to which the technology supports natural sensorimotor contingencies for perception. Higher immersion systems are those that enable more complete perceptual-motor interactions with the virtual environments (VEs). Presence, in contrast, constitutes a subjective psychological phenomenon described as the illusion of 'being there' in the VE (Slater, 2018; Triberti et al., 2025). Higher immersion systems generally facilitate stronger presence experiences, yet the relationship between immersion and presence is probabilistic rather than deterministic. Presence depends on multiple factors beyond technological characteristics, including narrative content, individual user characteristics, and the alignment between user intentions and environmental affordances (Triberti et al., 2025).

Interaction represents the third fundamental characteristic of VR, referring to users' capacity to actively manipulate and navigate the virtual environment in real-time. Together, these three characteristics facilitate VR's capacity to elicit realistic experiences (Riva et al., 2020; Slater, 2009).

When conceptualizing environments along a continuum, VR represents the antithesis of completely authentic environments, as traditionally described by Milgram and colleagues (Milgram et al., 1995). However, based on the possibility of individuals feeling immersed, present, and interacting with the surroundings, VR demonstrates utility to recreate authentic contexts. Figure 1 shows a simplified representation of the continuum.

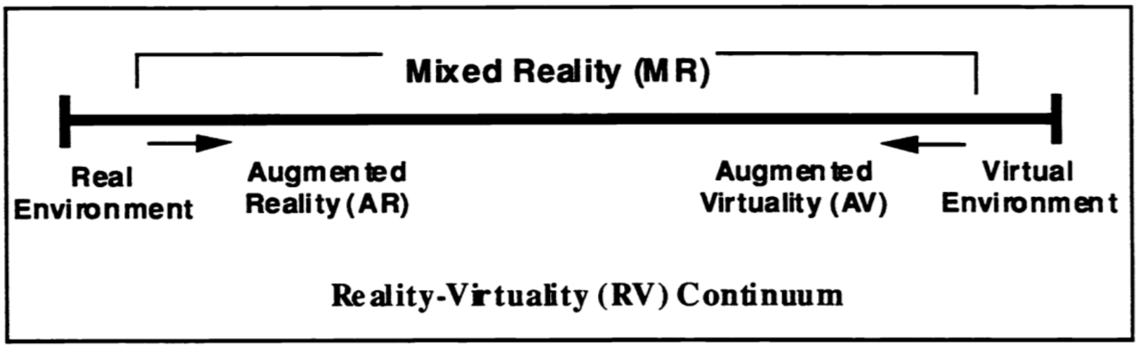

Figure 1. Representation of the VR continuum proposed by Milgram et al. (1995).

VR can be categorized through various frameworks based on the integration with the real world, immersion levels, and the nature of VE.

With respect to real-world integration, since the classification by Milgram et al., the authors established three distinct categories (Kumari and Polke, 2018). (i) *VR* constitutes a fully immersive environment that entirely replaces the users' reality with an alternative one. (ii) *Augmented reality* (AR) superimposes digital content onto the physical world, enhancing the user's perception of their surroundings without fully immersing them in a virtual world. (iii) *Mixed reality* (MR) represents a blending between real and virtual elements, allowing a smooth integration of virtual items into the actual environment through real-time interactions between digital and physical components. In MR contexts, physical elements and the real environment itself play a role, becoming boundaries and creating occlusions for virtual content. Nonetheless, the term *Extended Reality (XR)* is being increasingly employed in recent literature as an umbrella term to include all these categories, acknowledging the wide range of experiences that combine the real and virtual worlds in various ways (Rajamurugu and Yaknesh, 2025).

Literature supported another categorization of VR according to immersion level: (i) non-immersive, (ii) semi-immersive, and (iii) fully immersive (Winkler et al., 2020). This classification is determined by the different employed technologies that significantly impact users' perceived immersion (Cipresso et al., 2018; Slater, 2009) and their sense of presence within the VE (Riva, 2018; Riva and Mantovani, 2014). Non-immersive systems involve desktop-based technologies where users interact with the VE through conventional screens without specialized VR hardware. Interaction is done using gamepads, keyboards, and a mouse as one would for video games. Feedback and engagement with the environment are limited by the screen, so the user's movements are disconnected from the digital space. Consequently, there is a separation of the real and digital space since the user's position and the digital VR space are clearly apart. In semi-immersive VR, the user is surrounded by bigger screens or projection systems while maintaining partial connection to the real world; when they turn beyond the screens, the immersive effect decreases, disengaging users. Typically, the interactions were provided through motion-tracking technologies, which record physical movements and allow for more embodied interaction. Moreover, systems frequently include specialized controllers that make it easier to manipulate virtual objects in the projected space, such as joysticks or steering wheels. The most common situation in which semi-immersive scenarios exist is for simulation purposes (e.g., flight, car, and racing simulators). Fully immersive experiences employ specialized hardware to completely immerse users in VEs. Examples include head-mounted displays (HMDs) and four-wall Cave Automated Virtual Environment (CAVE; Cruz-Neira et al., 1993) configurations. These technologies produce a deep sense of being in virtual space by successfully isolating users from the outside world. Users benefit from complete visual immersion in addition to cutting-edge interaction modalities like hand tracking, haptic feedback devices, and motion capture technology. By using natural movements rather than input commands, users can connect with virtual objects and environments intuitively as the interaction moves toward embodied cognition, where their physical movements directly correlate to actions within the VE.

Regarding the nature of VE, VR can still be classified into three categories (Izard et al., 2018). (i) *Computer-generated VR* where environments are entirely produced with computer graphics. This framework offers three-dimensional, immersive experiences that allow users to interact with fully virtual worlds. In contrast, (ii) *360-degree VR* utilizes panoramic images and videos to create environments. These media, often known as immersive videos or spherical videos, are produced by capturing live action in spherical cameras. They differ from 2D content in the depiction of the entire world rather than just a small portion of it. While 360-degree media may not offer the interactive depth of computer-generated VR, it delivers a heightened sense of realism. On one hand, computer-generated VR has great benefits in customization capabilities tailored to specific objectives and high reproducibility. However, 360-degree media excel in depicting real-life situations with authenticity, allowing researchers to observe how people respond to stimuli that closely correspond to those encountered in actual life circumstances (Mancuso et al., 2024). This is the reason that makes 360-degree an emerging field in the VR panorama. (iii) *Hybrid VR* represents a synthesis of the previous two approaches, where computer-generated elements coexist with 360-degree media, effectively combining the advantages of both methodologies.

Table 2 summarizes the mentioned types of VR classifications.

Table 2. VR classifications

	Categories	**Description**
Integration with the real world	Virtual reality	Reality is entirely replaced by a simulated one
	Augmented reality	Overlays digital content onto the real world
	Mixed reality	Integration of digital and real-world elements in real-time
Immersion levels	Non-Immersive	Low degree of immersion due to desktop-based VR experiences
	Semi-Immersive	Medium degree of immersion while users maintain partial connection to the real world
	Fully Immersive	High degree of immersion. Users are completely in the virtual environment
Nature of virtual environments	Computer-Generated	Environments entirely created by computer graphics

	360-Degree	360-degree images and videos to create a virtual environment
	Hybrid	Computer-generated elements coexist with 360-degree images and videos

Application of virtual reality

Since its inception in the mid-1960s, when Ivan Sutherland attempted to conceptualize VR and designed the first HMD (Sutherland, 1968), this field has experienced remarkable growth. Particularly in the last decade, VR has increasingly attracted significant investment and public interest, culminating in unprecedented accessibility. Contemporary products such as PlayStation VR and advanced HMDs have opened this technology, transforming it from a futuristic concept into a rapidly expanding commercial reality accessible to the average consumer. This evolution has created significant potential in a variety of businesses (Abdelmaged, 2021). The implementation of VR across multiple sectors might be facilitated by the low-cost availability of VR; on the other hand, by its advantageous features (e.g., multi-sensory engagement, risk-free simulation capabilities, and customizable environments). In educational settings, VR creates interactive learning possibilities that enhance student engagement and knowledge retention across diverse subjects (AlGerafi et al., 2023; Tang et al., 2022). The immersive nature of VEs provides students with concrete experiences, limited in traditional learning methods, making it possible to approach students with special educational needs.

VR is also being utilized in the entertainment sector for immersive games and narratives, creating new avenues for audience engagement (Von Itzstein et al., 2024). VR is being used in the tourism industry to provide virtual tours of locations worldwide (Oncioiu and Priescu, 2022), and in the fields of architecture and engineering, it is being utilized for collaborative planning and detailed design visualization (Lemle et al., 2015), which results in more effective project development and execution. Moreover, the versatility of VR technology lies in its ability to simulate unsafe and unusual situations such as medical emergencies, military operations, and dangerous industrial scenarios. In this framework, specialists can learn and train their ability to capitalize on experiences without exposure to real physical risks.

Virtual reality in healthcare

Beyond these applications, VR has gained substantial attention in the healthcare sector due to its versatility and effectiveness as a clinical tool. This technology serves as a resource for clinicians in both therapies (Emmelkamp and Meyerbröker, 2021), and educational training (Cho et al., 2015; Freeman et al., 2017; Izard et al., 2018; Lange et al., 2010; Riva et al., 2019), enabling a customized therapeutic approach.

Nowadays, a proliferation of impactful tools emerges with different therapeutic objectives. These applications span numerous areas, such as mood modulation and psychological intervention (Liu et al., 2020), balance and gait improvement for mobility disorders (Muhla et al., 2022; Yin et al., 2016), limb function enhancement for rehabilitative purposes (Kiper et al., 2018), specific disorder treatment, such as eating disorders (Meschberger-Annweiler et al., 2023), addressing aging-related challenges, whether physiological or pathological conditions (Bruni et al., 2022; Mirelman et al., 2013; Pedroli et al., 2018; Serino et al., 2017). These applications might help to enhance the quality of life, improve healthcare delivery, and reduce social costs throughout the lifespan (Lutz et al., 2008; World Health Organization, 2020). Beyond direct therapeutic interventions, VR applications also support both patients and clinicians in clinical decision-making processes, facilitating health record access and enhancing communication between healthcare providers (Lin et al., 2013). VR demonstrates particular value also in telemedicine applications, significantly enhancing healthcare accessibility for diverse populations facing geographical, physical, or economic barriers (Bruni et al., 2023; Turvey and Fortney, 2017). In this direction, VR aids patients in managing their health status through mobile health applications. These instruments might help patients improve adherence to diet, exercise, and chronic diseases, while also enhancing interactions with caregivers (Galavi et al., 2024), offering new possibilities for patient monitoring and treatments.

From the most common use in exposure therapy for anxiety disorders to innovative diagnostic instruments, VR results in an effective option due to several key advantages, such as enhanced ecological validity. VR enables alignment between the content of the environment and collected stimuli, allowing clinicians to observe behaviors in more naturalistic contexts while maintaining experimental control. Moreover, compared to conventional assessment techniques, the use of activities that closely resemble real-life scenarios yields more ecologically valid measurements of behaviors and functional state, which may result in better prognostic indications (Pieri et al., 2023). Further, clinicians maintain comprehensive control over the therapeutic experience, as VR software enables manipulation of both environmental variables (i.e., scene elements) and user-related parameters (e.g., reaction times as performance metrics). This facilitates standardized stimulus presentation and error-free computation of performance metrics, enhancing the reliability of technological instruments. Finally, the immersive nature of

VR, combined with its capabilities to foster presence and interaction, significantly enhances user motivation and treatment engagement (Bell et al., 2020). Treatment discontinuation is usually prevented by gamification elements – an approach coming from video games – in which the user's entertainment is not the aim of the experience, but it is a means that can prevent dropouts and support compliance with the treatment, as demonstrated by Goumopoulos and colleagues (Goumopoulos et al., 2023)with their "Fun Park Tour" application, designed to support motor and cognitive training for fall prevention in older adults.

Although VR seems to be a beneficial approach, its implementation faces several challenges. First, adverse physiological reactions (i.e., cybersickness) might be experienced during VR exposure. This phenomenon manifests as a constellation of symptoms including oculomotor discomfort (e.g., eye fatigue, visual strain), disorientation (e.g., dizziness, vertigo), and nausea. It is due to a sensory conflict between the visual and vestibular systems (Gallagher and Ferrè, 2018), where the visual system receives information suggesting movement while the vestibular system detects no corresponding physical motion. Vergence and accommodation conflicts may also contribute to these symptoms. It is due to the eyes' convergence to a virtual object, unmatched with the focal distance required to view it clearly (Wang and Lin, 2021). While VR hardware quality continues to improve, as well as the best practices for VR environment design, cybersickness remains influenced by individual susceptibility due to various human factors that cannot be fully controlled (Chang et al., 2020). This risk requires consideration in healthcare contexts where target populations often include older people and patients characterized by many frailties. To this end, despite VR systems demonstrating good usability and acceptability amongst these patients (Tuena et al., 2020), VR applications may not be suitable for all clinical scenarios or patient populations. Thus, the development of effective VR tools for healthcare requires meticulous attention to UX evaluation.

User Experience in VR Applications for Health

As VR tools continue their expansion across multiple healthcare domains, the integration of comprehensive UX evaluation becomes essential to realize their full potential as a safe, accessible, and effective clinical intervention. Such evaluation should encompass multiple dimensions, including usability metrics, cybersickness susceptibility across different patient demographics, cognitive load requirements, and interface intuitiveness for both clinicians and patients. Through systematic UX evaluation, developers can identify potential barriers to adoption, minimize adverse effects, and optimize therapeutic efficacy. In particular, it is especially important to assess UX when designing VR-based applications for clinical usage for several reasons. Firstly, design errors can compromise patient safety. An accurate analysis of the experience ensures that applications are accessible and usable by intended users. According to Strada and colleagues (Strada et al., 2013), this is particularly noticeable in aging people who may have limited familiarity with technology or may feel uncomfortable using it. For instance, a user-friendly interface can boost patient confidence and decrease fear, which can improve adherence to clinical programs and produce more accurate outcomes (Galavi et al., 2024).

The application of VR technology can be further restricted by a number of obstacles, chief among them being cognitive and physical limitations in clinical populations. Working memory, spatial cognition, attention, language, reasoning, motor speed, flexibility, hand-eye coordination, and strength are all areas where these people may struggle, which could make VR interaction challenging. Moreover, declines in sensory, mental, and physical abilities may result from physiological changes linked to aging (World Health Organization, 2021). These difficulties are further threatened by clinical illnesses such as neurological disorders or frailty (Fried et al., 2001; Winblad et al., 2004). From a pragmatic perspective, application features must be tailored to these specific contexts and end-user characteristics. Quality of VR graphics (e.g., resolution, movement of visual elements, shapes, color contrast), synchronization (i.e., the delay between the users' movement into the VR environment and what is expected to be seen), as well as user interface layouts, are elements that could be influenced by the context of the application. While a playful interaction - featuring a great number of colors and sounds - might be crucial for successful gaming software, the same design elements might be perceived as inadequate (i.e., overwhelming or disorienting) in aging people or patients. Wang and colleagues demonstrated this concept in their comparative study of VR environments, both urban and natural (Wang et al., 2022). Unexpectedly, they found urban settings to be more restorative than parks. The authors suggested this could be attributed to the park's design, featuring bright colors and intense saturation that might cause visual discomfort. This finding emphasizes the importance of carefully drawing VR experiments, especially when dealing with complex experiences. The validity and efficacy of these studies can be significantly enhanced by paying close attention to design components that ensure the achievement of the intended therapeutic outcomes.

Further, when users find themselves in a VE that meets their needs, they tend to respond more genuinely and naturally, providing more reliable data for assessment and improving treatment outcomes. Considering UX factors can also help prevent cybersickness. A well-crafted UX can help reduce the risk of worsening patients' already delicate conditions, making VR interventions not just safer but also more comfortable. Moreover, a well-designed UX is essential to the efficacy of VR-based applications through its influence on the sense of presence (Riva et al., 2020). While early conceptualizations emphasized technological features as primary drivers of presence (media presence approach), contemporary research reveals that presence is fundamentally a psychological phenomenon shaped by multiple interconnected factors: content and narrative structure of VEs, individual user characteristics (e.g., age, gender, personality traits, prior experiences, and cultural background), and the alignment between users' intentional structures and environmental affordance (Triberti et al., 2025). This means that compelling storylines and meaningful goals may enhance presence more effectively than mere graphical sophistication, and the same VE may elicit varying degrees of presence across different users based on their individual characteristics and cultural backgrounds. This is reflected in clinical practice, where presence facilitates emotional engagement with therapeutic stimuli in exposure therapy for anxiety disorders (Riva et al., 2020), for example. However, presence in clinical VR also presents challenges, as its intensity can amplify both therapeutic benefits and potential adverse effects, particularly in vulnerable populations. From a UX perspective, optimizing presence requires integrated attention to narrative design, personalization strategies that accommodate individual variations, and interface usability that supports natural interaction, intuitive interaction aligned with users' therapeutic goals (Triberti and Riva, 2016).

Motivation and engagement also impact healthcare outcomes using VR because of the complexity of medical procedures and the need to avoid treatment discontinuation, as well as various aspects of user experience (Berni and Borgianni, 2021; Hassenzahl and Tractinsky, 2006). On the one hand, strong levels of motivation and engagement can significantly increase user happiness by creating favorable emotional bonds with the technology (Van der Heijden, 2004). On the other hand, users are more likely to use an application regularly and as intended when they believe it to be both user-friendly and inherently entertaining, which could improve health outcomes and diagnostic accuracy.

It is evident how evaluating and optimizing UX can help overcome barriers to fully exploit the potential offered by VR, while enabling the customization of clinical offers. Users are resistant to embracing technology that is incompatible with their everyday activities, or that is misaligned with their preferences and needs (Che et al., 2019; van Gemert-Pijnen, 2018). Considering factors like accessibility, inclusivity, and safety, this user-centered approach may produce more targeted and successful interventions. Preliminary evaluation also provides for the identification of potential issues related to clinical conditions that might cause discomfort, allowing for appropriate adjustments. Identifying and addressing early issues during the development process provides a cost-effective advantage compared to expensive redesigns or poor adoption rates.

In conclusion, investing in and improving UX seems to provide emphasis on patient-centered care, which might result in a major improvement in the identification and management of a range of physiological and pathological conditions. This approach improves the caliber and accessibility of healthcare services by guaranteeing that the advantages of VR may be extended to a variety of patient populations while minimizing related risks.

Based on these considerations and acknowledging the limitations in current UX assessment practices, this research intends to clarify the concept of UX in clinical contexts and develop a novel evaluation framework specifically designed to address the needs of patient-centered VR applications. The creation of a specialized questionnaire for evaluating UX in healthcare VR-based applications is the work's final goal. This work will lead the way to establish standardized methodologies for assessing VR interventions that can accommodate the distinctive needs and vulnerabilities of clinical populations, thereby facilitating the comparison between the diverse available VR solutions in healthcare settings. Figure 2 provides a schematic overview of the four-phase research pathway undertaken in this doctoral work, illustrating how each study builds upon the previous to culminate in the validated instrument.

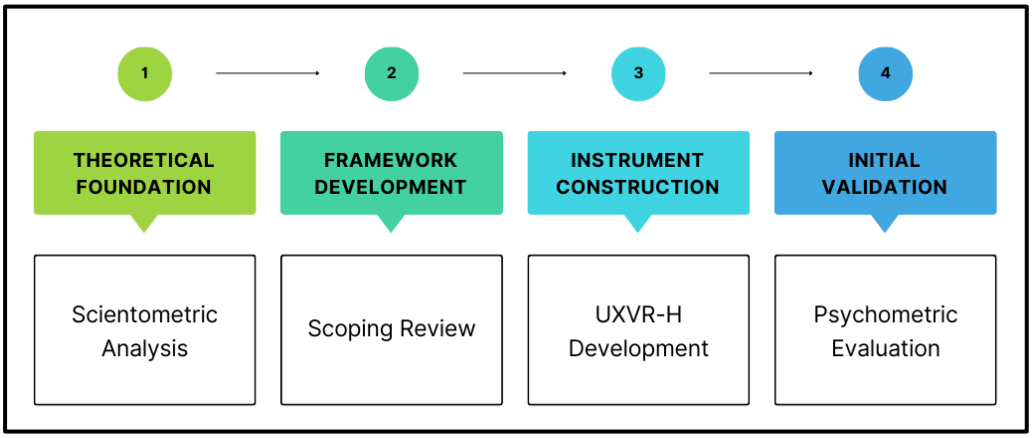

Fig. 2. Schematic overview of the research phases.

EXPERIMENTAL STUDIES

Through the analysis of the experimental studies conducted throughout my doctoral research, this section delineates the methodological trajectory that culminated in the development of the

User eXperience questionnaire for Virtual Reality applications in Healthcare (UXVR-H). The research pathway presented herein enlightens the empirical foundations and theoretical frameworks that informed the questionnaire's conceptualization, design, and validation process.

A Scientometric Analysis of User Experience: Exploring Evolution, Emerging Themes, and Impacts

Introduction

The conceptual framework of UX remains notably fluid, resulting in a fragmented theoretical landscape characterized by diverse focal points, including emotion, affect, experience, value, enjoyment, and aesthetic appreciation (Hart et al., 2012; Hassenzahl, 2013). Since its definition in the field of human-computer interaction and technology design, the foundational concepts underlying UX and their interpretations have continued to evolve dynamically within modern technological contexts.

Within this heterogeneous theoretical background, researchers' consensus converges on the critical importance of UX evaluation, as it fundamentally determines user-product interaction patterns. UX demonstrably influences multiple outcomes, including user satisfaction, motivation, loyalty, technology adoption rates, engagement metrics, operational effectiveness, efficiency parameters, and commercial viability. Conversely, suboptimal UX evaluations frequently precipitate user frustration, product abandonment, and consequent customer attrition (Professor and Khalid, 2006).

It becomes more important than ever to thoroughly assess UX to develop solutions that are not only practical and intuitive but also psychologically fulfilling. This approach results in the creation of products and services that truly satisfy the requirements and expectations of consumers. However, due to its multidimensionality, it is critical to consider different components and theoretical viewpoints of UX, which can differ greatly depending on the objectives of the study, the contextual circumstances, and the disciplinary framework. While theoretical pluralism reflects the inherent richness and multidimensionality of human experience that UX methodologies attempt to capture and enhance through design principles (Li et al., 2022; Liu, 2020), its conceptual heterogeneity frequently results in fragmented understanding and subjective interpretations of UX. This underscores both the complexity and methodological challenges inherent in capturing its comprehensive scope and significance.

Given the rapid pace at which technology is developing and the growing significance of UX, it is essential to understand how the field has changed and identify significant trends. This study employs a quantitative bibliometric approach and network visualization to analyze research trends, identify intellectual hotspots, and explore emerging challenges in the UX domain.

Previous studies have concentrated on particular facets or eras, which has limited their overall relevance throughout the evolution of the field (Li et al., 2022; Liu, 2020). Our methodological framework provides a wide-ranging global perspective of both historical and current research areas, ongoing intellectual paths, and the significant contributions of leading researchers and academic journals to the field. This detailed analysis yields numerous advantages, such as a deeper understanding of the discipline, insights into how various UX-related topics connect and evolve, identification of gaps that need further exploration, the creation of innovative research frameworks, and a stronger position for the field within the larger academic community (Bruni et al., 2025a).

Material and methods

1.2.1 Data collection

We used the scientific database Web of Science (WoS) Core Collection to input the query "User Experience" or UX. The search terms used regarded articles published during the whole timespan covered until February 15th, 2024. WoS core collection is composed of: Citation Indexes, Science Citation Index Expanded (SCI-EXPANDED) − 1970-present, Social Sciences Citation Index (SSCI) − 1970-present, Arts and Humanities Citation Index (A & HCI) − 1975-present, Conference Proceedings Citation Index-Science (CPCI-S) − 1990-present, Conference Proceedings Citation Index-Social Science & Humanities (CPCI-SSH) − 1990-present, Book Citation Index– Science (BKCI-S) − 2009-present, Book Citation Index – Social Sciences & Humanities (BKCI-SSH) − 2009-present, Emerging Sources Citation Index (ESCI) − 2015-present, Chemical Indexes, Current Chemical Reactions (CCR-EXPANDED) − 2009-present (Includes Institut National de la Propriete Industrielle structure data back to 1840), Index Chemicus (IC) − 2009-present.

1.2.2. Data analysis

The search strategy yielded 27,082 records, each containing fields including: author, title, abstract, and all references, which were required for the analysis. The analytical framework employed two primary software tools: the R-based bibliometrix package and CiteSpace 6.1.R3 Advanced (operating on Java Runtime v.8 update 91, build 1.8.0 91-b15) for conducting statistical analyses and generating visualizations of co-citation and citation networks, respectively.

Data analysis entailed descriptive analysis, citation analysis, co-citation, co-author, and co-word analysis. Moreover, visualization techniques were integrated to enhance the clarity of findings.

Results

Descriptive analysis reveals that UX research emerged as a scientific field in 1985, with publication output growing substantially from 5 initial results to a peak of 2,888 documents in 2023. Among the total publications, proceeding papers constitute the predominant document type, accounting for 50% of all papers in the dataset publications.

Subject category statistics from WoS indicate computer science as the leading research domain, followed by engineering. This disciplinary focus is confirmed through source analysis, which identifies IEEE Access as the most productive publisher in UX research with a total of 308 documents, followed by the International Journal of Human–Computer Interaction (203 documents), Applied Sciences-Basel (201 documents), Sensors (189 publications), and Multimedia Tools and Applications (149 documents). Moreover, according to Bradford's Law, IEEE also appears as the most relevant journal in the field of UX, providing the core of UX publications. Notably, several cross-disciplinary journals appear among the top 20 most frequent sources since the early 2000s, demonstrating increased interest in UX themes. These include Sustainability, Journal of Medical Internet Research, Frontiers in Psychology, and PLoSONE, as shown in Figure 3.

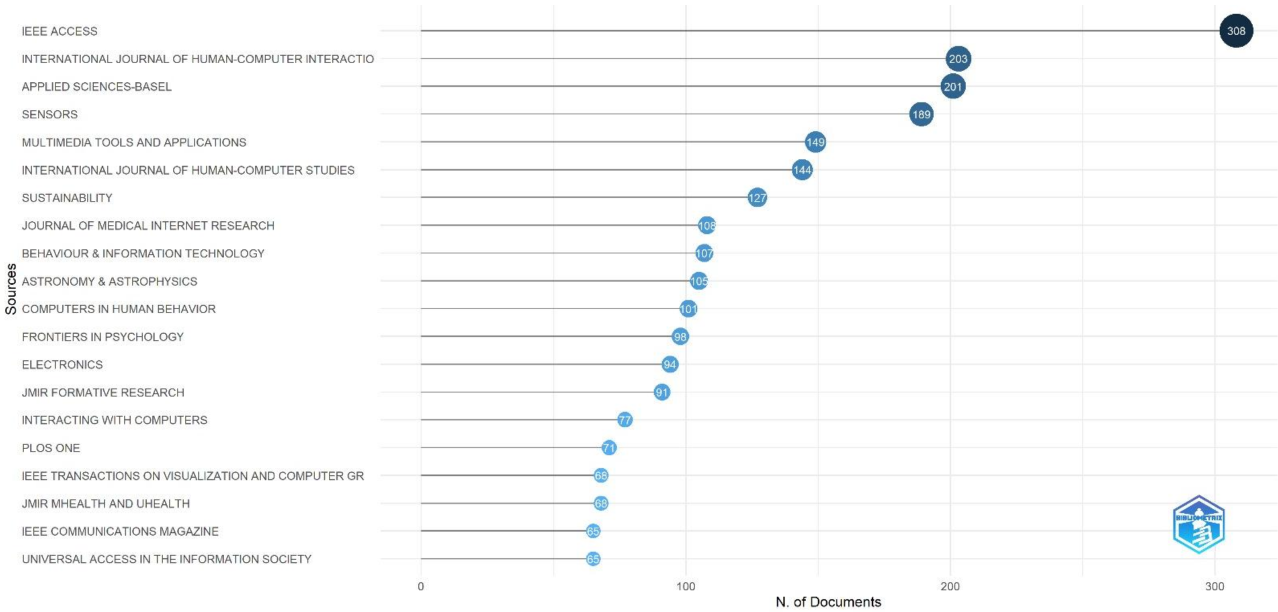

Figure 3. Most relevant sources based on the number of documents published.

Citation analysis

Citation tracking was conducted based on publications, and sources to enable the assessment of journals' impact and community recognition within the research domain. The h-index analysis provides insights into source importance based on citation frequency within the dataset. From this perspective, Computers in Human Behavior- a journal examining computer use from psychological and related disciplines perspectives – emerges as the most influential journal with an h-index of 39. The International Journal of Human-Computer Studies, which focuses on the design and use of interactive computer technology in the fields of human factors and ergonomics, artificial intelligence, industrial and manufacturing engineering, follows with an h- h-index of 36. Subsequent rankings include Astronomy & Astrophysics, IEEE Communication Magazine, and Journal of Medical Internet Research, followed by h-indexes of 34, 32, and 29, respectively, as shown in Figure 4.

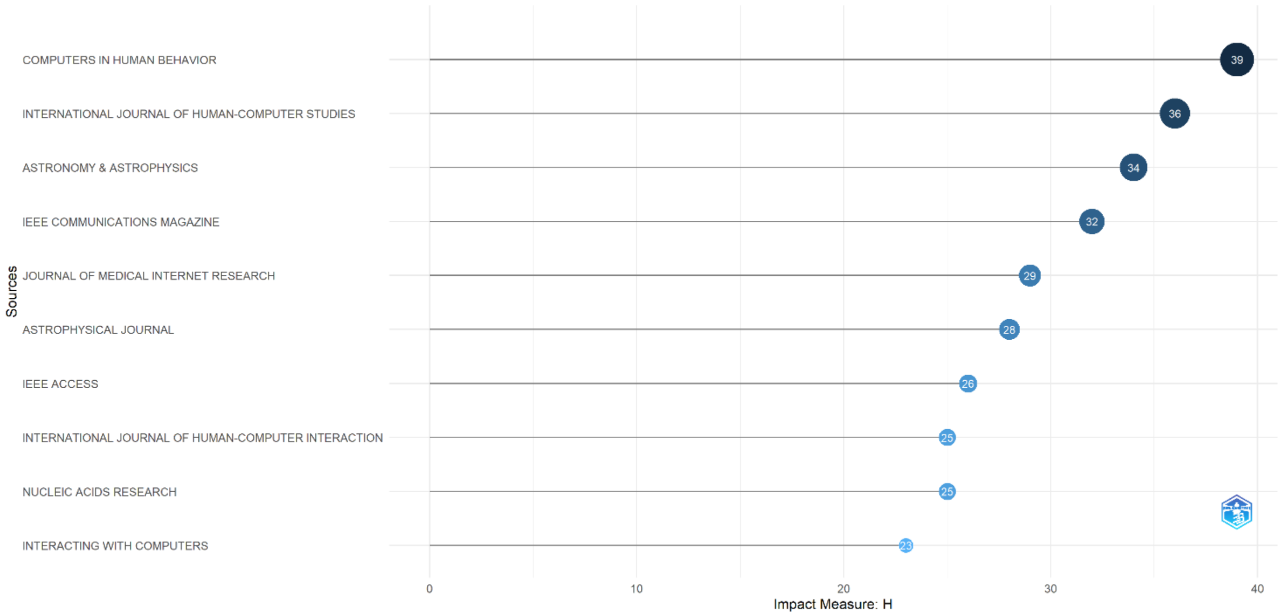

Figure 4. Top 10 list of the most influential journals based on h-index.

Document-level analysis reveals that Hassenzahl's studies, published in Behaviour & Information Technology (Hassenzahl and Tractinsky, 2006), and Interacting with Computers (Hassenzahl et al., 2010) represent the most cited documents with 781 and 352 local citations, respectively (see Figure 5).

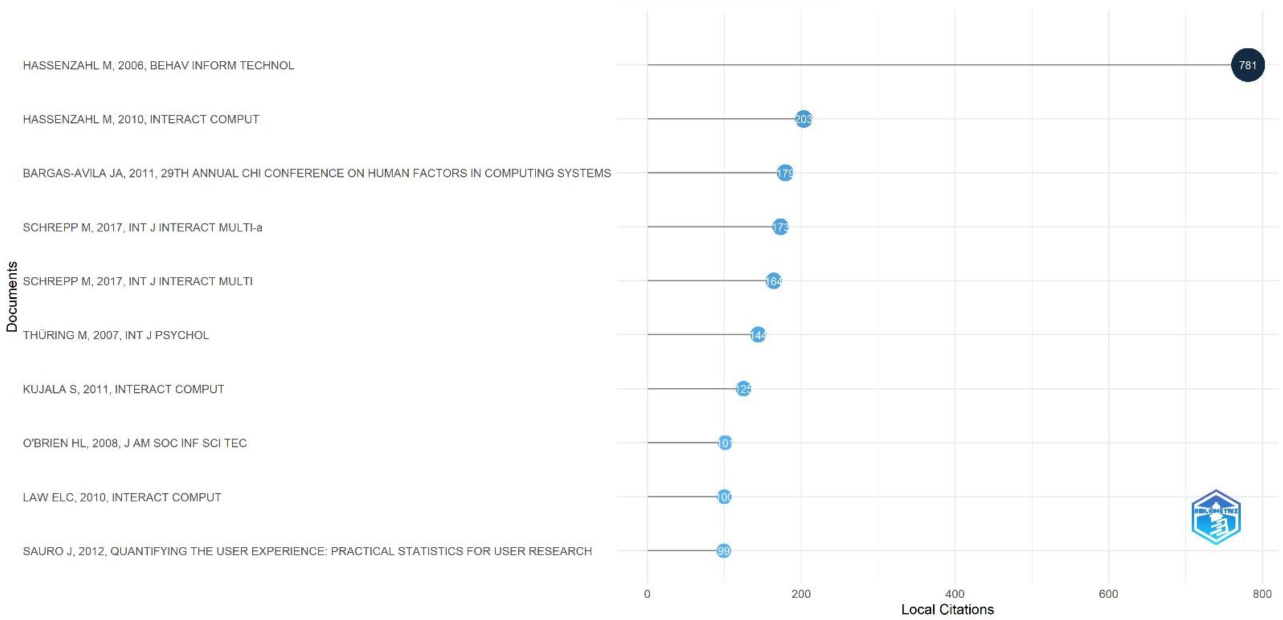

Figure 5. Top 10–most cited papers based on the number of local citations.

Co-citation analysis

Co-citation analysis was performed through sources (co-cited journals). By examining how frequently two journals are cited together in other works, co-citation analysis assesses the relationship between them. The stronger the thematic relationship between the two items is thought to be, the more often they are co-cited. The structural characteristics of co-citation networks were assessed using two metrics: betweenness centrality and PageRank algorithms. Betweenness centrality measures a node's ability to connect various network components along various pathways (Brandes, 2001; Freeman et al., 2017). Higher betweenness centrality values indicate critical roles in promoting information flow across network segments. The PageRank algorithm assessed the relative importance of sources, authors, and documents within citation networks by evaluating their relevance based on received citations. Detailed algorithmic descriptions are provided in Aria and Cuccurullo (2017) (Aria and Cuccurullo, 2017).

Journal co-citation analysis

The journal co-citation analysis identifies five distinct clusters of sources with significant as shown in Figure 6. The red cluster group sources related to informatics, engineering, and computer science, including IEE Access, Sensor-Basel, and Multimedia and Tools and Applications. IEE Access ranks highest in this cluster with a betweenness centrality of 155.66. Notably, this cluster includes arXiv. It is an online open-access repository of preprints and eprints works in the fields of mathematics, physics, astronomy, electrical engineering, computer science, quantitative biology, statistics, mathematical finance, and economics. They were approved for posting after moderation (but not peer review). The blue cluster encompasses journals focused on physics, astronomy, earth, and planetary sciences, with Monthly Notices of the Royal Astronomical Society as the top-ranked journal (betweenness: 0.01). The green cluster groups journals on computer networks, communications, and computer science applications in electronics and engineering, with IEEE Communications Magazine leading this cluster (betweenness: 4.35). The yellow and orange clusters each contain single journals—Nature and Science, respectively—with betweenness values of 126.11 and 124.05.

The purple cluster represents the largest grouping, encompassing multidisciplinary sources including thesis, miscellaneous publications, and social science journals. Lecture Notes in Computer Science – a miscellaneous collection - tops this cluster with a betweenness of 14.58. Lecture Notes in Computer Science also achieves the highest PageRank (0.06), based on citation count, followed by International Journal of Human-Computer Studies (PageRank: 0.05) and Computers in Human Behavior (PageRank: 0.05).

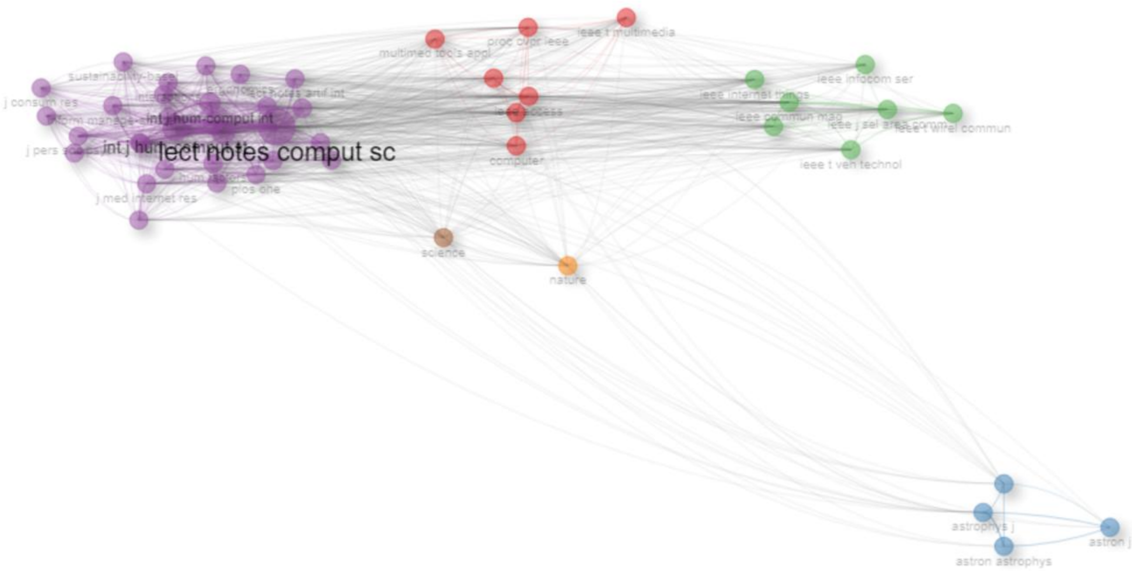

Figure 6. Co-citation network of journals

Co-words analysis

Co-word analysis was implemented to examine the conceptual architecture of the research field. This approach utilized the most significant keywords from documents to construct semantic maps, facilitating comprehension of the cognitive structure within the UX domain; specifically, revealing how key concepts, themes, and ideas are organized and interconnected (Callon et al., 1983).

Besides the "user experience", other relevant keywords appear in the dataset. The appearance of user experience keywords (5,329 times) is followed by "usability" with 1,203 occurrences (Figure 7). The prominence of these terms might reflect their historical synonymous usage and ongoing conceptual confusion in literature. Additional significant keywords include "virtual reality" (1,122 occurrences), "augmented reality" (688 occurrences), "human-computer interaction" (404 occurrences), and "machine learning" (381 occurrences), highlighting UX research's strong connection to emerging technologies.

Figure 7. Most frequent keywords.

Analyzing the frequency of keywords over time (i.e., trend topic analysis), we provide insights into future research directions and disciplinary evolution within UX research. User experience became a trending topic between 2016 and 2019. Initially, the most frequent terms were related to astronomy and physics fields, while the first decade of the twenty-first century witnessed a shift toward computer science and informatics, culminating in the emergence of new technology-related themes. In particular, the research field seems to be divided into three distinct phases: the pre-digital era, which is characterized by early web technologies and astronomical keywords; the digital and web technologies; and the immersive era, which is marked by increasingly common keywords such as "augmented reality," "virtual reality," "artificial intelligence," and "metaverse," which reflect the intrinsic relationship between emerging technologies and UX research. Table 3 reveals this transition from pure scientific research to technology application fields centered on user experience, with the period 2020-2023 marking heightened interest in artificial intelligence and the metaverse.

Table 3. Most frequent keywords with temporal distribution quartiles (Q1, median, Q3) showing the chronological evolution of research terms in the dataset.

Keywords	frequency	year_q1	year_med	year_q3
user experience	5326	2016	2019	2021
virtual reality	1122	2019	2021	2022
augmented reality	688	2017	2020	2022
design	379	2014	2017	2020
evaluation	293	2014	2018	2021
artificial intelligence	223	2020	2022	2023
performance	136	2013	2016	2020
lte	84	2014	2015	2017
novae, cataclysmic variables	61	1998	2002	2010
metaverse	41	2022	2023	2023
circumstellar matter	40	2002	2010	2013
accretion, accretion disks	39	2000	2004	2010
semantic web	34	2012	2014	2016
web services	31	2008	2012	2018
stars : activity	27	2001	2003	2006
iptv	26	2008	2011	2014
ajax	21	2008	2009	2011
knowledge management	20	2010	2013	2017
polarization	15	1994	2000	2020
radio continuum : stars	14	1998	1999	2002
binaries, eclipsing	11	1994	1995	1996
ultraviolet, stars	11	1995	1996	1997
stars : individual : ux ari	11	1998	2001	2003
mpeg-21	8	2006	2008	2009

photometry	7	2000	2006	2018
stars, late-type	6	1992	1994	1997
infrared : stars	5	2006	2007	2007

Moreover, we examine how keywords have changed over the last ten years. User experience, virtual reality, and usability are the mature and well-developed core of the UX field, with high centrality (i.e., the degree of interconnection with other thematic areas) and density (i.e., the degree of development/field importance), as shown in Figure 8. These motor themes play a key role in propelling the field's development, exhibiting both theoretical maturity and strong integration within the research. Fields related to machine learning and artificial intelligence emerge as basic themes with highly relevant but underdeveloped topics that represent promising growth opportunities for future UX research. Their positioning suggests that while these technologies are gaining recognition as central to UX research, they have not yet reached full theoretical and methodological maturity within the domain. Niche themes, including mhealth, COVID-19, and smartphones, represent specialized areas that have achieved substantial development but remain peripheral to the main research domain. Themes like big data, cloud computing, deep learning, and quality of experience are positioned in the emerging or declining quadrant, creating an interesting duality. Their moderate centrality and low development density might point to either new themes that haven't caught on in UX research yet or technologies whose initial excitement for study has faded as they have become more popular. This positioning must be carefully monitored to distinguish between declining relevance and emerging opportunities. Notably, quality of experience appears strategically positioned for potential transition, suggesting its evolution toward greater centrality in UX research, reflecting the field's increasing emphasis on measurable, comprehensive experiential assessments beyond traditional usability metrics.

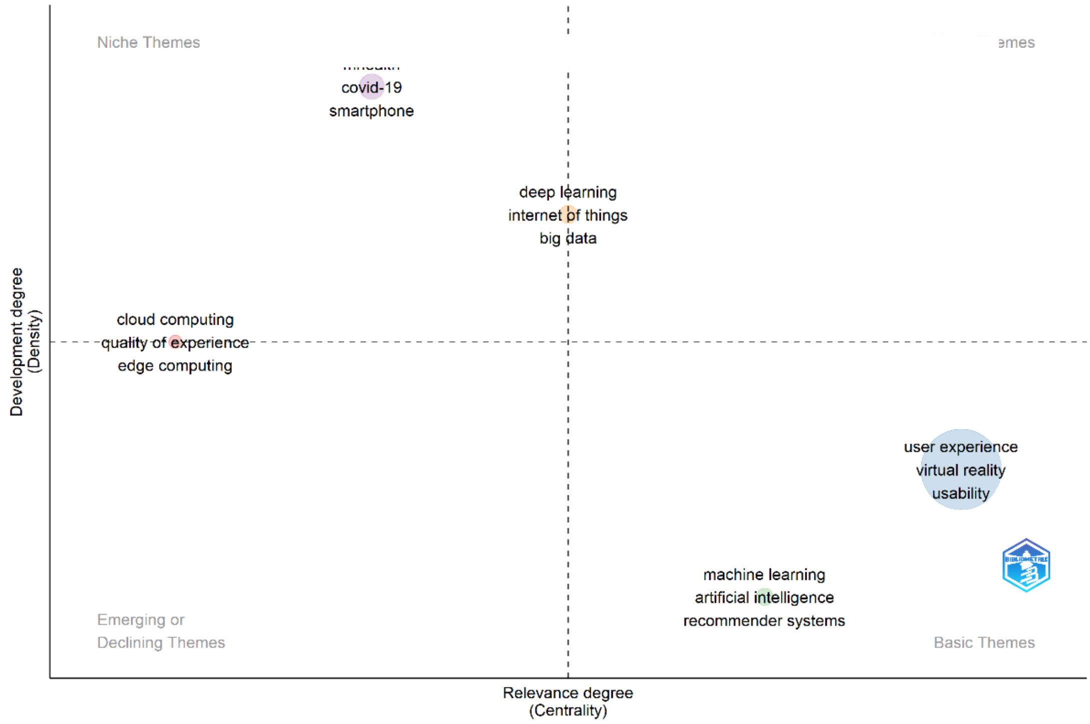

Figure 8. Thematic evolution map from 2014 to 2024.

Discussion

This work reveals a structured panorama of the UX field evolution, characterized by exponential growth in scientific production from 1985 to 2023 and increasing interdisciplinarity. With over 27,000 documents identified, the predominance of publications in computer science and engineering, together with the significant presence of multidisciplinary journals among the most productive sources, confirms the interdisciplinary nature of UX research. This variety of sources, which includes technical publications as well as specialized journals in psychology and medicine, illustrates how computer science and engineering methodologies are integrated with clinical behavioral research and how UX necessitates technical competencies to assess the interface with digital devices. Significantly, the field's conceptual development can be divided into three stages: the pre-digital period, which was dominated by themes related to physics and astronomy; the advent of digital and web technologies; and the present immersive period, which is marked by virtual and augmented reality, artificial intelligence, and the metaverse. This change is based on an expanded perspective that moves beyond technical aspects and includes a focus on people, as shown by keyword analysis, where "usability" was replaced with "user experience".

Citation analysis reveals that Hassenzahl's contributions (Hassenzahl and Tractinsky, 2006) consolidate the theoretical foundations of the UX domain. Hassenzahl's 2006 work stands out as a foundational contribution, representing the most cited document in the field and establishing key theoretical frameworks that continue to structure contemporary UX research.

Thematic analysis indicates strategic implications for the future advancement of the field. The placement of artificial intelligence and machine learning as core themes suggests that there is potential for a UX-specific framework for these technologies. The merging of topics related to VR and artificial intelligence into a single central theme also shows that UX research is moving away from traditional usability problems and toward more complicated interactions between people and technology. The strategic placement of "quality of experience" may also show that the field is becoming more interested in a more complete evaluation panorama that looks at not only functional usability but also user satisfaction and emotional engagement. The convergence of new technologies and clinical applications in UX research - as demonstrated by the inclusion of medical and psychology journals among the productive sources - represents another critical challenge. The rapid adoption of VR in clinical settings, particularly within neuropsychological domains (Alashram et al., 2019; Cavedoni et al., 2022; D'Cunha et al., 2019; Pieri et al., 2023), is hindered by the absence of customized UX evaluation instruments. The unique experiential components of therapeutic VR-based applications encompass aspects such as presence, immersion, and therapeutic efficacy, which surpass traditional focuses on usability metrics. The maturation of UX research, the consolidation of emerging technologies, and the growth of clinical applications all come together to create major opportunities for the development of specialized frameworks that can effectively address the complex nature of UX in therapeutic contexts, ultimately leading to more user-centered and effective interventions.

In conclusion, the analyzed patterns show a mature field in its theoretical basis but are dynamic in integrating the changes of new eras. Nowadays, this is expressed by a fundamental shift away from purely technical considerations and toward a more nuanced, human-centered approach that acknowledges the complexity of modern digital experiences.

How to evaluate User Experience in Digital Health? A Scoping Review of Tools in Virtual Reality Applications

Introduction

The rapid advancement of digital technologies has fundamentally transformed clinical practice across multiple healthcare domains, with Virtual Reality (VR) emerging as one of the most promising technological innovations in contemporary medicine. In the neuropsychological field, this transformation has been particularly pronounced, with growing interest in developing

VR-based tools for both clinical assessment and therapeutic intervention, alongside an expanding body of literature investigating their effectiveness in patient care (Pieri et al., 2023; Riva et al., 2020). VR application spans a wide range of clinical areas, including mood modulation (Liu et al., 2020), balance and gait improvement (Muhla et al., 2022; Yin et al., 2016), limb function enhancement (Kiper et al., 2018), treatment of eating disorders (Meschberger-Annweiler et al., 2023), and management of aging-related conditions, both physiological and pathological (Bruni et al., 2023; Mirelman et al., 2013; Pedroli et al., 2018; Serino et al., 2017). Moreover, VR demonstrates potential for enhancing quality of life, improving healthcare delivery, and reducing social costs throughout the lifespan(Lutz et al., 2008; World Health Organization, 2021). This technology can also facilitate health record accessibility and improve communication between patients and clinicians (Lin et al., 2013). Additionally, VR supports patients in managing their health status through diet and exercise programs, chronic disease management, and improved caregiver interactions (Galavi et al., 2024), expanding possibilities for both patient assessment and treatment interventions. The integration of VR technology into clinical practice represents a paradigm shift from traditional paper and pencil or computerized methods toward more immersive, interactive, and ecologically valid approaches. These technological advances offer unprecedented opportunities to create controlled yet realistic environments that can simulate real-world scenarios while maintaining the precision and standardization required for clinical applications. However, the successful implementation of such technologies necessitates careful evaluation of their impact on users, particularly given the unique characteristics and needs of clinical populations. Thus, evaluating UX in a clinical context appears crucial, as suggested by previous bibliometric analysis (Bruni et al., 2025a) that confirms the growing intersection between UX studies and clinical applications, highlighting VR as an emerging central theme.

Personal traits, contextual factors, presence, immersion, comfort, acceptance, and the complex interplay between technological functionality and therapeutic efficacy are just a few of the many dimensions that encompass UX evaluation, which goes beyond conventional usability metrics. Attention to the specific attributes of end-users and the nuances of their usage context is crucial to effectively delineate the concept and facilitate accurate assessment, as well as optimal design of technological instruments. These experiential factors can impact treatment engagement, compliance, and therapeutic outcomes, particularly for patients with neurological conditions and older adults. The consideration of UX in the design and early implementation phase of VR-based clinical tools might be a crucial component that differentiates successful, user-accepted applications from those that fail to achieve clinical adoption. A user-centered approach can result in more focused and successful interventions, enhancing patient outcomes and clinical workflow integration.

Despite the recognized importance of UX evaluation and its extensive conceptual development, current approaches to UX assessment in the clinical context remain fragmented. Various theoretical models emphasized different aspects such as technological features, emotional responses, perceived value, usability, and aesthetic considerations (Law et al., 2008), yet no comprehensive and systematic approach to its evaluation.

This review aims to present the available tools for evaluating UX in healthcare settings. While diverse methodologies exist for UX assessment, this study focuses primarily on questionnaires. The relative simplicity, rapidity, and ease of administration make them optimal to capture UX data in healthcare settings, where patient concerns and time constraints are crucial.

Aim

The primary objective is to synthesize existing literature and provide a comprehensive perspective on the convergence of VR technology, UX, and healthcare. While literature exists within each of these domains individually, to our knowledge, no systematic review has specifically addressed their intersection. This focus is particularly relevant given the future trajectory of UX research and its growing application in healthcare contexts (Familoni and Odetunde Babatunde, 2024). Specifically, this review analyzes the conceptual foundations of UX evaluation and examines the critical importance of systematic UX assessment in developing innovative clinical VR tools.

Methods

We proposed a scoping review methodology as the most appropriate strategy since it allowed for a thorough examination of the body of current literature, providing a broad perspective of the field, and allowing us to synthesize a variety of findings to identify and analyze knowledge gaps (Peters et al., 2015).

Literature search

This review was conducted on February 2nd, 2024, using three databases: PubMed, Web of Science, and Embase. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, extension for Scoping Reviews) guidelines (Tricco et al., 2018) were followed in reporting findings.

The bibliographic search employed the following terms combined with boolean operators: (measurement OR assessment OR evaluat*) AND ("user experience" OR UX) AND (instrument OR questionnaire OR scale) AND (VR OR "virtual reality" OR "360° video*" OR "360° image*" OR "360-degree video*" OR "360-degree image*" OR "spheri* video*" OR "360° technology" OR "360-degree technology" OR "360 degree technology" OR "immersive video*" OR "immersive image*" OR "360 degree medi*" OR "virtual-based" OR virtual OR immersive OR simul*).

The search strategy deliberately focused on explicit "user experience" and "UX" terminology rather than broader related constructs (e.g., usability, acceptability, user satisfaction). This decision was theoretically and methodologically motivated. First of all, preliminary searches revealed that including broader terms generated excessive results with limited relevance to a comprehensive experiential assessment. Then, our review aimed to identify how researchers explicitly conceptualize and operationalize UX as a distinct construct within healthcare VR contexts.

Results of the search strategies were subsequently refined through examination of the titles and abstracts of the identified records.

Articles screening

All publications identified through the literature search were imported into Ryyan software, where duplicate entries were automatically eliminated. After duplicate detection, two independent reviewers conducted the literature screening following a systematic two-stage process: initially, titles and abstracts underwent analysis, with full-text articles retrieved when at least one reviewer determined that an article potentially satisfied the inclusion criteria.

Eligibility was then confirmed by a thorough, comprehensive full-text examination. In case of disagreements among the reviewers, cooperative discussion was used to reach an agreement. I had consulted as a third independent reviewer to provide arbitration and make a final decision when dialogue fails to produce resolutions. A hierarchical set of criteria was developed and applied consistently throughout both the title/abstract screening and full-text screening phases:

1. Publications written in the English language;
2. Empirical research studies (excluding reviews, meta-analyses, conceptual papers, and study protocols);
3. Inclusion of human participants;

4. Adult populations (excluding pediatric and adolescent cohorts; no upper age restrictions);
5. Implementation of virtual reality or 360-degree media technologies within healthcare contexts;
6. Development or application of a questionnaire designed to evaluate user experience (irrespective of the specific operationalization of UX constructs).

Publications failing to meet any of the aforementioned conditions were systematically excluded based on the corresponding selection criterion that was not fulfilled.

Data extraction and synthesis

A designated researcher supervises the accuracy and comprehensiveness of the entire methodological procedure. Upon completion of the screening phase, the resulting articles were compiled, and data extraction was conducted using Excel spreadsheets. The following variables were extracted: participant populations involved in the UX evaluation, VR applications (including task categories, virtual environments, and technological platforms utilized), and UX evaluation (i.e., procedures and type of questionnaire).

Results

The screening process resulted in 17 articles that met the inclusion criteria, as depicted in Figure 9, which illustrates the comprehensive literature review workflow.

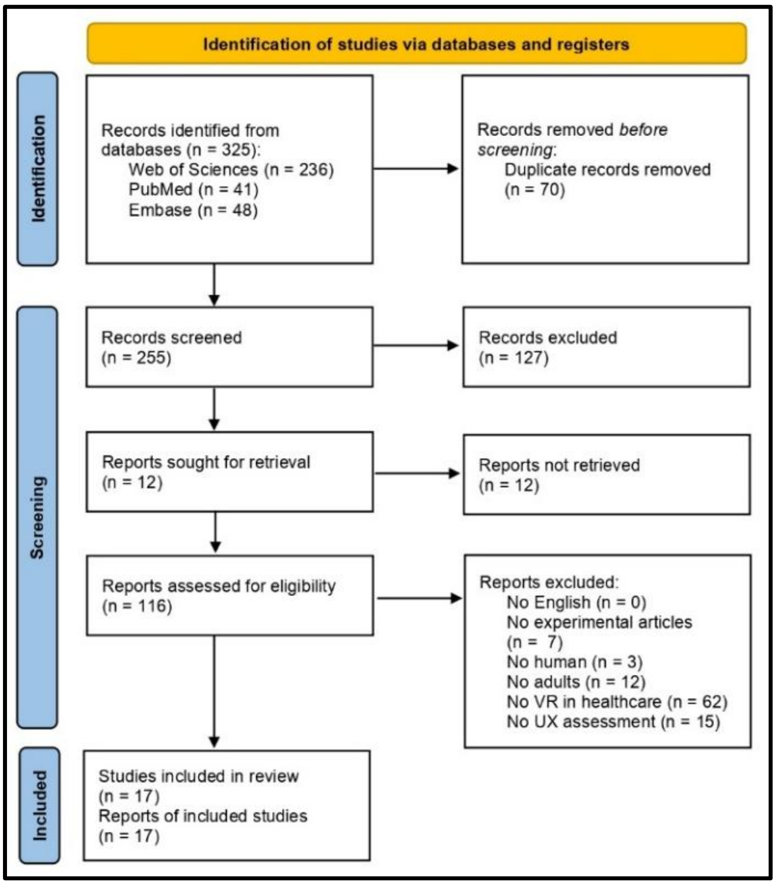

Fig. 9. Flow chart of the literature review process.

Our findings are presented and analyzed according to the variables systematically extracted during the data synthesis phase, as outlined in the preceding methodology section. Table 4 provides a comprehensive overview of these extracted variables. The subsequent sections present detailed analyses of each examined feature through dedicated paragraphs that explore the specific characteristics and findings related to each variable.

Table 4. The table summarizes studies and extracted information.

Paper	Population	VR applications	VR technology	UX assessment procedure	UX evaluation
(D. Liao et al., 2018)	10 healthy adults	Interactive scenes for depression assessment in VR environment composed of natural environments (e.g., green grove) and clinical scenes similar to a real diagnosis	HMD to deliver immersive 360° content	After a VR session, users were asked to fill out UX questionnaires.	ad hoc survey

		environment with images, language, and defined gestures of characters.			
(Hamzeheinejad et al., 2019)	45 healthy adults	Natural landscapes (grassland, forest, streamland, and beach) where users performed walking abilities. It can include a virtual trainer as a companion during the walks to increase encouragement.	HMD to deliver immersive VR content integrated with a platform to walk (Cardiostrong Cross trainer)	Participants performed three consecutive days of intervention at the end, they performed UX questionnaire.	UEQ
(Kourtesis et al., 2019)	40 healthy adults	3 different games in which participants have to perform several occupations, such as a cook, car mechanic, and an office worker (Job Simulator), complete several mini-games like slingshot, or longbow (The Lab), and complete several imaginary home chores (Rick and Morty).	HMD and motion sensors to deliver immersive VR content	Participants performed 3 sessions of intervention (one per week), then they performed UX questionnaires.	Virtual Reality Neuroscience Questionnaire (VRNQ)
(Chen et al., 2020)	23 healthy adults	Users performed three exergames: (i) wall dodging (players must rapidly maneuver through holes in approaching walls); (ii) fruit picking (three fruits are randomly displayed on the screen, and users should try to catch a specified fruit by moving the body from side to side); (iii) rats stomping (participants score points by	Motion interaction sensors connected with a screen displaying sound feedback and motion interaction.	After a VR session, users were asked to fill out UX questionnaires.	User Experience Questionnaire (UEQ-S) short version

		stepping on rats that emerge from the holes).			
(Borgnis et al., 2021)	66 healthy adults	EXIT 360°: domestic photos as virtual environments in which participants have to perform seven subtasks of increasing complexity (e.g., observe a map and choose the right path to exit to the house, explore a room and select the correct person according to a specific instruction, solve a rebus, memorize a sequence of numbers and report them in reverse)	HMD to deliver immersive 360° content	After a VR session, users were asked to fill out UX questionnaires.	User Experience Questionnaire; ICT - SOPI; Flow Short Scale (three items); Intrinsic Motivation Inventory (subscale enjoyment—four items)
(Borgnis et al., 2022)	27 PD patients and 27 healthy adults	EXIT 360°: domestic photos as virtual environments in which participants have to perform seven subtasks of increasing complexity (e.g., observe a map and choose the right path to exit to the house, explore a room and select the correct person according to a specific instruction, solve a rebus, memorize a sequence of numbers and report them in reverse)	HMD to deliver immersive 360° content	After a VR session, users were asked to fill out UX questionnaires.	User Experience Questionnaire; ICT - SOPI; Flow Short Scale (three items); Intrinsic Motivation Inventory (subscale enjoyment—four items)
(Huygelier et al., 2022)	15 healthy adults and 7	Natural outdoor scene (vegetable garden, lake, and forest) presented in one of	HMD to deliver immersive VR content	In a pre-training phase, cybersickness symptoms were	SSQ ad hoc User Experience scale

	stroke patients	three lighting conditions (day, evening, or night). Users have performed visual discrimination tasks.		measured. After a VR session, users were asked to fill out all the UX questionnaires.	
(Mondellini et al., 2022)	15 MCI patients	Virtual supermarket filled with grocery items, and the cash-register scene, in which the users can pay for the items they have picked. Users had to shop, picking all the items presented on a list and putting them in a cart.	HMD to deliver immersive VR content	In a pre-training phase, cybersickness symptoms and intention to use the virtual reality system were measured. After a VR session, users were asked to fill out all the UX questionnaires.	Simulator Sickness Questionnaire (SSQ); International Test Commission—Sense of Presence Inventory (ITC-SOPI); Technology Acceptance Model 3 questionnaire
(Pardini et al., 2022)	20 healthy adults	360° natural environments characterized by auditory stimuli. Users may choose one of the proposed realistic scenarios (mountain, marine, and countryside environments) where they experience relaxation training.	HMD to deliver immersive 360° content	After a VR session, users were asked to fill out UX questionnaires.	ad hoc questionnaire
(Goumopoulos et al., 2023)	23 healthy older adults and healthcare experts	15 games from the GAME2AWE platform organized into two themes (Life on a Farm and Fun Park Tour). Activities are themed around farming (e.g., seeding and fertilizing a field, crop harvesting, insect repelling, and selling crops or purchasing resources)	GAME2AWE platform composed of movement tracking sensors and an HMD to deliver immersive VR content	UX measures were assessed in a pre-pilot phase by experts and final users, and in a pilot phase by end-users.	Short interviews and discussions with seniors and experts; System Usability Scale (SUS) questionnaire; Virtual Reality Sickness Questionnaire (VRSQ); Acceptance and

		and in a fun park (activities that require physical and cognitive skills), respectively.			Use of Technology (UTAUT) model questionnaire
(Liu et al., 2023)	30 elderly patients with post-stroke cognitive impairment	16 games grouped in 3 categories in wich users performed life skills training (cooking, cleaning a window, crossing a road, watering flowers), exergames (playing sqash, shooting antiaircraft guns, flying gliders, playing baseball) and entertaining games (bracking eggshells, swatting insects, lighting fireworks, whack a mole, pumpong un a ballon, flying a Kongming lantern, Fruit Ninja, bubble jab).	HMD and motion sensors to deliver immersive VR content	Participants were involved in 6 weeks of training, then the UX was evaluated.	self-reported questionnaire
(Lorentz et al., 2023)	35 patients with brain injury resulting in attention deficits	VR traveller: an attentional dysfunctions program composed of several modules, within the context of a virtual journey around the world.	HMD to deliver immersive VR content	After the testing modules (one time), participants filled out the questionnaires and were interviewed about their experience.	User Experience Questionnaire (UEQ)
(Meschberger-Annweiler et al., 2023)	60 college women	A room without any furniture except for a large mirror located in front of the participant and two boxes placed on the floor beside the participant. They see their whole image in the mirror (avatar) and	HMD and motion sensors to deliver immersive VR content.	After a VR session, users were asked to fill out UX questionnaires.	System Usability Scale (SUS)

		perform an attentional bias modification task procedure.			
(Rojo et al., 2023)	23 healthy adults and 22 with lower limb disorders	Pedaleo VR: three scenarios (sky, canyon valley, and sailing environment) in which participants had to control a vehicle by pedaling. Vehicles might be a light aircraft or a fishing vessel.	cycle-ergometer and HMD to deliver immersive 360° based content	After a VR session, users were asked to fill out UX questionnaires.	Intrinsic motivation inventory (three subscales); Credibility and expectancy questionnaire (CEQ); Simulator sickness questionnaire (SSQ); Presence questionnaire (PQ); 18-item short scale of Game user experience satisfaction scale (GUESS); System usability scale (SUS)
(Tölgyesi et al., 2023)	21 ADHD patients and 21 neurotypical participants	Zenctuary VR: a small garden surrounded by a forest and a gently flowing river. Users had to please in virtual space, interact with the environment as they wanted (various types of auditory, visual, and tactile responses to the users' actions were generated), or even just look around the garden populated with flowers, birds, cloud, batterflies, plants.	HMD to deliver immersive VR contents	After a VR session users were asked to fill UX questionnaires.	User eXperience in Immersive Virtual Environment questionnaire (UEIVE)
(Colombo et al., 2024)	14 patients with COPD	Virtual Park: park with graphical and audio elements typical of a natural environment (e.g., flowers, trees,	Cycle-ergometer and and a wide projected screen in front	Participants were involved in two weeks of training,	Modified version of the User Experience Questionnaire (UEQ);

		birds) where users simulate a bicycle ride.	of the bicycle to provide semi-immersive VR experience.	then UX was evaluated.	Short Flow State Scale 2 (SFSS-2)
(Espinola et al., 2024)	15 healthy adults	Hospital scenario simulating the experience of a morally challenging event related to the COVID-19 pandemic.	HMD to deliver immersive VR content	After a VR session users were asked to fill UX questionnaires.	Igroup Presence Questionnaire (IPQ); dropout rate; qualitative responses provided during the debrief

Population

The majority of included studies recruited non-clinical populations. Specifically, five studies enrolled young adults with mean ages ranging from 21.82 ± 1.84 to 34.2 ± 10.6 years (Espinola et al., 2024; Hamzeheinejad et al., 2019; Kourtesis et al., 2019; Meschberger-Annweiler et al., 2023; Pardini et al., 2022). One study included participants aged between 21 and 37 years (D. Liao et al., 2018). One study recruited two cohorts of healthy older adults (mean ages 71.3 ± 4.3 and 67.5 ± 5.8 years) alongside healthcare professionals from diverse medical disciplines, including physiotherapy, orthopedics, psychology, physical education, and sports science (Goumopoulos et al., 2023). Two additional studies focused on healthy older adults with mean ages of 53.5 ± 20.30 and 71.48 ± 4.09 years, respectively (Borgnis et al., 2021; Chen et al., 2020).

Eight studies enrolled participants with various clinical conditions. (Lorentz et al., 2023) examined patients with brain injury-related attention deficits (mean age 51.66 ± 17.8 years). One study investigated patients diagnosed with Mild Cognitive Impairment (MCI) (mean age 75.73 ± 6.36 years) (Mondellini et al., 2022), while (Liu et al., 2023) recruited individuals with post-stroke cognitive impairment (mean age 74.93 ± 6.81 years). (Colombo et al., 2024) focused on patients with mild to moderate Chronic Obstructive Pulmonary Disease (COPD) (mean age 71.29 ± 6.93 years). Another study compared ADHD patients (mean age 31.9 ± 8.8 years) with neurotypical controls (mean age 34.4 ± 8.9 years) (Tölgyesi et al., 2023). One investigation contrasted stroke patients (mean age 44 ± 19 years) with neurologically healthy controls (mean age 59.57 ± 9.78 years) (Huygelier et al., 2022). Patients with Parkinson's Disease were examined in one study (mean age 68.2 ± 9 years) alongside healthy controls (mean age 66.4 ± 10.5 years) (Borgnis et al., 2022). Finally, one study evaluated older adults and patients with lower limb disorders, with mean ages of 85.16 ± 5.93 and 61.10 ± 12.62 years, respectively (Rojo et al., 2023).

VR application

Most of the reviewed studies fall into the category of preliminary research investigations, primarily concentrating on acceptability and feasibility evaluations. These studies were designed to evaluate the feasibility of interventions with an emphasis on UX when interacting with technological applications. The applications are mainly aimed at enhancing cognitive and motor abilities, and most of those applications rely on computer-generated virtual environments to deliver therapeutic experiences.

Several studies employed head-mounted displays (HMDs) to deliver immersive experiences. Lorentz and colleagues (Lorentz et al., 2023) implemented a VR travel application that simulated a global journey encompassing three destinations: New York, the Great Barrier Reef in Australia, and the Brazilian Amazon. Participants engaged in attention training tasks requiring them to press response keys when skyscrapers illuminated, identify specific fish appearing on screen, and count and recall the number and location of butterflies. Another investigation (Tölgyesi et al., 2023) developed Zenctuary VR, which featured an interactive garden environment surrounded by forest and flowing water, populated with natural elements that provided auditory, visual, and tactile feedback in response to user interactions. One study created a virtual supermarket scenario for shopping simulation, requiring participants to navigate grocery purchasing tasks (Mondellini et al., 2022), while Espinola and collaborators (Espinola et al., 2024) developed a VR hospital environment designed to address stress among healthcare workers during the COVID-19 pandemic.

Other applications included the GAME2AWE platform, which aimed to improve cognitive and motor skills (Goumopoulos et al., 2023). It featured "Life on a Farm" and "Fun Park Tour" themes that were presented through VR and augmented reality, respectively. A virtual park was developed by Colombo and colleagues, who integrated cycle-ergometer equipment to simulate bicycle riding, thereby enhancing physical performance in projected natural landscapes (Colombo et al., 2024). Other implementations included attention bias modification tasks utilizing HMDs with body trackers (Meschberger-Annweiler et al., 2023), where the virtual environment consisted of a furnished room containing only a large mirror and two floor-positioned boxes, allowing participants to view their full avatar representation in the mirror.

Motion sensors combined with HMDs were employed by one study (Liu et al., 2023) to deliver comprehensive training programs featuring life skills and exergames. The training encompassed three categories of VR games totaling sixteen items: life skills training (cooking, window cleaning, road crossing, flower watering), exergames (squash, antiaircraft gun shooting, glider flying, baseball), and entertainment games (eggshell breaking, insect swatting, firework lighting, whack-a-mole, balloon pumping, Kongming lantern flying, Fruit Ninja, bubble popping). Another investigation (Huygelier et al., 2022) provided visual discrimination tasks within natural settings through a serious game featuring a vegetable garden, lake, and forest presented under three lighting conditions: day, evening, and night. Participants wearing HMDs performed specific visual discrimination tasks within these environments.

In one study, participants used screens with motion-activated interaction sensors to play exergames that included coordination exercises like fruit catching and wall dodging (Chen et al., 2020). In these VR games, players stepped on rodents that appeared on the ground and walked through gaps in virtual walls, using lateral body movements to capture virtual fruits. Additional studies provided more immersive experiences that involved the use of HMDs combined with walking platforms (Hamzeheinejad et al., 2019) or static pedaling ergometers (Rojo et al., 2023). One VR application offered three pedaling scenarios where participants steered virtual vehicles such as light aircraft and fishing boats through pedaling to navigate the sky, canyon, and sea. Another application offered four different natural sceneries: grassland, forest, streamland, and beach, with optional virtual trainers that motivate users during walking.

Three studies focused on the use of 360-degree media technology. One of these (Pardini et al., 2022) offered personalized relaxation experiences by presenting natural settings along with auditory stimuli. The two other studies (Borgnis et al., 2022, 2021) developed EXIT 360° using real-world domestic photographs to create complex cognitive tasks. This application contained several virtual settings where participants carried out seven subtasks of varying complexity. These subtasks involved opening doors by selecting options, exploring rooms, identifying people as instructed, recognizing pieces of furniture, completing rebuses, memorizing sequences of numbers in reverse order, and recalling other tasks. Virtual environments were delivered through HMDs.

Liao et al. (D. Liao et al., 2018) explored personalized virtual environment design using 360-degree content for depression detection, incorporating both natural and clinical scenarios. The authors developed a 360-degree media solution for HMDs featuring interactive scenes composed of natural environments (such as green groves) and clinical scenarios resembling authentic diagnostic environments with integrated images, language, and character gestures.

Kourtesis end colleagues (Kourtesis et al., 2019) focused on validating a questionnaire for assessing VR experience quality, using a triad of commercial games: "Job Simulator," "The Lab," and "Rick and Morty: Virtual Rickality." Each game presented different scenario types, including occupational, skill-testing, and domestic tasks. Players in "Job Simulator" assumed different job roles in the game, including receptionist, cook, and automotive mechanic, using an HMD. "The Lab" presented participants with mini-games such as slingshot and longbow challenges designed to assess different skills. "Rick and Morty: Virtual Rickality" required users to complete simulated domestic tasks within a fictional setting.

UX assessment

The majority of studies conducted participant interviews regarding their experiences immediately following VR sessions, primarily requiring completion of one or more questionnaires (Borgnis et al., 2022, 2021; Chen et al., 2020; Colombo et al., 2024; Goumopoulos et al., 2023; Hamzeheinejad et al., 2019; Kourtesis et al., 2019; D. Liao et al., 2018; Liu et al., 2023; Lorentz et al., 2023; Meschberger-Annweiler et al., 2023; Pardini et al., 2022; Rojo et al., 2023; Tölgyesi et al., 2023). In contrast, two investigations (Huygelier et al., 2022; Mondellini et al., 2022) measured cybersickness symptoms both pre- and post-intervention, in addition to administering various user experience scales following the VR experience.

Four studies (Chen et al., 2020; Colombo et al., 2024; Kourtesis et al., 2019; Lorentz et al., 2023) implemented the User Experience Questionnaire (UEQ) (Laugwitz et al., 2008) in its original (Hamzeheinejad et al., 2019; Lorentz et al., 2023), abbreviated (Chen et al., 2020), or author-modified (Colombo et al., 2024) versions. One study (Liu et al., 2023) supplemented the UEQ with the Short Flow State Scale 2 (SFSS-2) developed by Jackson et al. (Jackson and Marsh, 1996).

Four investigations employed purpose-designed questionnaires. Specifically, Liao et al. (D. Liao et al., 2018) developed a satisfaction survey utilizing a 5-point scale ranging from 1 (dissatisfaction) to 5 (satisfaction). Another study (Pardini et al., 2022) incorporated 8 items from the Virtual Reality Symptom Questionnaire (VRSQ) (Ames et al., 2005) and 37 items from both the Presence Questionnaire and Immersive Tendencies Questionnaire (Witmer and Singer, 1998). Liu et al. (Liu et al., 2023) constructed a self-report questionnaire comprising 14 items organized into three categories: smart device usage frequency, satisfaction levels, and occurrence and severity of adverse reactions during intervention. Additionally, one investigation (Huygelier et al., 2022) developed a custom user experience scale based on the

International Test Commission - Sense of Presence Inventory (ITC-SOPI) (Lessiter et al., 2001), Narrative Engagement Scale (Busselle and Bilandzic, 2009), and Intrinsic Motivation Inventory (McAuley et al., 1989), while also administering the Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993).

One study (Tölgyesi et al., 2023) assessed user experience through the User eXperience in Immersive Virtual Environment questionnaire (UEIVE) (Tcha-tokey et al., 2016). Another investigation (Mondellini et al., 2022) utilized the SSQ (Kennedy et al., 1993), ITC-SOPI (Lessiter et al., 2001), and the Technology Acceptance Model 3 questionnaire (Venkatesh and Bala, 2008). Espinola et al. (Espinola et al., 2024) implemented the Igroup Presence Questionnaire (IPQ) (Schubert, 2003) and additionally evaluated dropout rates and qualitative responses obtained during debriefing sessions to assess user experience. One study (Goumopoulos et al., 2023) evaluated both expert and user perspectives during the VR application design phase through brief interviews and discussions, subsequently requiring participants to complete the System Usability Scale (SUS) questionnaire (Brooke, 1996), VRSQ (Ames et al., 2005), and Acceptance and Use of Technology (UTAUT) model questionnaire (Venkatesh and Bala, 2008). The SUS was also employed by another investigation (Meschberger-Annweiler et al., 2023). Two studies (Borgnis et al., 2022, 2021) administered the UEQ (Laugwitz et al., 2008), ICT-SOPI (Lessiter et al., 2001), three items from the Flow Short Scale (Engeser et al., 2005), and four items from the Intrinsic Motivation Inventory (enjoyment subscale) (Deci et al., 1994). Finally, one study (Rojo et al., 2023) implemented the Intrinsic Motivation Inventory (Deci et al., 1994), Credibility and Expectancy Questionnaire (CEQ) (Devilly and Borkovec, 2000), SSQ (Kennedy et al., 1993), Presence Questionnaire (PQ) (Witmer and Singer, 1998), SUS (Brooke, 1996) and the 18-item abbreviated version of the Game User Experience Satisfaction Scale (GUESS) (Keebler et al., 2020). The Virtual Reality Neuroscience Questionnaire (VRNQ) was specifically designed and validated by one investigation (Kourtesis et al., 2019). Table 5 provides comprehensive details regarding the questionnaires and their associated variables.

Table 5. The table displays the instruments each study employed, and the specific variables chosen by the authors to assess UX.

Paper	**UX evaluation**	**Variables involved in UX evaluation**	**Evaluation time point**

(D. Liao et al., 2018)	ad hoc survey	Easy to learn, Interest, Complexity, Attractiveness, Naturalness, Definition, Visibility, Vertigo	Post VR intervention
(Hamzeheinejad et al., 2019)	User Experience Questionnaire (UEQ)	Attractiveness, Efficiency, Perspicuity, Dependability, Stimulation, Novelty	Post VR intervention
(Kourtesis et al., 2019)	Virtual Reality Neuroscience Questionnaire (VRNQ)	User experience (intensity of the immersion, the level of enjoyment, quality of the graphics, sound, and VR technology), Game mechanics, In-game assistance, VR symptoms and effects	Post VR intervention
(Chen et al., 2020)	User Experience Questionnaire short version (UEQ-S)	Pragmatic and Hedonic quality	Post VR intervention
(Borgnis et al., 2022, 2021)	User Experience Questionnaire (UEQ)	Attractiveness, Efficiency, Perspicuity, Dependability, Stimulation, Novelty	Post VR intervention
	International Test Commission -Sense of Presence Inventory (ITC-SOPI)	Spatial presence, Engagement, Naturalness, Side-effects	Post VR intervention
	Flow Short Scale (three items)	Abilities in coping with the task, Challenges, Challenge-skill balance	Post VR intervention
	Intrinsic Motivation Inventory (four items)	Enjoyment	Post VR intervention
(Huygelier et al., 2022)	Simulator sickness questionnaire (SSQ)	Nausea, Oculomotor disorders, Disorientation	Pre and post VR intervention
	ad hoc User Experience scale	Usability of the touch controllers and user interface, Amount of presence experienced, Experience of the narrative, and Motivation	Post VR intervention
(Mondellini et al., 2022)	Simulator Sickness Questionnaire (SSQ)	Nausea, Oculomotor disorders, and Disorientation	Pre and post VR intervention
	International Test Commission -Sense of	Spatial presence, Engagement, Naturalness, Side-effects	Post VR intervention

	Presence Inventory (ITC-SOPI)		
	Technology Acceptance Model 3 questionnaire	Perceived ease of use, Computer anxiety, Perceived enjoyment, Behavioral intention	Post VR intervention
(Pardini et al., 2022)	ad hoc questionnaire	General physical side effects (e.g., fatigue, headache, nausea, concentration difficulties), Visual effects (e.g., blurred vision, and tired eyes), Realism of the environments, engagement, Immersiveness, tools' Usability and quality of the interface, Emotional states, Satisfaction	Post VR intervention
(Goumopoulos et al., 2023)	Short interviews and discussions with seniors and experts	Whether such games could have a positive impact, and to identify any features that must be implemented in the future	Post VR intervention
	System Usability Scale (SUS) questionnaire	Effectiveness, Efficiency, Satisfaction	Post VR intervention
	Virtual Reality Sickness Questionnaire (VRSQ)	General discomfort, Fatigue, Eyestrain, Difficulty focusing, Headache, Fullness of head, Blurred vision, Dizzy, Vertigo;	Post VR intervention
	Acceptance and Use of Technology (UTAUT) model questionnaire	Performance expectancy, Effort expectancy, Social Influence, Facilitating conditions	Post VR intervention
(Liu et al., 2023)	self-reported questionnaire	3 parts: (i) how often individuals used smart devices before the intervention, (ii) satisfaction with equipment and training content, (iii) occurrence and degree of adverse reactions during the intervention	Post VR intervention
(Lorentz et al., 2023)	User Experience Questionnaire (UEQ)	Attractiveness, Efficiency, Perspicuity, Dependability, Stimulation, Novelty	Post VR intervention
(Meschberger-Annweiler et al., 2023)	System Usability Scale (SUS) questionnaire	Effectiveness, Efficiency, Satisfaction	Post VR intervention

(Rojo et al., 2023)	Intrinsic motivation inventory (three subscales)	Value/usefulness, Interest/enjoyment, Perceived choice	Post VR intervention
	Credibility and Expectancy Questionnaire (CEQ)	Credibility and Expectancy	Post VR intervention
	Simulator sickness questionnaire (SSQ)	Nausea, Oculomotor disorders, Disorientation	Post VR intervention
	Presence questionnaire (PQ)	Realism, Control, Quality of interface, Possibility to examine, Possibility to act, and Self-evaluation.	Post VR intervention
	18-item short scale of Game User Experience Satisfaction Scale (GUESS)	Usability/playability, Narratives, Play engrossment, Enjoyment, Creative freedom, Audio aesthetics, Personal gratification, Social connectivity, Visual aesthetics	Post VR intervention
	System Usability Scale (SUS)	Effectiveness, Efficiency, Satisfaction	Post VR intervention
(Tölgyesi et al., 2023)	User eXperience in Immersive Virtual Environment Questionnaire (UEIVE)	Presence, Engagement, Immersion, Flow, Skill, Emotion, Usability, Technology adoption, Judgment, Experience consequence	Post VR intervention
(Colombo et al., 2024)	Modified version of the User Experience Questionnaire (UEQ)	Attractiveness, Perspicuity, Stimulation, Novelty	Post VR intervention
	Short Flow State Scale 2 (SFSS-2)	Challenge–skill balance, Action–awareness, Clear goals, Unambiguous feedback, Concentration on the task at hand, Sense of control, Transformation of time, Loss of self-consciousness, Autotelic experience	Post VR intervention
(Espinola et al., 2024)	Igroup Presence Questionnaire (IPQ); dropout rate;	General presence, Spatial presence, Involvement, Experienced realism	Post VR intervention

	qualitative responses provided during the debrief		

Discussion

The present review examines questionnaires currently utilized to assess UX in VR-based applications in healthcare settings. Our investigation yields some significant findings: (i) while most studies tested the experience of clinical populations, particularly those with neurological conditions, focused on healthy adults; (ii) VR applications used for UX evaluation primarily utilized immersive, interactive, and computer-generated environments to train cognitive and motor abilities; (iii) UX evaluation usually occurred right after VR sessions; and (iv) researchers used multiple questionnaires to capture a comprehensive experiential dimension.

Methodological considerations and limitations of the literature

Several methodological limitations emerged from our analysis that warrant consideration. User feedback collection remains crucial, requiring involvement of appropriate end-user populations to ensure experimental samples reflect target demographics. Through direct engagement, clinicians and developers can obtain comprehensive insights into authentic needs, preferences, and implementation challenges within clinical settings. Nonetheless, merely one study in our review included expert assessment during the development phase. This factor, despite being frequently overlooked, represents equally significant consideration. A variety of clinical and technical experiences may yield insightful information if interdisciplinary viewpoints are included, including those of professionals, therapists, patients, and health researchers. Combining these different points of view might offer more thorough assessment frameworks, recognizing improvements potentially overlooked when consulting only end-users or designers. This comprehensive assessment approach more accurately takes into account safety concerns, long-term clinical efficacy, and the needs of healthcare system integration.

Another drawback is the preponderance of immediate post-intervention feedback collection. Although real-time user feedback yields insightful information, baseline conditions might hide some effects. This is especially true for negative physiological reactions like headache, nausea, and vertigo, which could be caused by participant circumstances or side effects of the VR program and could distort the results. Some research used tools such as the Simulator Sickness Questionnaire before and after the intervention to record changes related to experimental circumstances to solve this methodological issue (Huygelier et al., 2022; Mondellini et al., 2022).

Additionally, a crucial but frequently disregarded component of therapies is their long-term impact. Investigating the impact of persistence over time or across multiple sessions is still crucial, even though the examined studies concentrated on single-session outcomes. This is especially crucial for patient populations with chronic conditions who may show decreased exercise tolerance and motivation, which can have a substantial impact on intervention engagement (Colombo et al., 2024).

Furthermore, many instruments utilized in current research were designed for outdated technologies and may not fully encompass modern technological complexities (e.g., (Brooke, 1996; Busselle and Bilandzic, 2009; Jackson and Marsh, 1996). A notable paucity of cross-cultural validation exists for numerous questionnaires, particularly in non-English speaking contexts, where tools are often merely translated without appropriate validation processes. This absence of rigorous validity evaluation may compromise measurement accuracy across different contexts and result reliability. Many existing instruments were originally developed for traditional computer interfaces or conventional display settings (Brooke, 1996; Lessiter et al., 2001), raising questions about their appropriateness for evaluating novel VR interaction methods given their unique attributes such as immersion, presence, and spatial interaction. To fill this methodological gap, researchers generally employ numerous questionnaires, as this review shows. Their objective is to capture the wider range of characteristics that are thought to affect the overall VR experience. While this method facilitates a more holistic assessment, it also presents issues with respondent load and possible construct overlaps, highlighting the need for a more systematic approach to user experience evaluation.

The eight factors framework: a novel contribution

The central innovation of this review lies in the identification and systematization of eight key dimensions that consistently emerge across UX evaluations in healthcare VR-based applications. Through comprehensive analysis of questionnaire variables and expert consultation, we noted a substantial number of variables that can be grouped into several key dimensions that emerged as the most frequently assessed components, identifying core factors that provide a structured framework for UX assessment. The eight dimensions encompass: (i) *usability and functionality*, addressing technical aspects; variables including in this dimensions comprises efficiency (the product's capacity to enable rapid and optimal utilization), perspicuity (interface clarity and comprehensibility), dependability (user confidence and control when using the product, referring to system reliability and predictability), effectiveness (system's ability to enable users to accomplish specific tasks), and ease of learning and use; (ii) *aesthetics of design*, incorporating attractiveness (general aesthetics and appeal) and design elements such as innovation, creativity, and cutting-edge features (novelty); (iii) *engagement*, involving (whether the product captures user attention and engages participants), the extent to which users are mentally and emotionally invested in the experience (involvement), immersion, and flow states; (iv) *emotional state*, encompassing positive and negative feelings, satisfaction, and enjoyment elicited by VR experiences; (v) *presence*, identified as a core feature across numerous studies; (vi) *realism of environments*, referring to the degree of resemblance to reality; (vii) *side effects*, including both general physical symptoms (such as nausea, vertigo, fatigue, headache, disorientation) and visual complications (including oculomotor disorders, eyestrain, blurred vision); and (viii) *motivation and intention of use*, encompassing behavioral intention, perceived usefulness, and technology adoption variables. Figure 10 illustrates the categorization of variables into the aforementioned key factors.

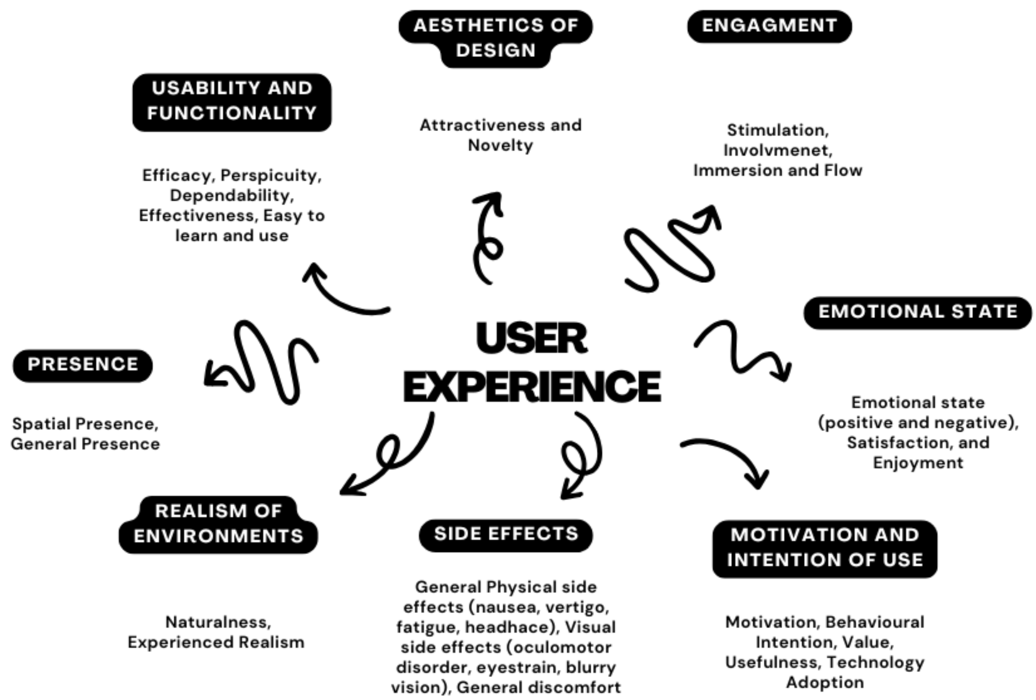

Fig. 10. Schematic representation of eight key factors comprising UX.

This eight-factor framework represents a significant advancement in the field by providing a comprehensive structure for UX evaluation in healthcare VR-based applications. Unlike previous approaches that employed multiple disparate questionnaires to capture various experiential aspects, this framework offers a systematic approach to understanding how users interact with and perceive VR environments. The framework's multifaceted nature enables researchers and developers to gain holistic insights while maintaining methodological rigor, ultimately contributing to more effective and user-centered VR healthcare applications.

Conclusion

In conclusion, while UX evaluation is increasingly recognized as a gold standard in the development and implementation of novel VR applications by healthcare specialists, this review highlights significant gaps that require attention. The efficacy and methodological rigor of UX evaluation techniques could be significantly improved by our consideration. In particular, the quality and applicability of VR healthcare applications may be enhanced using culturally validated instruments, the adoption of multidisciplinary evaluation approaches, and the implementation of longitudinal assessment protocols. Furthermore, by offering a methodical approach to UX assessment, the eight-factor framework that our analysis revealed makes a significant contribution to the field.

Nevertheless, this review presents some inherent limitations. Our analysis encompassed a diverse array of studies employing varied methodologies, which, while providing a broad perspective, may have introduced heterogeneity into our findings. Moreover, we specifically focused on VR applications in healthcare and questionnaire-based assessment methods; however, future investigation would help to expand the range to include other kinds of alternative UX evaluation techniques and technological equipment, including physiological measurements, behavioral analytics, and mixed-methods methodologies.

Further research is needed to validate the proposed eight-factor framework across diverse healthcare populations and VR applications, ultimately advancing the field toward more standardized, evidence-based UX evaluation practices in healthcare VR implementations.

User eXperience questionnaire for Virtual Reality applications in Healthcare: development of a new tool

Introduction

The concept of UX has evolved over recent decades into a multifarious field including several areas of human-centered interaction. UX research originally focused on technical features of device interaction and demanded technical know-how for development and evaluation. However, contemporary applications have transcended traditional boundaries, expanding into fields such as medicine and psychology, thereby reflecting a more comprehensive interdisciplinary approach (Bruni et al., 2025a). This expansion demonstrates the successful convergence of engineering and computer science ideas with clinical and behavioral research techniques, therefore changing the field's foundational paradigm. Rather than focusing exclusively on technical specifications and system performance, modern UX evaluation takes a broad view that includes many facets of user engagement and experience. However, an ongoing challenge in literature is the fundamental question of how to operationalize UX itself. Although researchers usually agree that UX comprises both practical (instrumental) and hedonic (non-instrumental) elements, its exact definition is elusive due to its interdisciplinary nature and the interplay of multiple factors (Maia and Furtado, 2016). UX complexity is also increased by contextual factors, temporal dynamics, and user attributes (i.e., personal and demographic traits) (Augstein et al., 2017; Nimmo et al., 2024; Parsons et al., 2004; Strada et al., 2013; Venkatesh and Morris, 2000). Furthermore, the relative importance of these dimensions varies depending on the system's purpose, user needs, and usage context. As a result, it is difficult to establish a single definition of UX. It appears as a heterogeneous concept that encompasses a variety of theoretical models and frameworks, varying from purely functional considerations to emotional and experiential ones.

In our previous work, we attempted to define the UX concept specifically for healthcare VR apps through a comprehensive literature review examining UX evaluation questionnaires. Our analysis identified key aspects that consistently emerged as the most frequently evaluated components during VR experiences in healthcare settings, proposing an eight-factor framework: usability and functionality, aesthetics of design, engagement, emotional state, motivation and Intention of use, side effects, realism of environments, and presence (shown in Figure 10 in the previous chapter).

Based on this model, we aim to develop a short and reliable questionnaire for evaluating the UX of VR applications in healthcare. Moreover, we would like to respond to the limitations of existing questionnaires and self-developed instruments, designing a specialized assessment tool specifically for healthcare VR applications, with particular emphasis on aging populations and rigorous questionnaire design principles. Our methodology employed a systematic approach that balances scientific rigor with practical utility. This paper details the design process of this instrument, responding directly to the field's need for a reliable UX assessment tool.

Methods

Questionnaire Development and Response Scale Selection

Based on our previous literature review results, we grouped all items of the most used UX questionnaire according to each of the eight factors identified in our review. Specifically, all questions related to a particular variable across different questionnaires were consolidated into single categories (e.g., all questions about sense of presence from all analyzed questionnaires were grouped). Subsequently, we eliminated repetitive questions and, through consultation with domain experts, conducted a careful selection process to determine which items to retain and which to exclude. Based on this review, we identified the most frequently used items for assessing the proposed factors, resulting in an initial pool of 271 items. Figure 11 clarifies in a diagram all phases of the process.

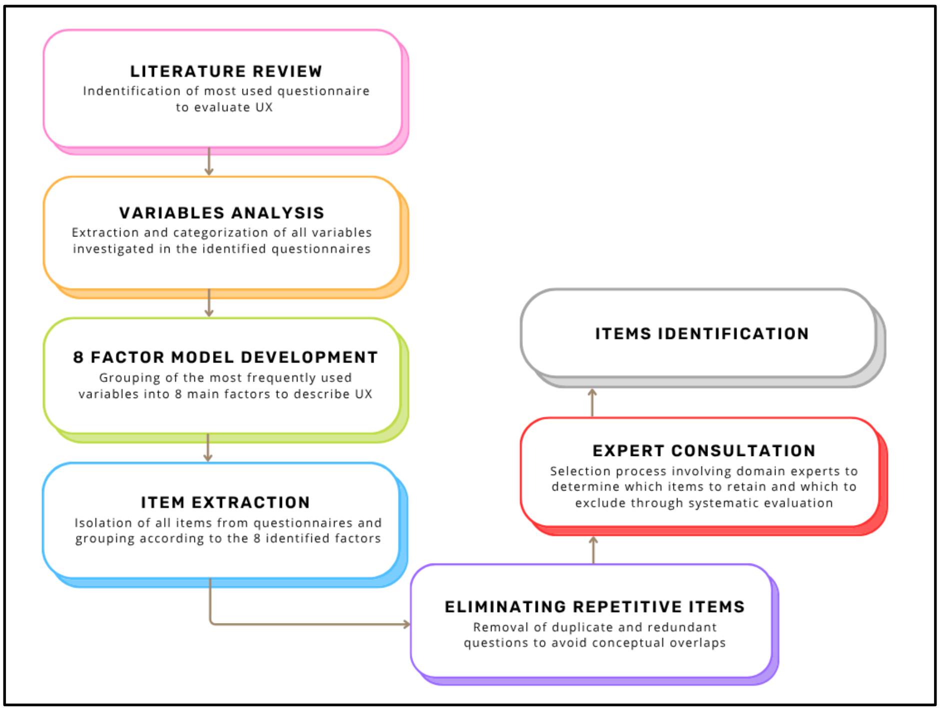

Fig. 11. Flowchart showing all phases of the process to identify preliminary items.

In a subsequent step, we revised each statement, considering the most common item formats within UX questionnaires, and established guidelines (Chiorri, 2011; Jespersen et al., 2018), such as simplicity, clarity, unambiguity, and the avoidance of vague quantifiers. We formulated items in a clear, specific manner, eliminating ambiguities and avoiding overly technical, specialist, or colloquial language. At this stage, we did not address content redundancies, refined wording, or item numerousness, as these aspects were planned for subsequent user-centered design phases. Regarding response options, we noted that UX literature typically employs Likert scales to measure agreement levels with the proposed statements. We selected a 5-point scale ranging from "not at all" to "extremely agree" as the optimal format, based on the research recommendations indicating that three-point scales restrict response diversity, whereas longer scales increase cognitive load. Evidence suggests that the optimal balance between cognitive demand and expressive capability lies within 5-7 options, preferably with an odd number of alternatives (Chiorri, 2011). We also developed specific instructions for the self-administered questionnaire to ensure standardized implementation across various settings. These instructions provide clear guidance for users, detailing the purpose of the questionnaire and completion procedures. Figure 12 presents the complete instruction set. These instructions were formulated to maximize response accuracy while maintaining accessibility for older adult populations, employing straightforward language and explicit directives that emphasize the subjective nature of the assessment.

User eXperience questionnaire for Virtual Reality applications in Healthcare (UXVR-H)

This questionnaire contains 24 statements about the system and the proposed experience. Indicate how much you agree or disagree with each statement, by circling the number that best represents your opinion on the scale from 1 to 5, where:

1 = Not at all **2** = Slightly **3** = Moderately **4** = Considerably **5** = Extremely

There are no right or wrong answers; your opinion is what matters.

Fig. 12. Instructions for the UXVR-H questionnaire.

Validation Process and Content Refinement

To ensure the quality and validity of our questionnaire, we implemented a comprehensive two-phase validation process. The first phase focused on user-centered assessment to evaluate item clarity and comprehensibility among the target population, while the second phase involved expert evaluation to establish content validity and theoretical alignment. This dual approach allowed us to balance practical usability with theoretical rigor, ensuring that the final instrument would be both accessible to end-users and methodologically sound.

1. User-Centered Assessment

Following item development, we implemented a reformulation technique wherein participants were asked to read and rephrase items in their own words. This iterative consultation process ensured appropriateness of questions by presenting preliminary items to potential end-users from the target population and requesting their rephrased versions. Given the specificity of the questionnaire application context, a cohort of Italian older adults (aged ≥ 60 years) was involved in this first reformulation phase to establish a comprehensive understanding of each item, evaluating both clarity and comprehensibility. Participants were encouraged to express opinions and concerns, including syntactic clarity, comprehension challenges, and item length. The initial pool of 271 items was divided into 14 subgroups (11 containing 20 items each and 3 containing 17 items each). Each subgroup was rephrased by 10 people. Items that even one participant could not reformulate in alignment with the researchers' original conceptualization were eliminated or redrafted. This participatory approach verified item clarity while generating alternative, more accessible formulations. As a result of this refinement process, we derived an initial version of the UXVR-H instrument comprising 66 items.

2. Expert Evaluation

The second phase included a systematic collection of expert evaluations to determine content validity. Our goal was to integrate experts' judgments on item relevance and representativeness with respect to the construct definition. Eight independent judges with expertise in UX and methodology performed comprehensive evaluations of the instrument, including qualitative analysis, construct definition, and dimensional structure, and quantitative evaluation of the relevance of each item. Experts were invited to give qualitative feedback on the overall instrument and items (e.g., observing grammar, using appropriate and correct words, applying correct and proper order of words in items, and appropriate scoring). Moreover, they were asked to evaluate the relevance of each item for defining UX using a 3-point Likert scale ranging from 1 (less relevant) to 3 (very relevant). Additionally, judges - blind to our hypotheses – classified each item according to the theoretical variable framework. Table 1 illustrates an example of the evaluation grid, translated from the original Italian version. It shows the item relevance rating scale (1-3) and the component classification dropdown menu used during the content validation process.

Content Validity Analysis

The expert evaluation process involved both qualitative and quantitative assessments. For quantitative content validity analysis, we employed two complementary indices: the Content Validity Ratio (CVR) and Content Validity Index (CVI). The CVR indicates the degree to which an item is necessary for operationalizing a construct within a set of items. The CVR varies between -1 and +1, with higher scores indicating greater agreement among panel members on the necessity of an item in the instrument. The critical value for CVR acceptance is determined by the Lawshe Table based on the number of panelists (Lawshe, 1975). For the present study, with 8 judges, the minimum value was 0.75, indicating significance at an acceptable level. The CVI provides a measure of content validity by calculating the proportion of experts who rated each item as relevant (rating 2 or 3). We calculated item-level CVI (I-CVI) and scale-level CVI (S-CVI). For I-CVI, we used the threshold of 0.78 as recommended for a panel of 6-10 experts (Polit and Beck, 2006). Items with I-CVI ≥ 0.78 were considered to have excellent content validity; moreover, we also considered items with I-CVI between 0.70-0.79 as acceptable but requiring potential modification (Shi et al., 2012), following expert feedback. We consider a value of 0.90 or higher as an indicator of excellent content validity of the overall scale (Polit and Beck, 2006). This index provides a global assessment of how well the entire instrument represents the examined construct.

Items were first evaluated considering the CVR threshold. This method ensures that only items judged as very relevant are retained, maintaining the necessary rigor for construct operationalization. Then, we performed the I-CVI analysis for items that did not meet the CVR threshold. This supplementary screening identifies items that, while not reaching the stringent CVR requirement, demonstrate overall relevance among experts and may benefit from modification rather than elimination. For item-dimension correspondence, we retained items that at least half of the judges classified to the appropriate theoretical dimension, based on our initial hypothesis. Fleiss' κ was used to evaluate inter-rater agreement for dimensional categorization.

Results

User-centered Assessment outcomes

The development process yielded an initial pool of 271 items; the reformulation phase involved 140 Italian older adults (aged ≥ 60 years, M = 67.8, SD = 5.2) of various education levels. They evaluated the items divided into 14 subgroups. Each subgroup was assessed by 10 participants to ensure a comprehensive evaluation of item clarity and comprehensibility. This process resulted in 205 items being eliminated due to comprehension issues or inappropriate formulation. Through 66 retained items, 23 were alternatively formulated based on participants' feedback to avoid technical language and ambiguous issues. The distribution of the remained 66 items across factors was as follows: usability and functionality: 8 items (12.1%), aesthetics of design: 3 items (4.5%), engagement: 10 items (15.1%), emotional state: 8 items (12.1%), presence: 10 (15.1%), realism of environments: 7 items (10.6%), side effects: 10 items (15.1%), motivation and Intention of use: 11 items (16.7%).

Expert Evaluation outcomes

Ten independent judges were invited to participate in the evaluation process. Eight experts responded and participated in the content validation phase. The expert panel comprised specialists in user experience, virtual reality researchers, psychometricians, and healthcare clinicians, ensuring comprehensive domain expertise for the validation process.

In the primary screening phase, 9 of the 66 examined items met the critical threshold established by Lawshe's table for 8 expert judges. This shows strong expert agreement on the relevance and need of these items for measuring UX in VR healthcare use. Thus, items were automatically retained without adjustments. For the remaining items that did not meet the CVR threshold, we performed I-CVI analysis. It revealed that 21 more items reached I-CVI ≥ 0.78, proving great content validity despite not meeting the CVR criteria. 9 of these items were removed despite excellent CVI ratings due to content redundancy or theoretical overlap that did not compromise the definition of the respective dimension. Expert consent that their removal would enhance rather than diminish the instrument's clarity. Table 6 presents examples of CVR and CVI calculations for items. Moreover, three additional items with an acceptable I-CVI score (0.75) were retained. These items demonstrated acceptable relevance but required modifications according to expert suggestions and research group consensus. The qualitative feedback focused on linguistic clarity, conceptual precision, and dimensional alignment.

Table 6. The table exemplified the CVR and CVI calculation for some selected items.

Item	**Expert rating item essential (3s)**	**Expert rating items 2-3s**	**CVR**	**I-CVI**	**Decision**
The system is easy to use	7 out of 8	8 out of 8	0.75	1.00	Retained
I found the graphics appealing	4 out of 8	8 out of 8	0.50	1	Modified
The experience amused me	4 out of 8	5 out of 8	0	0.63	Eliminated

Regarding the inter-rater agreement for dimensional classification yielded κ = 0.60, indicating good agreement among experts. However, in certain instances, items were reassigned to alternative dimensions based on expert consensus when the majority of judges consistently categorized them differently from the original theoretical framework. For instance, item 4 was reassigned from the "motivation and intention to use" dimension to the "usability" dimension following expert consensus. The final distribution of items across dimensions is presented in Table 7.

The final validation outcomes demonstrate robust psychometric properties. The S-CVI achieved a value of 0.91. which exceeds the established threshold of 0.90, confirming excellent content validity of the instrument according to established psychometric criteria

Table 7. The distribution of the final 24 items across 8 factors framework.

Variables	Numbers of Items	Example of Item retained
Usability and Functionality	4	Il Sistema è facile da usare
Aesthetics of Design	2	L'aspetto del prodotto è attraente
Engagement	3	Mi sentivo coinvolto nell'esperienza
Emotional state	3	L'esperienza mi ha soddisfatto
Presence	3	Mi sentivo corcondato dal mondo virtuale
Realism of Environments	3	L'ambiente mostrato mi sembrava reale
Side effects	3	Hoprovato nausea durante l'esperienza
Motivation and Intention of Use	3	Vorrei utilizzare il dispositivo proposto

The outcome of our development process culminated in a 24-item instrument that comprises the UXVR H (User Experience in Virtual Reality for Healthcare) self-administered questionnaire. Items are measured using a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). This instrument evaluates eight key dimensions of UX in VR-based applications for healthcare, as identified through our comprehensive theoretical framework and validated through expert consensus. Figure 13 presents the complete instruments with all items.

	Per niente	Poco	Abbastanza	Molto	Moltissimo
Il sistema proposto era facile da usare	1	2	3	4	5
L' aspetto dell'ambiente con cui ho interagito è attraente	1	2	3	4	5
L'ambiente con cui ho interagito era stimolante	1	2	3	4	5
Durante l'esperienza mi sono stancato più del solito	1	2	3	4	5
Avevo la sensazione di essere proprio dentro l'ambiente	1	2	3	4	5
L'esperienza mi ha annoiato	1	2	3	4	5
È stato facile imparare come usare il sistema proposto	1	2	3	4	5
Ho provato mal di testa durante l'esperienza	1	2	3	4	5
Mi sarebbe piaciuto che l'esperienza continuasse	1	2	3	4	5
Ho trovato il sistema proposto complicato da usare	1	2	3	4	5
L'ambiente mostrato mi sembrava reale	1	2	3	4	5
Durante l'esperienza ho sentito affaticamento della vista	1	2	3	4	5
Sentivo di poter interagire con l'ambiente mostrato	1	2	3	4	5
L'esperienza è stata frustrante	1	2	3	4	5
Durante l'esperienza ho perso la cognizione del tempo	1	2	3	4	5
Durante l'esperienza ho sudato più del solito	1	2	3	4	5
Mi sono sentito sicuro nell'interagire con l'ambiente mostrato	1	2	3	4	5
Ero completamente focalizzato sull'esperienza che stavo facendo	1	2	3	4	5
Durante l'esperienza ho avvertito un senso di vertigine	1	2	3	4	5
Avevo la sensazione di poter maneggiare gli oggetti nell'ambiente mostrato	1	2	3	4	5
Durante l'esperienza per me era irrilevante sapere cosa accadeva fuori	1	2	3	4	5
Ho provato nausea durante l'esperienza	1	2	3	4	5
L'esperienza mi ha soddisfatto	1	2	3	4	5
Mi sentivo competente nell'utilizzare il sistema proposto	1	2	3	4	5

	Not at all	Slightly	Moderately	Considerably	Extremely
The proposed system is easy to use	1	2	3	4	5
The visual aesthetics of the environment I interacted with was attractive	1	2	3	4	5
The environment I interacted with was stimulating	1	2	3	4	5
I felt tired during the experience	1	2	3	4	5
I felt as if I was physically present inside the environment	1	2	3	4	5
I found the experience boring	1	2	3	4	5
It was easy to learn how to use the proposed system	1	2	3	4	5
I had a headache during the experience	1	2	3	4	5
I wished the experience would continue	1	2	3	4	5
I found the proposed system difficult to use	1	2	3	4	5
The displayed environment seemed real to me	1	2	3	4	5
I felt eye strain during the experience	1	2	3	4	5
I felt I could interact with the displayed environment	1	2	3	4	5
The experience was frustrating	1	2	3	4	5
I lost track of time during the experience	1	2	3	4	5
I perspired more than usual during the experience	1	2	3	4	5
I felt confident interacting with the displayed environment	1	2	3	4	5
I was completely focused on the experience	1	2	3	4	5
I felt dizziness during the experience	1	2	3	4	5
I felt like I could handle objects in the displayed environment	1	2	3	4	5
During the experience, it was irrelevant for me to know what was happening in the real world	1	2	3	4	5
I felt nausea during the experience	1	2	3	4	5
I was satisfied with the experience	1	2	3	4	5
I feel competent in using the proposed system	1	2	3	4	5

Fig. 13. The UXVR-H questionnaire in both English and Italian versions.

Discussion

The UXVR-H represents a concise, self-administered questionnaire (5-10 minutes administration time), specifically designed to evaluate the quality of experience in VR-based applications for older adults. The instrument is intended to be applicable across both pathological and physiological contexts.

The value of this questionnaire encompasses multiple theoretical and practical dimensions. The UXVR-H questionnaire addresses three-dimensional experiential qualities afforded by VR environments and their distinctive phenomenological characteristics. This represents a significant advancement over traditional instruments that were originally developed for two-dimensional experiences and merely adapted to VR contexts without proper theoretical reconsideration. Moreover, the design methodology of the questionnaire accommodates the specific needs and characteristics of the target population. For instance, critical design elements, including vocabulary selection, item complexity, questionnaire length, and cognitive load considerations, were optimized for older adult participants who may have varying levels of technological familiarity and potential age-related cognitive changes.

Another important consideration concerns the multifaceted approach proposed with the UXVR-H questionnaire. Within a single, efficiently administered instrument, researchers can assess numerous variables while minimizing participant burden and reducing assessment fatigue. The essential dimensions of user experience include interface efficacy and usability metrics, design element impact and aesthetic considerations, subjective emotional responses and affective states, presence and immersion levels, perceived naturalness and realism of virtual environments, potential adverse physical reactions and cybersickness symptoms, and user propensity for VR technology adoption and acceptance.

Nevertheless, it is important to acknowledge inherent limitations in purely quantitative assessment approaches. Quantitative data from UXVR-H fail to indicate specific pathways for system improvements or intervention modifications. For instance, a negative rating on item 6 ('I found the experience boring') identifies the presence of a problematic area but fails to specify what elements should be enhanced, modified, or redesigned. Therefore, quantitative measures should always be completed with qualitative data, such as interviews, focus groups, or observational protocols, offering the contextual depth needed for useful insights and system improvement. This integrated assessment strategy maximizes the utility of the quantitative data while addressing the diagnostic limitations of standardized questionnaires.

Preliminary Validation of the UXVR-H: A User Experience Questionnaire for Healthcare Virtual Reality

Introduction

UX evaluation is crucial in VR application development, as users who have a positive experience with the proposed applications - finding them easy to use and engaging, for example - are more likely to adopt them consistently and use them as intended (Van der Heijden, 2004). This sustained engagement can potentially lead to better diagnostic accuracy and improved health outcomes. However, comprehensively capturing the multitude and complexity of UX presents significant challenges. In an attempt to respond to this challenge, literature identifies several evaluation methods, from observational techniques to the use of multiple sensors for physiological measurements (Inan Nur et al., 2021). Even with this range of possibilities, collecting users' feedback through questionnaires remains the most frequently employed approach due to its simplicity in both execution and data analysis (Inan Nur et al., 2021; Schrepp, 2015). Nevertheless, this approach has its own set of limitations that warrant careful consideration.

Despite the increasing number of questionnaires available in literature (e.g., Brooke 1996, Chin et al. 1987, Lewis 1995, Sauro and Dumas 2009, Lessiter et al 2001, Witmer and Singer 1998), most of them focus on an isolated aspect of VR experience, failing to capture comprehensive UX measures. For instance, many studies employed the System Usability Scale (SUS), confounding this usability measure with broader UX evaluation; otherwise, authors combined multiple questionnaires, trying to capture different aspects such as sickness, presence, and technical issues. This approach results in a proliferation of disparate instruments that inadequately capture UX's multifaceted nature, risking the reduction of this complex phenomenon to disconnected constructs merely grouped under the UX umbrella. Moreover, employing multiple questionnaires proves inconvenient regarding time and cost-effectiveness. A significant proportion of available tools is also outdated (e.g., Brooke, 1996; Jackson and Marsh, 1996; McAuley et al., 1989) and inadequately addresses modern technologies' complexities. These instruments were originally designed for two-dimensional interfaces, yet are frequently employed across diverse research contexts without appropriate adaptation. Given that contemporary VR tools offer fundamentally different experiential paradigms, the appropriateness of these measures for evaluating novel VR interaction methods is questionable, creating methodological gaps through misalignment between study requirements and measured constructs in the questionnaires (Holden and Karsh, 2010). Nevertheless, standard questionnaires offer the advantage of extensive use across multiple studies, enabling reliable measurement and facilitating cross-study comparisons.

Alternatively, authors may create custom questionnaires according to the general psychometric guidelines. While this approach provides flexibility to address unique features of specific VR applications, these instruments frequently lack adequate validation (Gallagher, 1974) and reliability (Lacker & Lessig, 1980). Since studies presenting VR-based tools typically prioritize technological validation over psychometric rigor, researchers often preclude reliable psychometric analysis. Consequently, the reliability and validity of findings from such studies remain questionable. Moreover, some of those custom questionnaires sometimes consist of yet untranslated versions, lacking cross-cultural validation, especially in non-English speaking contexts. Questionnaires were merely translated into a different language without cultural adaptation, potentially compromising validity in rigorous UX evaluations. For instance, in the Italian research landscape, there is a conspicuous absence of tools, to the best of our knowledge (Bruni et al., 2025). Furthermore, custom questionnaires lack standardized items across studies, hindering result comparison and standardization efforts. However, the most critical limitation of both well-validated and ad-hoc questionnaires predominantly reflects researchers' subjective UX conceptualizations rather than theoretically grounded and validated constructs, highlighting the field's current fragmentation and lack of consensus regarding UX definition and measurement.

These methodological gaps highlight the need for developing psychometrically sound questionnaires tailored specifically for virtual reality in healthcare. These tools should capture the unique aspects of immersive healthcare experiences while supporting the scientific standards necessary for robust evaluation.

Building on our previous research, this study aims to test the framework of the UXVR-H questionnaire through a preliminary analysis. The UXVR-H represents a concise and reliable instrument specifically designed for evaluating UX in VR-based healthcare applications.

We examine all items through complementary analytical approaches: confirmatory factor analysis on the multi-item factors, alongside descriptive and correlational analyses for all items, including the single-item factors. This preliminary study will inform the design of subsequent comprehensive validation research that will integrate all theoretical factors into a unified structural model.

Material and methods

UXVR-H questionnaire

The UXVR-H questionnaire comprises 24 items with a five-point Likert scale. Items are allocated across eight constructs: usability and functionality (4 items), design aesthetics (1 item), engagement (4 items), emotional states (3 items), presence (4 items), environmental realism (1 item), side effects (6 items), and motivation and intention to use (1 item).

Participants

A total of 126 participants were recruited (83 females and 43 males) with a mean of 64,34 years of age (SD = 8,68) and a mean of 14,20 years of education (SD = 4,42). The sample size was determined based on the recommended participant-to-item ratio to ensure adequate stability for the confirmatory factor analysis (CFA) procedure (Kyriazos, 2018). This approach aligns with established guidelines for preliminary validation studies utilizing moderate-sized samples (White, 2022). Participants were recruited using a mixed strategy to maximize representation. The primary recruitment involves a consecutive sampling within the IRCCS Istituto Auxologico Italiano - Department of Medicine, Neurology and Rehabilitation at the hospital. This was supplemented by a snowball sampling to reach even cases that do not access hospital services. All participants voluntarily took part in the experiment without any compensation. Inclusion criteria comprised individuals aged 50 years or older without documented neurological or psychological pathologies. All participants provided written informed consent in accordance with the Declaration of Helsinki and participated voluntarily without financial compensation.

Procedure

Participants completed an immersive VR experience delivered through a head-mounted display (HMD), during which they performed cognitive tasks (e.g., verbal and visuospatial memory exercises and visual search tasks). The proposed VR application was designed from photos and videos recorded using specific instrumentation (Insta360 X3) that are managed using Unity software and achieved with text and sound elements to create the final tool. The experience provided 8 exercises with varying cognitive demands among activities. All were interconnected exercises embedded within the scenario of a village excursion, to provide cohesive and realistic frameworks within which participants complete specific exercises that mimic daily requirements. The VE presented moderate visual complexity and muted color palettes designed to minimize potential overstimulation, and the application provided clear visual and auditory feedback for instructions, correct and incorrect responses. The VR application prioritized functional task completion and narrative elements over entertainment value or gamification features. The interaction modality relied primarily on head-orientation-based selection combined with handheld controller button presses for response confirmation. This relatively simple interaction paradigm was intentionally chosen to accommodate the motor and cognitive capabilities of the older adult target population.

The VR session lasted approximately 20 minutes. Following the VR experience, participants completed the UXVR-H questionnaire in a paper-and-pencil format. Each session was conducted individually in a controlled laboratory setting with the supervision of a psychologist. Responses from all participants were digitized and entered into Microsoft Excel spreadsheets, with double-entry verification procedures implemented to ensure data accuracy. Statistical analyses were performed using Jamovi 2.6 (The Jamovi Project, 2024). Statistical significance was set at $\alpha = 0.05$.

Statistical methods

Before all the analyses, we screened data for missing values, and participants with extensive missing responses were excluded. The remaining sporadic missing data were handled using Full Information Maximum Likelihood (FIML). Descriptive statistics (mean, standard deviation, skewness, kurtosis) were computed for all items to assess distributional properties and identify potential violations of normality assumptions. Given the complex structure of the UXVR-H questionnaire, with eight theoretical factors varying in item composition (ranging from 1 to 6 items per factor), we adopted different approaches to analyze multi- and single-item variables (Kline, 2023). CFA was conducted on five multi-item factors (Usability and Functionality, Engagement, Emotional State, Presence, and Side Effects) for a total of 21 items. Internal consistency for each factor was evaluated using Cronbach's alpha (α) - to measure the coherence with which items assess the same construct - and McDonald's omega (ω), to provide a more accurate reliability estimation (Revelle and Zinbarg, 2009). Then, item-rest correlations (i.e., correlations between each item and the corrected factor total) were examined to assess individual item contributions. We consider a value above 0.30 as adequate (Hair, 2009). Reliability was also considered, and factors failing to meet acceptable reliability thresholds ($\alpha > 0.70$) were excluded from the CFA model. Then, a multi-factor CFA model was tested using Maximum Likelihood estimation with FIML for missing data, with all items that passed the reliability tests. Model fit was evaluated using multiple indices: chi-square test, Comparative Fit Index (CFI; acceptable > 0.90), Root Mean Square Error of Approximation (RMSEA; acceptable < 0.08), and Standardized Root Mean Square Residual (SRMR; acceptable < 0.08). Convergent validity was checked through standardized factor loadings (target > 0.50), composite reliability (CR; target > 0.70), and average variance extracted (AVE; target > 0.50). The Fornell-Larcker criterion was used for discriminant validity. It requires that the square root of each factor's AVE exceeds its correlations with other factors, ensuring that factors measure distinct constructs (Fornell and Larcker, 1981).

For items treated as single indicators, we computed correlations with factor scores from the CFA to assess their nomological validity and how they fit into the larger UX framework. This analysis examined whether single items correlated with factors in theoretically expected patterns.

Results

Descriptive data

After data screening, three participants were excluded due to extensive data missing (> 60%). The final sample (N=123) presented a Missing Completely At Random pattern of missing data (Rubin, 1976), with an overall missing rate of 0.2%. No item showed systematic missing patterns, with all items having <5% missing responses.

Normality assessment via skewness e kurtosis (thresholds: |skewness| < 2, |kurtosis| < 7) revealed appropriate distribution of most items (West et al., 1995). Items 19 (dizziness) and 22 (nausea) showed severe normality violation, ceiling effects, and minimal variability. These items were excluded due to a lack of discriminant power. Table 8 shows descriptives.

Table 8. The descriptive data of 24 items.

							Skewness		Kurtosis	
	N	**Missing**	**Mean**	**SD**	**Min**	**Max**	**Skewness**	**SE**	**Kurtosis**	**SE**
Item 1	123	0	3.49	0.803	2	5	0.1367	0.218	-0.4283	0.433
Item 2	123	0	3.75	0.902	1	5	-0.5005	0.218	0.2113	0.433
Item 3	123	0	3.67	0.979	1	5	-0.4759	0.218	-0.0733	0.433
Item 4_Reverse	123	0	4.15	1.030	1	5	-1.1231	0.218	0.6729	0.433
Item 5	123	0	3.98	0.887	1	5	-0.6129	0.218	0.0488	0.433
Item 6_Reverse	122	1	4.83	0.476	2	5	-3.2960	0.219	12.6825	0.435
Item 7	122	1	3.45	0.946	1	5	0.1737	0.219	-0.3338	0.435
Item 8_Reverse	122	1	4.79	0.485	3	5	-2.2558	0.219	4.4422	0.435
Item 9	122	1	3.00	1.240	1	5	-0.0265	0.219	-0.9368	0.435
Item 10_Reverse	123	0	4.13	0.877	1	5	-1.1463	0.218	1.6724	0.433
Item 11	123	0	3.50	0.900	1	5	-0.2870	0.218	-0.1040	0.433

Item 12_Reverse	123	0	4.37	0.960	1	5	-1.5287	0.218	1.5216	0.433
Item 13	122	1	3.39	0.905	1	5	-0.1181	0.219	0.0748	0.435
Item 14_Reverse	123	0	4.60	0.744	1	5	-2.2398	0.218	5.6898	0.433
Item 15	122	1	2.47	1.200	1	5	0.3407	0.219	-0.9024	0.435
Item 16_Reverse	123	0	4.71	0.662	2	5	-2.5267	0.218	6.2783	0.433
Item 17	123	0	3.44	0.968	1	5	0.0649	0.218	-0.7123	0.433
Item 18	123	0	4.06	0.843	1	5	-0.6103	0.218	0.1934	0.433
Item 19_Reverse	123	0	4.77	0.722	1	5	-4.0064	0.218	17.2156	0.433
Item 20	123	0	2.76	1.195	1	5	0.1163	0.218	-0.8404	0.433
Item 21	123	0	3.41	1.179	1	5	-0.5039	0.218	-0.5127	0.433
Item 22_Reverse	122	1	4.80	0.655	1	5	-4.2518	0.219	20.1352	0.435
Item 23	123	0	4.05	0.922	1	5	-0.7356	0.218	0.0230	0.433
Item 24	123	0	3.00	1.008	1	5	0.2439	0.218	-0.3415	0.433

Correlation analysis identified Items 16 (excessive sweating) and 21 (irrelevance of external events) as problematic, with inadequate intra-factor correlations (Item 16: r = 0.008-0.208; Item 21: r = 0.138-0.189), as shown in Figure 14. Both items were excluded, yielding 17 multi-item indicators for subsequent analyses.

Correlation Matrix - Presence

		Item 5	Item 13	Item 20	Item 21
Item 5	Pearson's r	—			
	df	—			
	p-value	—			
Item 13	Pearson's r	0.603***	—		
	df	120	—		
	p-value	<.001	—		
Item 20	Pearson's r	0.337***	0.435***	—	
	df	121	120	—	
	p-value	<.001	<.001	—	
Item 21	Pearson's r	0.155	0.138	0.138	—
	df	121	120	121	—
	p-value	0.086	0.131	0.127	—

Note. * p < .05, ** p < .01, *** p < .001

Correlation Matrix – Side Effects

		Item 4_Reverse	Item 8_Reverse	Item 12_Reverse	Item 16_Reverse
Item 4_Reverse	Pearson's r	—			
	df	—			
	p-value	—			
Item 8_Reverse	Pearson's r	0.390***	—		
	df	120	—		
	p-value	<.001	—		
Item 12_Reverse	Pearson's r	0.426***	0.449***	—	
	df	121	120	—	
	p-value	<.001	<.001	—	
Item 16_Reverse	Pearson's r	0.208*	0.008	0.208*	—
	df	121	120	121	—
	p-value	0.021	0.927	0.021	—

Note. * p < .05, ** p < .01, *** p < .001

Fig. 14. Correlation matrix of Presence and Side Effects factors.

Multi-item Factor Analysis

We analyzed multi-item factors: Usability and Functionality, Engagement, Emotional State, Presence, and Side Effects. Internal consistency evaluation revealed adequate reliability for three factors: Usability and Functionality ($\alpha = 0.776$, $\omega = 0.785$, item-rest range: 0.488-0.657), Presence ($\alpha = 0.697$, $\omega = 0.737$, item-rest range: 0.433-0.616), and Side Effects ($\alpha = 0.639$, $\omega = 0.687$, item-rest range: 0.475-0.504). While the Side Effects factor showed an alpha slightly below the conventional 0.70 threshold, all items have adequate item-rest correlations, and McDonald's omega was close to the acceptable threshold. This supports retention of the factor, given its theoretical importance for VR healthcare safety monitoring.

The Engagement ($\alpha = 0.564$, $\omega = 0.579$) and Emotional State ($\alpha = 0.533$, $\omega = 0.575$) failed to reach an acceptable reliability threshold, indicating conceptual heterogeneity among their items. This was confirmed by inter-item correlation analyses, showing that items measured diverse aspects rather than unified constructs. Consequently, the seven items from these factors were excluded from the CFA. Table 9 shows the results of the reliability analysis. The final dataset for CFA comprised 10 items distributed across three factors.

Table 9. The table summarizes reliability analysis results for the Multi-Item Factor.

	α	**ω**	**Item**	**Item-rest (r)**	**α if dropped**	**Decision**
Usability and Functionality	0.78	0.79	1	0.65	0.69	Included in CFA
			7	0.66	0.68	
			10	0.56	0.73	
			17	0.49	0.77	
Engagment	0.55	0.59	3	0.38	0.49	Excluded from CFA
			9	0.38	0.49	
			15	0.30	0.56	
			18	0.41	0.48	
Emotional State	0.53	0.56	6	0.38	0.45	Excluded from CFA
			14	0.32	0.47	
			23	0.42	0.34	
Presence	0.70	0.74	5	0.53	0.59	Included in CFA

			13	0.62	0.49	
			20	0.43	0.75	
Side Effects	0.64	0.69	4	0.48	0.53	Included in CFA
			8	0.50	0.59	
			12	0.50	0.46	

Note. α = Cronbach's alpha; ω = McDonald's omega

The three-factor correlated model was tested using Maximum Likelihood estimation with Full Information Maximum Likelihood (FIML) for missing data. The model specified three latent factors—Usability and Functionality (4 items), Presence (3 items), and Side Effects (3 items)—with correlations among factors freely estimated. Model fit was acceptable, as shown by the chi-square test: $\chi^2(32) = 57.7$, $p = 0.004$. The CFI exceeded the acceptable threshold of 0.90, indicating good incremental fit (CFI= 0.921). The RMSEA was marginally above the 0.08 threshold (RMSEA= 0.081, 90% CI [0.046, 0.114]), while the SRMR met the acceptable criterion (SRMR = 0.074). All standardized factor loadings were significant ($p < .001$) and ranged from 0.510 to 0.852, exceeding the minimum threshold (see Table 10).

Table 10. The table summarizes standardized factor loadings from Confirmatory Factor Analysis.

Factors	Indicator	Estimate	SE	Z	p	Stand. Estimate
Usability and Functionality	Item 1	0.600	0.0686	8.75	<.001	0.750
	Item 7	0.703	0.0809	8.68	<.001	0.747
	Item 10_Reverse	0.604	0.0768	7.86	<.001	0.691
	Item 17	0.550	0.0893	6.16	<.001	0.571
Presence	Item 5	0.618	0.0851	7.26	<.001	0.700
	Item 13	0.768	0.0900	8.52	<.001	0.852
	Item 20	0.607	0.1141	5.32	<.001	0.510
Side Effects	Item 4_Reverse	0.661	0.1033	6.40	<.001	0.645
	Item 8_Reverse	0.281	0.0479	5.86	<.001	0.582
	Item 12_Reverse	0.685	0.0956	7.17	<.001	0.716

Note. Factor = latent construct; Indicator = observed item; Estimate = unstandardized factor loading (regression coefficient); SE = standard error; z = z-statistic (Estimate/SE); p = p-value for significance test; Stand. Estimate = standardized factor loading.

Analyzing factor correlations, we noted moderate associations between Usability and Presence (r = 0.321, p = 0.003) and between Usability and Side Effects (r = 0.616, p < .001), while Presence and Side Effects were uncorrelated (r = 0.044, p = 0.719). These patterns support an ecological relationship through Usability and Side Effects, whereby a well-designed system generates less physical discomfort. All factors exceeded reliability thresholds: Usability (CR = 0.786), Presence (CR = 0.736), and Side Effects (CR = 0.686). However, AVE values fell below the recommended threshold: Usability (AVE = 0.481), Presence (AVE = 0.492), and Side Effects (AVE = 0.423), indicating that the factors explained less than 50% of item variance. While this suggests limited convergent validity, the adequate factor loadings (all >0.50) and composite reliability values support the factors' internal consistency. Discriminant validity was confirmed. The √AVE for Usability (0.694) exceeded its correlations with Presence and Side Effects. Similarly, √AVE for Presence (0.701) exceeded its correlations with Usability and Side Effects. The √AVE for Side Effects (0.650) exceeded its correlation with both other factors. Although the a moderately high correlation between Usability and Side Effects, the Fornell-Larcker criterion confirms that the factors measure distinct constructs (see Table 11).

Table 11. The table summarizes data on discriminant validity.

Factor	CR	AVE	Usability and Functionality	Presence	Side Effects
Usability and Functionality	0.786	0.481	(0.694)		
Presence	0.736	0.492	0.321**	(0.701)	
Side Effects	0.686	0.423	0.616***	0.044	(0.650)

Note. Diagonal (bold, parentheses): √AVE. CR = Composite Reliability. AVE = Average Variance Extracted. **p < .01. ***p < .001

Single-item factor Analysis

Three items were designed as single-item indicators to capture specific UX dimensions: Aesthetics and Design (Item 2), Realism of Environment (Item 11), and Motivation and Intention of use (Item 24). Nomological validity was assessed by examining correlations between these items and factor scores derived from the three-factor CFA model, as well as intercorrelations among the single items themselves (Table 12).

Results confirmed theoretically expected patterns of correlations with multi-item factors. Item 11 showed a strong association with Presence ($r = 0.618$, $p < .001$), while its correlation to Usability and Functionality was modest ($r = 0.254$, $p = .005$), and there was no significant relationship with Side Effects ($r = 0.142$, $p = .119$). This pattern confirms that perceived realism is a key element for presence, more than comfort and usability factors, supporting the idea that there is a difference between how the environment looks and experiential immersion (how users feel and engage with it). The correlation magnitude ($r = 0.618$) results, appropriate for the redundancy threshold, indicating that realism captures 38% shared variance with Presence while retaining 62% unique variance. Item 2 correlates with both Usability ($r = 0.418$, $p < .001$) and Presence ($r = 0.472$, $p < .001$), along with a weaker association to Side Effects ($r = 0.281$, $p = .002$). This indicates that aesthetic quality plays a role in both the practical and experiential aspects of UX. The moderate correlation values ($r = 0.42; 0.47$) fit within the ideal range for related but distinct constructs, showing a good integration within the UX framework without any issues of overlapping with a single factor. Item 24 had a correlation with Usability ($r = 0.559$, $p < .001$), with weaker associations to Presence ($r = 0.182$, $p = .045$) and Side Effects ($r = 0.276$, $p = .002$). The relationship with usability suggests that the desire to continue using the VR system is mainly influenced by ease of use rather than immersive quality. The correlation values ($r = 0.559$) are also in the optimal range for related but non-redundant constructs.

The three single-item indicators showed consistent intercorrelations that support their distinctiveness. Aesthetics and Realism had a moderate correlation ($r = 0.501$, $p < .001$), suggesting there's some overlap between visual appeal and how true-to-life the environment feels, yet they still maintain their distinct concepts. Aesthetics were more closely correlated to Motivation and Intention of Use ($r = 0.415$, $p < .001$) compared to Realism ($r = 0.262$, $p = .003$). This might imply that aesthetic quality is a more important driver of sustained engagement than photorealistic fidelity in healthcare VR contexts. This pattern emphasizes the practical, utility-driven aspect of adopting healthcare VR, where design attractiveness and usability might be more important than immersive realism. All single-item factor correlations were within acceptable ranges, demonstrating both convergent and discriminant validity. These indicators also demonstrated adequate nomological validity through expected correlation patterns, reinforcing their value as standalone indicators that offer unique insights beyond the multi-item factors.

Table 12. Single-Item Correlations with CFA Factor Scores for Nomological Validity.

		Item 2	Item 11	Item 24
Inter-item correlation				
Item 2	Pearson's r	—		
	df	—		
	p-value	—		
Item 11	Pearson's r	0.501***	—	
	df	121	—	
	p-value	<.001	—	
Item 24	Pearson's r	0.415***	0.262**	—
	df	121	121	—
	p-value	<.001	0.003	—
Factor Score Correlations				
Usability and Functionality_score	Pearson's r	0.418***	0.254**	0.559***
	df	120	120	120
	p-value	<.001	0.005	<.001
Presence_score	Pearson's r	0.472***	0.618***	0.182*
	df	120	120	120
	p-value	<.001	<.001	0.045
SideEffect_score	Pearson's r	0.281**	0.142	0.276**
	df	120	120	120
	p-value	0.002	0.119	0.002

Note. Correlation thresholds: $r > 0.70$ = potential redundancy; $r = 0.30\text{-}0.70$ = appropriate relatedness; $r < 0.30$ = high independence. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Discussion

The present study aimed to provide the scientific community with a methodologically robust and psychometrically sound instrument for evaluating UX in VR-based healthcare applications. Through systematic validation procedures including reliability analysis, confirmatory factor analysis, and nomological validity assessment, we identified an empirically supported framework for operationalizing UX in immersive healthcare contexts.

Eight different factors were included in our first theoretical model, which was based on a thorough literature study (Bruni et al., 2025b). However, empirical analysis revealed a more condensed six-dimensional structure as more appropriate for capturing the key elements of UX in healthcare VR experiences. The final UXVR-H questionnaire comprises 13 items distributed across six factors: three multi-item factors (Usability and Functionality, 4 items; Presence, 3 items; and Side Effects, 3 items), and three single-item indicators (Aesthetics and Design, Environmental Realism, and Usage Motivation).

The retention of usability and functionality, presence, and side effects as multi-item factors confirms their central role in healthcare VR experiences. The stronger correlation observed between usability and side effects compared to the usability-presence correlation reveals an important design principle: pragmatic design quality primarily influences physical comfort and the mitigation of adverse reactions rather than experiential immersion. This pattern suggests that usability optimization serves dual critical functions for therapeutic compliance and sustained use of VR in healthcare, enhancing functional performance while simultaneously promoting comfort and tolerability. The moderate correlation between usability and presence confirms that while interface design and technological quality facilitate presence experience, presence cannot be reduced to mere usability or technical sophistication, in line with the inner presence theory (Triberti et al., 2025). Presence proves essential for therapeutic effectiveness across diverse healthcare VR applications, including pain management, exposure therapy, and cognitive rehabilitation (Riva et al., 2020; Serino et al., 2017; Triberti and Riva, 2016). This dissociation underscores the need for an integrated attention to multiple design dimensions beyond interface usability, including narrative coherence, personalization strategies, and the provision of meaningful interaction opportunities aligned with therapeutic goals.

Due to their insufficient internal consistency, two of eight initially theorized factors (Engagement and Emotional State) were eliminated throughout the validation procedure. The observed heterogeneity among items intended to measure engagement suggests that this construct may be multifaceted rather than unidimensional in VR healthcare settings. Traditional engagement measures derived from gaming contexts assess flow, temporal dissociation, and focused attention (Brockmyer et al., 2009) that we transform into questions such as 'during the experience I lost track of time' or 'I was completely focused on the experience I was having'. However, our findings suggest that healthcare VR experiences elicit qualitatively different engagement patterns. This divergence likely stems from the different motivational structures underlying healthcare versus entertainment VR use. Entertainment VR succeeds through hedonic absorption. In contrast, healthcare VR operates through utilitarian pathways where users remain goal-oriented and task-compliant. This pattern is consistent with recent literature in technology acceptance for health interventions that distinguish between hedonic and utilitarian engagement pathways (Gu et al., 2021; Venkatesh et al., 2012). This suggests that in healthcare contexts, engagement may manifest through practical engagement with therapeutic activities rather than immersive absorption. These findings are further supported by the correlation found between usage motivation and usability features rather than presence, suggesting that perceived functional utility rather than experiential immersion is the main factor driving continued use intention. This pattern directly confirms predictions from the Unified Theory of Acceptance and Use of Technology (Venkatesh and Bala, 2008), which highlights effort and performance expectancy as the main factors that influence the adoption of health technology, especially among older adult populations (Wildenbos et al., 2018). Likewise, the exclusion of emotional state raises the possibility that affective reactions to VR in healthcare are more varied than initially conceptualized. Items intended to capture emotion (e.g., boredom, frustration, satisfaction) loaded on separate factors or showed inconsistent patterns, suggesting they tap into distinct affective dimensions rather than forming a coherent emotional state factor. This observation aligns with constructed emotion theory. It underlines that emotional experiences are complex and shaped by context. In the realm of healthcare VR, emotional responses could be influenced by specific tasks, perception of therapeutic progress, and individual coping mechanisms, leading to varied patterns that defy simple measurement (Barrett, 2017). Moreover, the temporal context of measurement may have substantially influenced the pattern of emotional responses observed. This study provides a single-session VR experience, a context different from repeated interventions extending across multiple sessions. During the first experience, emotional responses likely reflect a combination of technological novelty, curiosity, and task-related challenges, with negative emotions remaining muted. However, this pattern may shift considerably across repeated exposures, as demonstrated in longitudinal research (Karapanos et al., 2009). Frustration with persistent usability issues and boredom with repetitive content tend to emerge after multiple sessions, experiencing emotional patterns qualitatively distinct from first-session responses.

Similarly, the pattern observed for side effects should be influenced by the context of both application characteristics and temporal dynamics. Minimal visual design and straightforward interaction mechanics, while appropriate for the target population, may have reduced the likelihood of cybersickness symptoms compared to more visually complex or motion-intensive VR applications. However, cybersickness susceptibility may also exhibit temporal patterns, with some individuals experiencing increasing symptoms across repeated exposures. These application-specific characteristics should be considered when interpreting the validation findings and assessing the generalizability of the UXVR-H structure.

Despite some changes, the structure of UXVR-H offers a balance between psychometric rigor and practical efficiency, providing a comprehensive UX assessment while minimizing respondent burden. This research also contributes to the ongoing debate about the complexity of UX constructs and the necessity for precise and verified models. We demonstrated the improperness of theoretical general UX frameworks to capture specialized areas such as healthcare and VR. Our results emphasize the importance of context-aware UX measurement and conceptualization, particularly when transferring ideas across various application domains with different user goals and success criteria.

Another important contribution of our tool concerns practical implications for researchers, developers, and clinicians working with VR-based healthcare interventions. First, its brevity (13 items, approximately 3-5 minutes administration time) makes the UXVR-H questionnaire feasible for integration into clinical protocols where time constraints, patient cognitive load, and burden are challenging. Second, the instrument's structure provides actionable guidance for VR healthcare application design. The predominance of functional performance as a core dimension and its high correlation with usage motivation emphasize that ease of use is the primary driver of sustained engagement in healthcare contexts. Instead of prioritizing visually appealing or immersive elements, developers should prioritize fluid interaction mechanics, clear task instructions, and intuitive interfaces. This does not mean aesthetics are irrelevant—our single-item indicator confirms they contribute to overall UX - but rather that functional usability provides the foundation upon which other experiential qualities are built.

Although we provide an advantageous instrument, several limitations warrant acknowledgment. First, the sample comprised older adults recruited from a clinical setting, which may limit generalizability to younger populations or non-clinical VR applications. While this sampling strategy was appropriate given the instrument's intended healthcare application context and the demographic reality that many VR-based health interventions target aging populations, future validation studies should examine measurement invariance across age groups, clinical conditions, and application types to establish the instrument's broader applicability. Second, participants experienced a specific VR application focused on cognitive training tasks, which may have influenced the salience and interpretation of certain UX dimensions. The specific technological configuration (HMD-based immersive VR with controller interaction) represents only one modality within the broader spectrum of VR technologies (e.g., cave automatic virtual environments, augmented reality, mixed reality). Future research should validate the UXVR-H across diverse VR healthcare applications and technological configurations to establish its robustness across the heterogeneous healthcare VR landscape. Third, the exclusion of Engagement and Emotional State factors, while empirically supported, suggests these constructs may require reconceptualization rather than abandonment. Moreover, emotional responses, engagement patterns, and even cybersickness susceptibility may evolve considerably across repeated VR exposures. Future research employing repeated-measures designs across multiple sessions would illuminate how UX factor structure and item performance evolve, potentially revealing that certain constructs require distinct operationalizations for initial versus sustained use contexts, and whether initial symptom profiles predict longer-term tolerance, enabling personalized approaches to exposure duration and session frequency. Finally, the limited convergent validity and moderate sample size suggest that replication with larger samples would provide more stable parameter estimates and potentially improve model fit.

DISCUSSION

Within the framework of UX, this concept remains highly heterogeneous, characterized by different disciplinary conceptualizations without consensus on clear operational definitions (Abro et al., 2015; Basri et al., 2016). This theoretical fragmentation contrasts with the practical need to carefully evaluate the interaction resulting from the use of specific products such as VR-based applications. Examining UX can inform clinicians about important aspects, including identification of usability barriers, accommodation of patients' specific needs and limitations, enhanced treatment adherence, sustained patient engagement, and long-term adoption of proposed intervention. Despite the rapid expansion of VR applications in healthcare (Abdelmaged, 2021), theoretical frameworks and instruments supporting UX assessment remain scarce. Factors such as peculiar VR features (e.g., presence), users' peculiarity characteristics, including cognitive limitations or age-relevant vulnerabilities, require consideration.

While several methodologies exist (e.g., physiological, observational, and self-report methods) (Inan Nur et al., 2021), questionnaires represent the best practice due to their optimal combination of efficiency, ease of administration, and adaptability to clinical contexts. These features make them particularly suitable for aging and patients who frequently experience burden and fatigue, limiting adherence to more sophisticated assessment solutions (Kula et al., 2019; Schrepp, 2015). Questionnaires also prove appropriate for healthcare environments where time constraints are important, while offering standardized data collection that facilitates quantitative analysis and cross-study comparison. However, available questionnaires present several challenges: most were originally developed for two-dimensional interfaces and subsequently adapted for use in VR contexts without appropriate theoretical reconsideration (Bruni et al., 2025b). Moreover, most instruments lack rigorous psychometric validation, particularly for aging populations.

Based on this evidence, the present doctoral dissertation addressed these gaps through four complementary studies, creating a theoretical framework and an evaluation tool (the UXVR-H) that is tailored for clinical settings with rigorous psychometric requirements. This work provides significant contributions from both theoretical and methodological perspectives.

The theoretical contribution of this thesis arises from a scientometric analysis documenting UX's transformation from a technical discipline to an interdisciplinary field, identifying a progressive expansion from purely functional optimization toward holistic consideration of the human-technology relationship (Triberti and Brivio, 2017), recognizing the complexity of contemporary digital experiences (Bruni et al., 2025a). The scoping review enabled identification of key dimensional factors assessed in VR applications. Building upon this theoretical framework, we used a rigorous, multi-phase methodological approach to develop the User eXperience questionnaire for Virtual Reality applications in Healthcare (UXVR-H). To ensure accessibility and methodological rigor, the construction process combined user-centered assessments with expert consultation, as well as psychometric analysis. The UXVR-H questionnaire comprises 13 items distributed across six factors: usability and functionality, presence, side Effects, aesthetics and design, realism of environment, motivation and intention of use.

To our knowledge, this is the first questionnaire created especially for VR experiences in healthcare within the Italian linguistic and cultural context, constructed through rigorous methodological processes rather than mere translation. Beyond its theoretical contributions, this research has a practical impact. The framework and derived tool have the potential to enhance the design of VR applications for healthcare by offering more precise and pertinent user feedback. Furthermore, the standardized evaluation method makes it easier to compare different VR solutions in a meaningful way, which could improve evidence-based decision-making when it comes to the use of healthcare technology. The UXVR-H questionnaire emerges as a promising instrument balanced between respondent burden considerations and coverage of key

UX dimensions. Its brevity (13 items), adequate reliability, acceptable model fit and nomological validity position UXVR-H as a valuable proposal to the limited collection of VR-specific assessment tools. Moreover, our work demonstrated that UX dimensionality in general-purpose VR applications might not be the same as in healthcare settings. When assessing overall experience and usage intentions, pragmatic factors (comfort, usability) prove more critical than hedonistic ones (engagement, emotional intensity). This pattern supports the idea that VR experience in healthcare necessitates specialized evaluation frameworks, considering the priorities and characteristics of clinical contexts than adaptation of generic UX measures (Gu et al., 2021; Venkatesh et al., 2012).

Notwithstanding these contributions, it is important to note a few limitations. First, it is important to acknowledge the inherent limitations of questionnaire-based evaluation. While quantitative measures can indicate positive or negative UX, they do not provide direct guidance on specific improvement pathways. Although the UXVR-H points out problematic areas, it fails to capture which components need to be improved or redesigned. Therefore, to provide the information required for actionable insight and system optimization, quantitative assessment should always be integrated with qualitative methods. Second, consistent with the current literature (Maia and Furtado, 2016), this instrument focuses on behavioral episodes and momentary assessments, failing to capture the evolution of experience over prolonged interaction. The temporal dimension of UX (i.e., how perceptions and attitudes change over time) remains unexamined. Similarly, standardized questionnaires fail to consider the situated context in which technology use actually occurs. By design, questionnaires employ fixed items and response formats intended for consideration across diverse users and settings, thereby prioritizing reliability and comparability rather than contextual factors that profoundly shape real-world technology use (Triberti and Brivio, 2017). In healthcare specifically, numerous contextual variables might influence UX but remain invisible to standardized instruments. For example, the questionnaire cannot reveal that patients felt rushed because the VR session occurred immediately before another scheduled appointment, or that they hesitated to report discomfort due to concerns about disappointing the therapist. These contextual factors could be potentially decisive for therapeutic engagement and long-term adoption. Contextual inquiry methods, such as situated interviews, should complement questionnaire-based assessment by revealing how technology becomes integrated (or fails to integrate) within the complex ecology of healthcare practice (Holtzblatt and Beyer, 1997). Finally, without cultural adaptation and validation, the cross-cultural validity and international applicability were limited because it was created especially for the Italian aging context. This cultural specificity, while ensuring relevance for the target population, restricts generalizability to other linguistic and cultural contexts. Cross-cultural differences in perceptions of technology, healthcare expectations, and communication styles may influence item interpretation and the factor structure, necessitating careful validation before international implementation. However, while the present questionnaire is specifically tailored for the aging population in clinical contexts, the methodological principles remain transferable to different populations and settings. For instance, adaptation in the pediatric populations would require substantial modifications during the design phase (e.g., simplifying response scales or involving both children and their caregivers in validation processes), yet the systematic development framework provides a repeatable template. The contextual specificity of UXVR-H represents both a limitation in terms of generalizability and a strength in ensuring relevance for targets. Importantly, the development methodology itself constitutes a generalizable contribution: the systematic process of development, from theoretical conceptualization to empirical validation, provides a rigorous model that can guide future instrument development.

Furthermore, the current validation study represents an initial step rather than a definitive psychometric evaluation of the instrument, and additional analyses and considerations are needed to establish a complete psychometric and clinical foundation.

Future developments should prioritize several key areas. First of all, there is a need for comprehensive psychometric validation of the UXVR-H through larger samples. Our findings would benefit from integration into a unified structural equation model incorporating all eight theoretical factors simultaneously, using advanced modeling techniques to accommodate both multi-item and single-item constructs. Cross-cultural validation across different linguistic and cultural contexts would ensure international applicability. Additionally, acknowledging the critical dimension of time in UX research (Kujala et al., 2011; Law et al., 2009), longitudinal research examining how UX evolves through multiple sessions and prolonged interventions would provide valuable insights, particularly for clinical populations such as chronic conditions who may require extended therapeutic engagement with VR systems. Such temporal analyses could reveal adaptation patterns and long-term acceptance factors not captured in single-session evaluations. Further, criterion validity studies examining relationships between UXVR-H scores and objective outcomes - including treatment adherence, therapeutic efficacy, and technology adoption rates - would establish the instrument's predictive utility for clinical practice. Finally, examination of the instrument's performance across diverse VR application types, such as pain management or exposure therapy, would establish generalizability boundaries and potentially identify deeper application-specific needs to enhance measurement precision.

With continued refinement, the UXVR-H has the potential to become a standard assessment tool supporting evidence-based VR healthcare development and evaluation. The methodological framework also provides a perspective for further investigations. Considering the presented design model, researchers might develop some variants of the UXVR-H, providing a family of instruments specific to populations and contexts. This tool family perspective would preserve methodological comparability while allowing the contextual sensitivity needed to capture the nuances of user experience across different application domains.

CONCLUSION

This doctoral thesis successfully bridged User Experience, Virtual Reality, and Healthcare, providing both theoretical foundations and practical tools for evaluating innovative clinical experiences. The eight-factor framework and UXVR-H questionnaire represent significant methodological contributions guiding future development of patient-centered VR technologies. Similarly, the rigorous process design represents a template for future instruments.

While representing initial steps toward comprehensive understanding, the present research provides a robust platform for future research due to its advancement in theoretical, practical, and methodological acknowledgment. As the field evolves rapidly, the UXVR-H offers standardized assessment, facilitating evidence-based decision-making and cross-study comparisons. Ultimately, this work advances personalized, effective healthcare interventions, supporting VR technologies that are sophisticated yet genuinely responsive to patient needs, capabilities, and experiences.

References

Abbasi, M.Q., Weng, J., Wang, Y., Rafique, I., Wang, X., Lew, P., 2012. Modeling and evaluating user interface aesthetics employing ISO 25010 quality standard, in: 2012 Eighth International Conference on the Quality of Information and Communications Technology. IEEE, pp. 303–306.

Abdelmaged, M.A.M., 2021. Implementation of virtual reality in healthcare, entertainment, tourism, education, and retail sectors.

Abro, A., Sulaiman, S., Mahmood, A.K., Khan, M., 2015. Understanding factors influencing User Experience of interactive systems: A literature review 10, 18175–18185.

Alashram, A.R., Annino, G., Padua, E., Romagnoli, C., Mercuri, N.B., 2019. Cognitive rehabilitation post traumatic brain injury: A systematic review for emerging use of virtual reality technology. Journal of Clinical Neuroscience 66, 209–219. https://doi.org/10.1016/j.jocn.2019.04.026

Alben, L., 1996. Quality of experience: defining the criteria for effective interaction design. interactions 3, 11–15.

Albert, B., Tullis, T., 2022. Measuring the user experience: Collecting, analyzing, and presenting UX metrics. Morgan Kaufmann.

AlGerafi, M.A.M., Zhou, Y., Oubibi, M., Wijaya, T.T., 2023. Unlocking the potential: A comprehensive evaluation of augmented reality and virtual reality in education. Electronics (Basel) 12, 3953.

Ames, S.L., Wolffsohn, J.S., McBrien, N.A., 2005. The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display. Optom Vis Sci 82, 168–176. https://doi.org/10.1097/01.opx.0000156307.95086.6

Aria, M., Cuccurullo, C., 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. J Informetr 11, 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Augstein, M., Neumayr, T., Kern, D., Kurschl, W., Altmann, J., Burger, T., 2017. An analysis and modeling framework for personalized interaction, in: Companion Proceedings of the 22nd International Conference on Intelligent User Interfaces. pp. 57–60.

Backhaus, N., Trapp, A.K., Thüring, M., 2018. Skeuomorph versus flat design: user experience and age-related preferences, in: Design, User Experience, and Usability: Designing Interactions: 7th International Conference, DUXU 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part II 7. Springer, pp. 527–542.

Barrett, L.F., 2017. The theory of constructed emotion: an active inference account of interoception and categorization. Soc Cogn Affect Neurosci 12, 1–23.

Basri, N.H., Mara, K.P., Noor, N.L., Adilah, W., Adnan, W., 2016. Conceptualizing and Understanding User Experience. 2016 4th International Conference on User Science and Engineering (i-USEr) 81–84. https://doi.org/10.1109/IUSER.2016.7857938

Bell, I.H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., Valmaggia, L., 2020. Virtual reality as a clinical tool in mental health research and practice. Dialogues Clin Neurosci 22, 169–177.

Berni, A., Borgianni, Y., 2021. From the definition of user experience to a framework to classify its applications in design. Proceedings of the Design Society 1, 1627–1636. https://doi.org/10.1017/pds.2021.424

Borgnis, F., Baglio, F., Pedroli, E., Rossetto, F., Isernia, S., Uccellatore, L., Riva, G., Cipresso, P., 2021. EXecutive-Functions Innovative Tool (EXIT 360◦): A Usability and User Experience Study of an Original 360◦-Based Assessment Instrument. Sensors 21. https://doi.org/https://doi.org/10.3390/s21175867

Borgnis, F., Baglio, F., Pedroli, E., Rossetto, F., Meloni, M., Riva, G., Cipresso, P., 2022. A Psychometric Tool for Evaluating Executive Functions in Parkinson's Disease. J Clin Med 11. https://doi.org/10.3390/jcm11051153

Brandes, U., 2001. A faster algorithm for betweenness centrality. Journal of mathematical sociology 25, 163–177.

Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K.M., Pidruzny, J.N., 2009. The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing. J Exp Soc Psychol 45, 624–634.

Brooke, J., 1996. SUS: A “Quick and Dirty” Usability Scale. Usability Evaluation In Industry 207–212. https://doi.org/10.1201/9781498710411-35

Bruni, F., Borghesi, F., Mancuso, V., Cipresso, P., Pedroli, E., 2025a. A Scientometric Analysis of User Experience: Exploring Evolution, Emerging Themes, and Impacts. Int J Hum Comput Interact 1–20.

Bruni, F., Mancuso, V., Greci, L., Arlati, S., Cavallo, M., Riva, G., Goulene, K., Stramba-Badiale, M., Pedroli, E., 2023. A cross-platform application for the ecological and remote assessment of memory impairment in aging: ECO-MEMORY. Virtual Real 27, 2757–2767. https://doi.org/10.1007/s10055-023-00826-8

Bruni, F., Mancuso, V., Pedroli, E., 2025b. How to evaluate User Experience in Digital Health? A Scoping Review of questionnaires in Virtual Reality Applications. Front Digit Health 7, 1561364.

Bruni, F., Mancuso, V., Stramba-Badiale, C., Greci, L., Cavallo, M., Borghesi, F., Riva, G., Cipresso, P., Stramba-Badiale, M., Pedroli, E., 2022. ObReco-2: Two-step validation of a tool to assess memory deficits using 360° videos. Front Aging Neurosci 14, 1–10. https://doi.org/10.3389/fnagi.2022.875748

Busselle, R., Bilandzic, H., 2009. Measuring narrative engagement. Media Psychol 12, 321–347.

Callon, M., Courtial, J.-P., Turner, W.A., Bauin, S., 1983. From translations to problematic networks: An introduction to co-word analysis. Social science information 22, 191–235.

Cavedoni, S., Cipresso, P., Mancuso, V., Bruni, F., Pedroli, E., 2022. Virtual reality for the assessment and rehabilitation of neglect: where are we now? A 6-year review update, Virtual Reality. Springer London. https://doi.org/10.1007/s10055-022-00648-0

Chang, E., Hyun Taek, K., and Yoo, B., 2020. Virtual Reality Sickness: A Review of Causes and Measurements. Int J Hum Comput Interact 36, 1658–1682. https://doi.org/10.1080/10447318.2020.1778351

Che, X., Ma, S., Yu, Q., Yue, C., 2019. An event-based user experience evaluation method for virtual reality applications. Proceedings - 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Internet of People and Smart City Innovation, SmartWorld/UIC/ATC/SCALCOM/IOP/SCI 2019 767–774. https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00164

Chen, M., Tang, Q., Xu, S., Leng, P., Pan, Z., 2020. Design and evaluation of an augmented reality-based exergame system to reduce fall risk in the elderly. Int J Environ Res Public Health 17, 1–10. https://doi.org/10.3390/ijerph17197208

Chiorri, C., 2011. Teoria e tecnica psicometrica: costruire un test psicologico. McGraw-Hill.

Cho, K.H., Kim, M.K., Lee, H.J., Lee, W.H., 2015. Virtual reality training with cognitive load improves walking function in chronic stroke patients. Tohoku Journal of Experimental Medicine 236, 273–280. https://doi.org/10.1620/tjem.236.273

Cipresso, P., Giglioli, I.A.C., Raya, M.A., Riva, G., 2018. The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature . Frontiers in Psychology .

Colombo, V., Mondellini, M., Fumagalli, A., Aliverti, A., Sacco, M., 2024. A virtual reality-based endurance training program for COPD patients: acceptability and user experience. Disabil Rehabil Assist Technol 19, 1590–1599. https://doi.org/10.1080/17483107.2023.2219699

Davis, F.D., Bagozzi, R.P., Warshaw, P.R., 1989. User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Manage Sci 35, 982–1003. https://doi.org/10.1287/mnsc.35.8.982

D'Cunha, N.M., Nguyen, D., Naumovski, N., McKune, A.J., Kellett, J., Georgousopoulou, E.N., Frost, J., Isbel, S., 2019. A Mini-Review of Virtual Reality-Based Interventions to Promote Well-Being for People Living with Dementia and Mild Cognitive Impairment. Gerontology 65, 430–440. https://doi.org/10.1159/000500040

de Souza Lima, A.L., von Wangenheim, C.G., Borgatto, A.F., 2022. Assessment of visual aesthetics through human judgments: a systematic mapping, in: Proceedings of the 21st Brazilian Symposium on Human Factors in Computing Systems. pp. 1–14.

Deci, E.L., Eghrari, H., Patrick, B.C., Leone, D.R., 1994. Facilitating internalization: The self-determination theory perspective. J Pers 62, 119–142.

Devilly, G.J., Borkovec, T.D., 2000. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 31, 73–86.

Emmelkamp, P.M.G., Meyerbröker, K., 2021. Virtual reality therapy in mental health. Annu Rev Clin Psychol 17, 495–519.

Engeser, S., Rheinberg, F., Vollmeyer, R., Bischoff, J., 2005. Motivation, flow-Erleben und Lernleistung in universitären Lernsettings. Zeitschrift für pädagogische Psychologie 19, 159–172.

Espinola, C.W., Nguyen, B., Torres, A., Sim, W., Rueda, A., Beavers, L., Campbell, D.M., Jung, H., Lou, W., Kapralos, B., Peter, E., Dubrowski, A., Krishnan, S., Bhat, V., 2024. Digital Interventions for Stress Among Frontline Health Care Workers: Results From a Pilot Feasibility Cohort Trial. JMIR Serious Games 12, 1–16. https://doi.org/10.2196/42813

Familoni, B.T., Odetunde Babatunde, S., 2024. USER EXPERIENCE (UX) DESIGN IN MEDICAL PRODUCTS: THEORETICAL FOUNDATIONS AND DEVELOPMENT BEST PRACTICES. Engineering Science & Technology Journal 5, 1125–1148. https://doi.org/10.51594/estj/v5i3.975

Finstad, K., 2010. The usability metric for user experience. Interact Comput 22, 323–327.

Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. Journal of marketing research 18, 39–50.

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., Slater, M., 2017. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 47, 2393–2400.

Fried, L.P., Tangen, C.M., Walston, J., Newman, A.B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W.J., Burke, G., 2001. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 56, M146–M157.

Fu, B., Noy, N.F., Storey, M.-A., 2016. Eye tracking the user experience–an evaluation of ontology visualization techniques. Semant Web 8, 23–41.

Galavi, Z., Norouzi, S., Khajouei, R., 2024. Heuristics used for evaluating the usability of mobile health applications: A systematic literature review. Digit Health 10. https://doi.org/10.1177/20552076241253539

Gallagher, M., Ferrè, E.R., 2018. Cybersickness: a multisensory integration perspective. Multisens Res 31, 645–674.

Garfield, E., 2004. Historiographic mapping of knowledge domains literature. J Inf Sci 30, 119–145.

Gaver, B., Martin, H., 2000. Alternatives: exploring information appliances through conceptual design proposals, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 209–216.

Glanznig, M., 2012. User experience research: Modelling and describing the subjective. Interdisciplinary Description of Complex Systems: INDECS 10, 235–247.

Godovykh, M., Tasci, A.D.A., 2020. Customer experience in tourism: A review of definitions, components, and measurements. Tour Manag Perspect 35, 100694.

Goumopoulos, C., Drakakis, E., Gklavakis, D., 2023. Feasibility and Acceptance of Augmented and Virtual Reality Exergames to Train Motor and Cognitive Skills of Elderly. Computers 12. https://doi.org/10.3390/computers12030052

Greene, S., Thapliyal, H., Caban-Holt, A., 2016. A survey of affective computing for stress detection: Evaluating technologies in stress detection for better health. IEEE Consumer Electronics Magazine 5, 44–56.

Gu, D., Khan, S., Khan, I.U., Khan, S.U., Xie, Y., Li, X., Zhang, G., 2021. Assessing the adoption of e-health technology in a developing country: an extension of the UTAUT model. Sage Open 11, 21582440211027564.

Hair, J.F., 2009. Multivariate data analysis.

Hamzeheinejad, N., Roth, D., Gotz, D., Weilbach, F., Latoschik, M.E., 2019. Physiological effectivity and user experience of immersive gait rehabilitation. 26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings 1421–1429. https://doi.org/10.1109/VR.2019.8797763

Hart, J., Sutcliffe, A., De Angeli, A., 2012. Using affect to evaluate user engagement, in: CHI'12 Extended Abstracts on Human Factors in Computing Systems. pp. 1811–1834.

Hartmann, J., Sutcliffe, A., Angeli, A. De, 2008. Towards a theory of user judgment of aesthetics and user interface quality. ACM Transactions on Computer-Human Interaction (TOCHI) 15, 1–30.

Hassenzahl, M., 2013. User experience and experience design. The encyclopedia of human-computer interaction 2, 1–14.

Hassenzahl, M., 2010. Experience design: Technology for all the right reasons. Morgan & Claypool Publishers.

Hassenzahl, M., 2007. The hedonic/pragmatic model of user experience. Towards a UX manifesto 10, 2007.

Hassenzahl, M., 2004. The interplay of beauty, goodness, and usability in interactive products. Hum Comput Interact 19, 319–349.

Hassenzahl, M., Diefenbach, S., Göritz, A., 2010. Needs, affect, and interactive products–Facets of user experience. Interact Comput 22, 353–362.

Hassenzahl, M., Tractinsky, N., 2006. User experience - A research agenda. Behaviour and Information Technology 25, 91–97. https://doi.org/10.1080/01449290500330331

Hinderks, A., Meiners, A.-L., Mayo, F.J.D., Thomaschewski, J., 2019. Interpreting the Results from the User Experience Questionnaire (UEQ) using Importance-Performance Analysis (IPA)., in: WEBIST. p. 15th.

Holden, R.J., Karsh, B.-T., 2010. The technology acceptance model: its past and its future in health care. J Biomed Inform 43, 159–172.

Holtzblatt, K., Beyer, H., 1997. Contextual design: defining customer-centered systems. Elsevier.

Hussain, J., Ali Khan, W., Hur, T., Muhammad Bilal, H.S., Bang, J., Ul Hassan, A., Afzal, M., Lee, S., 2018. A multimodal deep log-based user experience (UX) platform for UX evaluation. Sensors 18, 1622.

Huygelier, H., Schraepen, B., Lafosse, C., Vaes, N., Schillebeeckx, F., Michiels, K., Note, E., Vanden Abeele, V., van Ee, R., Gillebert, C.R., 2022. An immersive virtual reality game to train spatial attention orientation after stroke: A feasibility study. Applied Neuropsychology:Adult 29, 915–935. https://doi.org/10.1080/23279095.2020.1821030

Inan Nur, A., B. Santoso, H., O. Hadi Putra, P., 2021. The method and metric of user experience evaluation: a systematic literature review, in: Proceedings of the 2021 10th International Conference on Software and Computer Applications. pp. 307–317.

Iso, W., 1998. 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs). The international organization for standardization 45, 22.

Izard, S.G., Juanes, J.A., García Peñalvo, F.J., Estella, J.M.G., Ledesma, M.J.S., Ruisoto, P., 2018. Virtual reality as an educational and training tool for medicine. J Med Syst 42, 1–5.

Jackson, S.A., Marsh, H.W., 1996. Development and validation of a scale to measure optimal experience: The Flow State Scale. J Sport Exerc Psychol 18, 17–35.

Jespersen, L.N., Michelsen, S.I., Holstein, B.E., Tjørnhøj-Thomsen, T., Due, P., 2018. Conceptualization, operationalization, and content validity of the EQOL-questionnaire measuring quality of life and participation for persons with disabilities. Health Qual Life Outcomes 16. https://doi.org/10.1186/s12955-018-1024-6

Jordan, P.W., 2000. Designing pleasurable products: An introduction to the new human factors. CRC press.

Karapanos, E., Hassenzahl, M., Martens, J.-B., 2008. User Experience Over Time 3561–3566.

Karapanos, E., Zimmerman, J., Forlizzi, J., Martens, J.-B., 2009. User experience over time: an initial framework, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 729–738.

Keebler, J.R., Shelstad, W.J., Smith, D.C., Chaparro, B.S., Phan, M.H., 2020. Validation of the GUESS-18: a short version of the Game User Experience Satisfaction Scale (GUESS). J Usability Stud 16, 49.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G., 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. Int J Aviat Psychol 3, 203–220.

Kiper, P., Szczudlik, A., Agostini, M., Opara, J., Nowobilski, R., Ventura, L., Tonin, P., Turolla, A., 2018. Virtual Reality for Upper Limb Rehabilitation in Subacute and Chronic Stroke: A Randomized Controlled Trial. Arch Phys Med Rehabil 99, 834-842.e4. https://doi.org/10.1016/j.apmr.2018.01.023

Kline, R.B., 2023. Principles and practice of structural equation modeling. Guilford publications.

Koch, S.H., Proynova, R., Paech, B., Wetter, T., 2013. How to approximate users' values while preserving privacy: experiences with using attitudes towards work tasks as proxies for personal value elicitation. Ethics Inf Technol 15, 45–61.

Kollmorgen, J., Schrepp, M., Thomaschewski, J., 2022. Influence of demographic variables and usage behaviour on the perceived user experience, in: International Conference on Web Information Systems and Technologies. Springer, pp. 186–208.

Kourtesis, P., Collina, S., Doumas, L.A.A., MacPherson, S.E., 2019. Validation of the Virtual Reality Neuroscience Questionnaire: Maximum Duration of Immersive Virtual Reality Sessions Without the Presence of Pertinent Adverse Symptomatology. Front Hum Neurosci 13, 1–13. https://doi.org/10.3389/fnhum.2019.00417

Kujala, S., 2013. Lost in Time : The Meaning of Temporal Aspects in User Experience 559–564.

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Sinnelä, A., 2011. Identifying hedonic factors in long-term user experience, in: Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces. pp. 1–8.

Kula, I., Branaghan, R.J., Atkinson, R.K., Roscoe, R.D., 2019. Assessing user experience via biometric sensor affect detection, in: Human Performance Technology: Concepts, Methodologies, Tools, and Applications. IGI Global, pp. 698–714.

Kumari, S., Polke, N., 2018. Implementation issues of augmented reality and virtual reality: A survey, in: International Conference on Intelligent Data Communication Technologies and Internet of Things. Springer, pp. 853–861.

Kurosu, M., Hashizume, A., Ueno, Y., 2018. User experience evaluation by ERM: experience recollection method, in: Human-Computer Interaction. Theories, Methods, and Human Issues: 20th International Conference, HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part I 20. Springer, pp. 138–147.

Kyriazos, T.A., 2018. Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. Psychology 9, 2207.

Lange, B.S., Requejo, P., Flynn, S.M., Rizzo, A.A., Valero-Cuevas, F.J., Baker, L., Winstein, C., 2010. The potential of virtual reality and gaming to assist successful aging with disability. Physical Medicine and Rehabilitation Clinics 21, 339–356.

Laugwitz, B., Held, T., Schrepp, M., 2008. Construction and Evaluation of a User Experience Questionnaire BT - HCI and Usability for Education and Work, in: Holzinger, A. (Ed.), . Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 63–76.

Lavie, T., Tractinsky, N., 2004. Assessing dimensions of perceived visual aesthetics of web sites. Int J Hum Comput Stud 60, 269–298.

Law, E., Roto, V., Kort, J., Technology, C., Hassenzahl, M., Psychology, E., 2008. Towards a Shared Definition of User Experience 2395–2398.

Law, E.L., Leicester, L.E., Hassenzahl, M., 2009. Understanding , Scoping and Defining User eXperience : A Survey Approach 719–728.

Lawshe, C.H., 1975. A quantitative approach to content validity. Pers Psychol 28, 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lemle, E., Bomkamp, K., Williams, M.K., Cutbirth, E., Lemle, E., Bomkamp, K., Williams, M.K., Cutbirth, E., 2015. Virtual reality and the future of entertainment. Two bit circus and the future of entertainment 25–37.

Lessiter, J., Freeman, J., Keogh, E., Davidoff, J., 2001. A cross-media presence questionnaire: The ITC-sense of presence inventory. Presence: Teleoperators and Virtual Environments 10, 282–297. https://doi.org/10.1162/105474601300343612

Lewis, C., 1982. Using the “thinking-aloud” method in cognitive interface design. Yorktown Heights, N.Y. : IBM T.J. Watson Research Center,.

Lewis, J.R., 2014. Usability: lessons learned… and yet to be learned. Int J Hum Comput Interact 30, 663–684.

Lewis, J.R., Sauro, J., 2018. Item benchmarks for the system usability scale. J Usability Stud 13.

Li, Q., Chen, Z., Wu, T., 2024. Examining the Influence of Front-Facing Camera Layout on the Aesthetic Experience of Smartphone User Interfaces, in: International Conference on Human-Computer Interaction. Springer, pp. 164–170.

Li, R., Zhang, H., Liu, C., Qian, Z.C., Zhang, L., 2022. Bibliometric and Visualized Analysis of User Experience Design Research: From 1999 to 2019. Sage Open 12, 21582440221087264. https://doi.org/10.1177/21582440221087266

Liao, D., Shu, L., Huang, Y., Yang, J., Xu, X., 2018. Scenes Design in Virtual Reality for Depression Assessment, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer International Publishing. https://doi.org/10.1007/978-3-319-91584-5_10

Liao, J., Lou, J., Wu, Q., Zou, M., Zheng, L., 2018. A review of age-related characteristics for touch-based performance and experience, in: Human Aspects of IT for the Aged Population. Acceptance, Communication and Participation: 4th International Conference, ITAP 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part I 4. Springer, pp. 279–296.

Lin, M.K., Nicolini, O., Waxenegger, H., Galloway, G.J., Ullmann, J.F.P., Janke, A.L., 2013. Interpretation of medical imaging data with a mobile application: a mobile digital imaging processing environment. Front Neurol 4, 85. https://doi.org/10.3389/fneur.2013.00085

Liu, Q., Wang, Y., Yao, M.Z., Tang, Q., Yang, Y., 2020. The Effects of Viewing an Uplifting 360-Degree Video on Emotional Well-Being Among Elderly Adults and College Students Under Immersive Virtual Reality and Smartphone Conditions. Cyberpsychol Behav Soc Netw 23, 157–164. https://doi.org/10.1089/cyber.2019.0273

Liu, Y., 2020. A Scientometric Analysis of User Experience Research Related to Green and Digital Transformation, in: 2020 Management Science Informatization and Economic Innovation Development Conference (MSIEID). pp. 377–380. https://doi.org/10.1109/MSIEID52046.2020.00080

Liu, Y.-J., Yu, M., Zhao, G., Song, J., Ge, Y., Shi, Y., 2017. Real-time movie-induced discrete emotion recognition from EEG signals. IEEE Trans Affect Comput 9, 550–562.

Liu, Z., He, Z., Yuan, J., Lin, H., Fu, C., Zhang, Y., Wang, N., Li, G., Bu, J., Chen, M., Jia, J., 2023. Application of Immersive Virtual-Reality-Based Puzzle Games in Elderly

Patients with Post-Stroke Cognitive Impairment: A Pilot Study. Brain Sci 13. https://doi.org/10.3390/brainsci13010079

Lorentz, L., Simone, M., Zimmermann, M., Studer, B., Suchan, B., Althausen, A., Estocinova, J., Müller, K., Lendt, M., 2023. Evaluation of a VR prototype for neuropsychological rehabilitation of attentional functions. Virtual Real 27, 187–199. https://doi.org/10.1007/s10055-021-00534-1

Lun, L., Zetian, D., Hoe, T.W., Juan, X., Jiaxin, D., Fulai, W., 2024. Factors influencing user intentions on interactive websites: Insights from the technology acceptance model. IEEE Access.

Lutz, W., Sanderson, W., Scherbov, S., 2008. The coming acceleration of global population ageing. Nature 451, 716–719.

Maia, C.L.B., Furtado, E.S., 2016. A systematic review about user experience evaluation, in: Design, User Experience, and Usability: Design Thinking and Methods: 5th International Conference, DUXU 2016, Held as Part of HCI International 2016, Toronto, Canada, July 17–22, 2016, Proceedings, Part I 5. Springer, pp. 445–455.

Mancuso, V., Borghesi, F., Bruni, F., Pedroli, E., Cipresso, P., 2024. Mapping the landscape of research on 360-degree videos and images: a network and cluster analysis. Virtual Real 28. https://doi.org/10.1007/s10055-024-01002-2

Matus, N., Rusu, C., Cano, S., 2021. Student eXperience: a systematic literature review. Applied Sciences 11, 9543.

McAuley, E., Duncan, T., Tammen, V. V, 1989. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. Res Q Exerc Sport 60, 48–58.

Meschberger-Annweiler, F.A., Ascione, M., Porras-Garcia, B., Ferrer-Garcia, M., Moreno-Sanchez, M., Miquel-Nabau, H., Serrano-Troncoso, E., Carulla-Roig, M., Gutiérrez-Maldonado, J., 2023. An Attentional Bias Modification Task, through Virtual Reality and Eye-Tracking Technologies, to Enhance the Treatment of Anorexia Nervosa. J Clin Med 12. https://doi.org/10.3390/jcm12062185

Milgram, P., Takemura, H., Utsumi, A., Kishino, F., 1995. Augmented reality: A class of displays on the reality-virtuality continuum, in: Telemanipulator and Telepresence Technologies. Spie, pp. 282–292.

Minge, M., Fr, S., 2008. Dynamics of user experience. ... Goals and Strategies for Studying User Experience ... 2–6.

Mirelman, A., Rochester, L., Reelick, M., Nieuwhof, F., Pelosin, E., Abbruzzese, G., Dockx, K., Nieuwboer, A., Hausdorff, J.M., 2013. V-TIME: A treadmill training program augmented by virtual reality to decrease fall risk in older adults: Study design of a randomized controlled trial. BMC Neurol 13. https://doi.org/10.1186/1471-2377-13-15 LK - http://QT8BH6HW4W.search.serialssolutions.com?sid=EMBASE&issn=14712377&id=doi:10.1186%2F1471-2377-13-15&atitle=V-TIME%3A+A+treadmill+training+program+augmented+by+virtual+reality+to+decrease+fall+risk+in+older+adults%3A+Study+design+of+a+randomized+controlled+trial&stitle=BMC+Neurol.&title=BMC+Neurology&volume=13&issue=&spage=&epage=&aulast=Mirelman&aufirst=Anat&auinit=A.&aufull=Mirelman+A.&coden=BNMEC&isbn=&pages=-&date=2013&auinit1=A&auinitm=

Mondellini, M., Arlati, S., Gapeyeva, H., Lees, K., Märitz, I., Pizzagalli, S.L., Otto, T., Sacco, M., Teder-Braschinsky, A., 2022. User Experience during an Immersive Virtual Reality-Based Cognitive Task: A Comparison between Estonian and Italian Older Adults with MCI. Sensors 22. https://doi.org/10.3390/s22218249

Muhla, F., Duclos, K., Clanché, F., Meyer, P., Maïaux, S., Colnat-Coulbois, S., Gauchard, G.C., 2022. Does the Management of Visual and Audible Motion Information during an Immersive Virtual Reality Timed Up and Go Test Impact Locomotor Performance in the Elderly? Gerontology 68, 456–464. https://doi.org/10.1159/000517286

Nielsen, J., 1994. Usability engineering. Morgan Kaufmann.

Nimmo, R., Constantinides, M., Zhou, K., Quercia, D., Stumpf, S., 2024. User Characteristics in Explainable AI: The Rabbit Hole of Personalization?, in:

Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. pp. 1–13.

Norman, D.A., Draper, S.W., 1986. User centered system design; new perspectives on human-computer interaction. L. Erlbaum Associates Inc.

Noroozi, F., Marjanovic, M., Njegus, A., Escalera, S., Anbarjafari, G., 2017. Audio-visual emotion recognition in video clips. IEEE Trans Affect Comput 10, 60–75.

Nourbakhsh, N., Chen, F., Wang, Y., Calvo, R.A., 2017. Detecting users' cognitive load by galvanic skin response with affective interference. ACM Transactions on Interactive Intelligent Systems (TiiS) 7, 1–20.

Ofosu-Asare, Y., 2024. Enhancing user experience in online learning environments: Design, evaluation, and usability techniques. Journal of Graphic Engineering and Design 15, 45–59.

Oncioiu, I., Priescu, I., 2022. The use of virtual reality in tourism destinations as a tool to develop tourist behavior perspective. Sustainability 14, 4191.

Pardini, S., Gabrielli, S., Dianti, M., Novara, C., Zucco, G.M., Mich, O., Forti, S., 2022. The Role of Personalization in the User Experience, Preferences and Engagement with Virtual Reality Environments for Relaxation. Int J Environ Res Public Health 19. https://doi.org/10.3390/ijerph19127237

Parsons, T.D., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., Buckwalter, J.G., Rizzo, A.A., 2004. Sex differences in mental rotation and spatial rotation in a virtual environment. Neuropsychologia 42, 555–562. https://doi.org/10.1016/j.neuropsychologia.2003.08.014

Pedroli, E., Greci, L., Colombo, D., Serino, S., Cipresso, P., Arlati, S., Mondellini, M., Boilini, L., Giussani, V., Goulene, K., Agostoni, M., Sacco, M., Stramba-Badiale, M., Riva, G., Gaggioli, A., 2018. Characteristics, usability, and users experience of a system combining cognitive and physical therapy in a virtual environment: Positive bike. Sensors (Switzerland) 18. https://doi.org/10.3390/s18072343

Peng, G.C., Annansingh, F., 2015. Experiences in applying mixed-methods approach in information systems research, in: Research Methods: Concepts, Methodologies, Tools, and Applications. IGI Global Scientific Publishing, pp. 910–936.

Peters, M.D.J., Godfrey, C.M., Khalil, H., McInerney, P., Parker, D., Soares, C.B., 2015. Guidance for conducting systematic scoping reviews. JBI Evid Implement 13.

Pieri, L., Tosi, G., Romano, D., 2023. Virtual reality technology in neuropsychological testing: A systematic review. J Neuropsychol 00, 1–18. https://doi.org/10.1111/JNP.12304

Pohlmeyer, A.E., Hecht, M., Blessing, L., 2009. User experience lifecycle model ContinUE [continuous user experience]. Der Mensch im Mittepunkt technischer Systeme. Fortschritt-Berichte VDI Reihe 22, 314–317.

Polit, D., Beck, C., 2006. The Content Validity Index: Are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 29, 489–497. https://doi.org/10.1002/nur.20147

Professor, M., Khalid, H., 2006. Handbook of Human Factors and Ergonomics, Third Edition. pp. 543–572. https://doi.org/10.1002/0470048204.ch21

Rajamurugu, N., Yaknesh, S., 2025. Extended Reality: A Conceptual Description. Introduction to Extended Reality (XR) Technologies 1–16.

Revelle, W., Zinbarg, R.E., 2009. Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. Psychometrika 74, 145–154.

Riva, G., 2018. The neuroscience of body memory: From the self through the space to the others. Cortex 104, 241–260.

Riva, G., Mancuso, V., Cavedoni, S., Stramba-Badiale, C., 2020. Virtual reality in neurorehabilitation: a review of its effects on multiple cognitive domains. Expert Rev Med Devices 17, 1035–1061. https://doi.org/10.1080/17434440.2020.1825939

Riva, G., Mantovani, F., 2014. Extending the Self through the Tools and the Others: a General Framework for Presence and Social Presence in Mediated Interactions, in:

Interacting with Presence: HCI and the Sense of Presence in Computer-Mediated Environments. pp. 9–31. https://doi.org/10.2478/9783110409697.1

Riva, G., Wiederhold, B.K., Mantovani, F., 2019. Neuroscience of Virtual Reality: From Virtual Exposure to Embodied Medicine. Cyberpsychol Behav Soc Netw 22, 82–96. https://doi.org/10.1089/cyber.2017.29099.gri

Rojo, A., Castrillo, A., López, C., Perea, L., Alnajjar, F., Moreno, J.C., Raya, R., 2023. PedaleoVR: Usability study of a virtual reality application for cycling exercise in patients with lower limb disorders and elderly people. PLoS One 18, 1–20. https://doi.org/10.1371/journal.pone.0280743

Rubin, D.B., 1976. Inference and missing data. Biometrika 63, 581–592.

Salimun, C., bin Abu Seman, E.A., Shishehchi, S., 2021. Predicting Aesthetic Preferences: Does the Big-Five Matters? International Journal of Advanced Computer Science and Applications 12.

Scherr, S.A., Elberzhager, F., Holl, K., 2017. An automated feedback-based approach to support mobile app development, in: 2017 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA). IEEE, pp. 44–51.

Schrepp, M., 2015. User experience questionnaire handbook. All you need to know to apply the UEQ successfully in your project 50–52.

Schrepp, M., Held, T., Laugwitz, B., 2006. The influence of hedonic quality on the attractiveness of user interfaces of business management software. Interact Comput 18, 1055–1069.

Schrepp, M., Hinderks, A., 2017. Design and evaluation of a short version of the user experience questionnaire (UEQ-S).

Schrepp, M., Hinderks, A., Thomaschewski, J., 2017. Construction of a benchmark for the user experience questionnaire (UEQ). International Journal of Interactive Multimedia and Artificial Intelligence 4, 40–44.

Schubert, T.W., 2003. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. Z. für Medienpsychologie 15, 69–71.

Schulze, M., Kollmorgen, J., Meiners, A.L., Hinderks, A., Winter, D., Santoso, H.B., Thomaschewski, J., 2023. On the Importance of UX Quality Aspects for Different Product Categories. International Journal of Interactive Multimedia and Artificial Intelligence 8, 232–246. https://doi.org/10.9781/ijimai.2023.03.001

Serino, S., Pedroli, E., Tuena, C., De Leo, G., Stramba-Badiale, M., Goulene, K., Mariotti, N.G., Riva, G., 2017. A novel virtual reality-based training protocol for the enhancement of the “mental frame syncing” in individuals with Alzheimer’s disease: A development-of-concept trial. Front Aging Neurosci 9, 1–12. https://doi.org/10.3389/fnagi.2017.00240

Shi, J., Mo, X., Sun, Z., 2012. Content validity index in scale development. Zhong Nan Da Xue Xue Bao Yi Xue Ban 37, 152–155. https://doi.org/10.3969/j.issn.1672-7347.2012.02.007

Siddiqi, M.H., Alam, M.G.R., Hong, C.S., Khan, A.M., Choo, H., 2016. A novel maximum entropy markov model for human facial expression recognition. PLoS One 11, e0162702.

Slater, M., 2018. Immersion and the illusion of presence in virtual reality. British Journal of Psychology. https://doi.org/10.1111/bjop.12305

Slater, M., 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos Trans R Soc Lond B Biol Sci 364, 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Strada, C., Brivio, E., Galimberti, C., 2013. 8. Digital education for older generation non native users: a focused ethnography study of a pilot project. QWERTY-Interdisciplinary Journal of Technology, Culture and Education 8, 44–57.

Sutherland, I.E., 1968. A head-mounted three dimensional display, in: Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I. pp. 757–764.

Tang, Y.M., Chau, K.Y., Kwok, A.P.K., Zhu, T., Ma, X., 2022. A systematic review of immersive technology applications for medical practice and education - Trends, application areas, recipients, teaching contents, evaluation methods, and performance. Educ Res Rev 35, 100429. https://doi.org/https://doi.org/10.1016/j.edurev.2021.100429

Taylor, F.W., 1911. The principles of scientific management. NuVision Publications, LLC.

Tcha-tokey, K., Christmann, O., Loup-escande, E., Richir, S., 2016. Proposition and Validation of a Questionnaire to Measure the User Experience in Immersive Virtual Environments Science Arts & Métiers (SAM) is an open access repository that collects the work of Arts et Métiers ParisTech. https://doi.org/10.20870/IJVR.2016.16.1.2880

The jamovi project, 2024. The jamovi project (2024).

Tölgyesi, B., Bakk, Á.K., Barkóczi, M., Buri, B., Szabó, A., Tobai, B., Sadie, B., Cserjési, R., Georgieva, I., Roth, C., 2023. “Virtual reality nature as our next retreat?”: User experience testing of a simulated natural environment in virtual reality. ACM International Conference Proceeding Series 448–453. https://doi.org/10.1145/3603555.3608543

Triberti, S., Brivio, E., 2017. User experience: psicologia degli oggetti, degli utenti e dei contesti d’uso. Maggioli Editore.

Triberti, S., Riva, G., 2016. Being present in action: a theoretical model about the “interlocking” between intentions and environmental affordances. Front Psychol 6, 2052.

Triberti, S., Sapone, C., Riva, G., 2025. Being there but where? Sense of presence theory for virtual reality applications. Humanit Soc Sci Commun 12, 79. https://doi.org/10.1057/s41599-025-04380-3

Tricco, A.C., Lillie, E., Zarin, W., O’Brien, K.K., Colquhoun, H., Levac, D., Moher, D., Peters, M.D.J., Horsley, T., Weeks, L., Hempel, S., Akl, E.A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M.G., Garritty, C., Lewin, S., Godfrey, C.M., Macdonald, M.T., Langlois, E. V, Soares-Weiser, K., Moriarty, J., Clifford, T., Tunçalp, Ö., Straus, S.E., 2018. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 169, 467–473. https://doi.org/10.7326/M18-0850

Tuena, C., Pedroli, E., Trimarchi, P.D., Gallucci, A., Chiappini, M., Goulene, K., Gaggioli, A., Riva, G., Lattanzio, F., Giunco, F., Stramba-Badiale, M., 2020. Usability issues of clinical and research applications of virtual reality in older people: A systematic review. Front Hum Neurosci 14. https://doi.org/10.3389/fnhum.2020.00093

Turvey, C., Fortney, J., 2017. The Use of Telemedicine and Mobile Technology to Promote Population Health and Population Management for Psychiatric Disorders. Curr Psychiatry Rep 19. https://doi.org/10.1007/s11920-017-0844-0

Van der Heijden, H., 2004. User acceptance of hedonic information systems. MIS quarterly 695–704.

van Gemert-Pijnen, 2018. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. Routledge.

Venkatesh, V., Bala, H., 2008. Technology acceptance model 3 and a research agenda on interventions. Decision sciences 39, 273–315.

Venkatesh, V., Morris, M.G., 2000. Why don’t men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior. MIS quarterly 115–139.

Venkatesh, V., Thong, J.Y.L., Xu, X., 2012. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS quarterly 157–178.

Von Itzstein, G.S., Billinghurst, M., Smith, R.T., Thomas, B.H., 2024. Augmented reality entertainment: Taking gaming out of the box, in: Encyclopedia of Computer Graphics and Games. Springer, pp. 162–170.

Wang, Y.-J., Lin, Y.-H., 2021. Liquid crystal technology for vergence-accommodation conflicts in augmented reality and virtual reality systems: a review. Liq Cryst Rev 9, 35–64.

Wang, Z., Li, Y., An, J., Dong, W., Li, H., Ma, H., Wang, J., Wu, J., Jiang, T., Wang, G., 2022. Effects of Restorative Environment and Presence on Anxiety and Depression Based on Interactive Virtual Reality Scenarios. Int J Environ Res Public Health 19. https://doi.org/10.3390/ijerph19137878

West, S.G., Finch, J.F., Curran, P.J., 1995. Structural equation models with nonnormal variables: Problems and remedies.

White, M., 2022. Sample size in quantitative instrument validation studies: A systematic review of articles published in Scopus, 2021. Heliyon 8.

Wickens, C.D., Gordon, S.E., Liu, Y., Lee, J., 2004. An introduction to human factors engineering. Pearson Prentice Hall Upper Saddle River, NJ.

Wildenbos, G.A., Peute, L., Jaspers, M., 2018. Aging barriers influencing mobile health usability for older adults: a literature based framework (MOLD-US). Int J Med Inform 114, 66–75.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.O., Nordberg, A., Bäckman, L., Albert, M., Almkvist, O., Arai, H., Basun, H., Blennow, K., De Leon, M., Decarli, C., Erkinjuntti, T., Giacobini, E., Graff, C., Hardy, J., Jack, C., Jorm, A., Ritchie, K., Van Duijn, C., Visser, P., Petersen, R.C., 2004. Mild cognitive impairment - Beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment. J Intern Med 256, 240–246. https://doi.org/10.1111/j.1365-2796.2004.01380.x

Winkler, N., Röthke, K., Siegfried, N., Benlian, A., 2020. Lose yourself in VR: exploring the effects of virtual reality on individuals' immersion.

Witmer, B.G., Singer, M.J., 1998. Measuring presence in virtual environments: A presence questionnaire. Presence: Teleoperators and Virtual Environments 7, 225–240. https://doi.org/10.1162/105474698565686

Wolf PhD, C., Jason, A., 2014. Defining patient experience. Patient Exp J 1, 7–19.

World Health Organization, 2021. Decade of healthy ageing: baseline report. World Health Organization.

Yi, M., Wang, Y., Tian, X., Xia, H., 2021. User experience of the mobile terminal customization system: the influence of interface design and educational background on personalized customization. Sensors 21, 2428.

Yin, C., Hsueh, Y.-H., Yeh, C.-Y., Lo, H.-C., Lan, Y.-T., 2016. A virtual reality-cycling training system for lower limb balance improvement. Biomed Res Int 2016.

Yoon, J., Kim, C., Kang, R., 2020. Positive user experience over product usage life cycle and the influence of demographic factors. International Journal of design 14, 85.

Zaman, B., Shrimpton-Smith, T., 2006. The FaceReader: Measuring instant fun of use, in: Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles. pp. 457–460.