Doctoral theses at NTNU, 2024:179

Asim Hameed

(In)Authentic VR

Quality Assessments of Interactivity in Virtual Reality

Doctoral thesis

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Information Technology and Electrical
Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

Asim Hameed

(In)Authentic VR

Quality Assessments of Interactivity in Virtual Reality

Thesis for the Degree of Philosophiae Doctor

Trondheim, May 2024

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering
Department of Electronic Systems

ISBN 978-82-326-7950-8 (printed ver.)
ISBN 978-82-326-7949-2 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2024:179

Printed by NTNU Grafisk senter

This thesis is submitted to the Norwegian University of Science and Technology (NTNU) for the fulfillment of requirements for the degree of Doctor of Philosophy.

The doctoral work started in March 2018 at the Department of Electronic Systems, NTNU, Trondheim, Norway. The work has been supervised by Professor Andrew Perkis and co-supervised by Professor Sebastian Möller.

The members of the assessment committee are as follows: Professor Mel Slater, University of Barcelona, Spain; Professor Rigmor Baraas, University of South-Eastern Norway, Kongsberg Norway; Professor Ceenu George, Technical University Berlin, Germany; and Associate Professor Ana Sanchez Laws, Norwegian University of Science & Technology, Norway.

The doctoral work was funded by the IE Faculty.

Dedicated to Momo, Baba, and Amajee

One is still what one is going to cease to be
and already what one is going to become.

– Jean-Paul Sartre

Preface

Immersive technologies like virtual reality (VR) rapidly transform how we experience media and interact with information. As these emerging mediums reshape notions of space, time, and reality, there is a pressing need to develop rigorous frameworks for understanding the multidimensional facets that collectively define their experiential qualities. This thesis offers timely theoretical explications and empirical insights to advance scholarly comprehension of immersive media experiences (IMEx) in their human-centric complexity.

The overarching motivation stems from recognizing that while technological capabilities are integral, they tell only part of the story. To holistically evaluate IMEx, cross-disciplinary perspectives that consider the confluence of system parameters with cognitive, perceptual, and behavioral processes must be adopted. Through conceptual consolidations coupled with systematic mixed-methods studies, the research presented here expands the toolkit for assessing user-cent kkered facets that shape experiential quality. A seminal contribution lies in formulating an overarching taxonomy delineating key aspects, elements, and features that characterize quality within immersive media. This thesis proposes authenticity as a complementary concept to presence and draws attention to their conceptual correlation within a user-centric framework. It is a reflective appraisal of coherence and congruence in their significance towards a virtual experience. A comprehensive taxonomy is then introduced to untangle the complex, interconnected aspects influencing VR experiences and to capture their multidimensional nature for a more systematic assessment.

Further, adopting observational techniques from human behavioral research, it investigates overt physical manifestations of perceived affordances and actions within VR environments. Comparisons with self-reported measurements reveal complex interplays between technology, cognition, and behavior. Shifting to input modalities, the thesis empirically examines oft-held assumptions about interaction naturalness through comparative assessments of hand-tracking and controllers for motor tasks. Findings expose subtle divergences between technological capabilities and user performance that warrant greater prudence in equating fidelity with quality.

A notable emphasis is highlighting the need to look beyond technological capabilities when evaluating immersive experiences. While engineering advancements in display resolution, field of view, and tracking accuracy are indispensable, the thesis cautions against equating fidelity with quality. It advocates a perspective that weighs the fidelity of technology against the fidelity of experience. The latter includes intricate cognitive and emotional processes associated with concepts such as presence, flow, and cognitive absorption that require further explanation. In this pursuit, the research applies cross-disciplinary tools like behavioral observation, affordance taxonomy, and pragmatic-hedonic modeling that uncover relationships and gaps not discernible through system parameters alone. It expands the methodological repertoire for investigating subjective and latent facets of user experience.

Collectively, this thesis expands and enriches the understanding of IMEx by integrating cross-disciplinary perspectives. The empirical findings offer original evidence, highlighting system potential and experience gaps. In illuminating relationships between technology and behavior, the thesis advocates complementing engineering advancements with deeper insights into human factors for optimizing fully immersive VR experiences. Overall, it provides timely contributions toward comprehensive quality assessment frameworks for emerging mediums that reshape perceptions of how we interact with data, tell stories, and socialize.

Acknowledgments

As this chapter concludes, I find myself humbled by the remarkable experiences and companionships that came to define it. I am grateful to so many people whose kind support and encouragement brought a lightness to this endeavor.

First and foremost, I express my deepest gratitude to my supervisors. To Professor Andrew Perkis, Department of Electronic Systems (NTNU), for his unwavering support and mentorship throughout my doctoral work. To Professor Sebastian Möller, Quality & Usability Labs (TU Berlin), for his dedication and meticulous feedback that helped shape this work.

I shall fondly remember the Signal Processing Group for the lively kitchen atmosphere and even livelier wine tastings. I value the support and assistance I received at the Department of Electronic Systems during my years at IES. I want to thank everyone who kept the coffee machines running and the fruit baskets fresh. I also appreciate the team at the Quality & Usability Labs for their help and facilitation during my research residency at TU Berlin. I shall cherish my time in Berlin for years to come.

To Stefan Uhrig, Shafaq Irshad, Øyvind Klungre, and others – for your thoughtful contributions to our discussions. I appreciate you sharing your insights and perspectives, which have added much-needed value to my work.

A shoutout to the rad group I met along the way and got to know as friends. You all made this time memorable. Your presence is enduring, and your impact is profound. To Ashkan, Johanne, Francois, and Victor – *You Is Kind, You Is Smart, You Is Important*.

My special gratitude goes to Trondheim, which has become home.

But most of all, I am indebted to the love and kindness of my family. To my parents for enabling me to write and be a part of my story. To my wife for believing in me, tolerating me, and reminding me that a cup of tea can fix everything. To my sisters, who've been my pillars of strength, my OG cheerleaders, always ready to offer a listening ear and some much-needed humor. To my Peppa Pig, Baby George, and Baby Alexander – you guys light me up!

List of Publications

This dissertation is based on the research studies attributed below. It was conducted in accordance with the research objectives defined in Section 1.1. There are 8 total papers, referred to in the text by their Roman numerals, that cover the contributions mentioned in Section 1.2. Seven papers have already been published, and one is under review.

I. Hameed, Asim, Sebastian Moeller, and Andrew Perkis. "A Holistic Quality Taxonomy for VR Experiences." in *Frontiers in Virtual Reality*, 2024: (in review).

II. Hameed, Asim, and Andrew Perkis. "Authenticity & Presence: Assessing the Quality of VR Experiences." in *Frontiers in Psychology*, 2024, 15: 1291650.

III. Hameed, Asim, and Andrew Perkis. "Spatial storytelling: Finding interdisciplinary immersion." in *11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5–8, 2018, Proceedings 11,* pp. 323-332. Springer International Publishing, 2018.

IV. IV Hameed, Asim, Shafaq Irshad, and Andrew Perkis. "Towards a quality framework for immersive media experiences: a holistic approach." in *12th International Conference on Interactive Digital Storytelling, ICIDS 2019, Little Cottonwood Canyon, UT, USA, November 19–22, 2019, Proceedings 12,* pp. 389-394. Springer International Publishing, 2019.

V. Hameed, Asim, and Andrew Perkis. "Affects of Perceived-Actions within Virtual Environments on User Behavior on the Outside." in *12th International Conference on Quality of Multimedia Experience (QoMEX),* pp. 1-6. IEEE, 2020.

VI. Hameed, Asim, and Andrew Perkis. "A Subjective and Behavioral Assessment of Affordances in Virtual Architectural Walkthroughs." in *Applied Sciences* 11, no. 17 (2021): 7846.

VII. Hameed, Asim, Andrew Perkis, and Sebastian Möller. "Evaluating hand-tracking interaction for performing motor-tasks in vr learning environments." in *13th International Conference on Quality of Multimedia Experience (QoMEX),* pp. 219-224. IEEE, 2021.

VIII. Hameed, Asim, Andrew Perkis, and Sebastian Möller. "How Good Are Virtual Hands? Influences of Input Modality on Motor Tasks in Virtual Reality." in *Journal of Environmental Psychology* 92 (2023): 102137.

List of Figures

List of Tables

Contents

Abbreviations

Abbreviations

ANOVA	Analysis of Variance
AR	Augmented Reality
AT	Active-Time
ATT	Attractiveness
AV	Augmented Virtuality
BORIS	Behavioral Observation Research Inter- active Software
CGI	Computer-Generated Imagery
DoF	Degrees of Freedom
DV	Dependent Variables
EN	Engagement
HMD	Head-Mounted Displays
HQ	Hedonic Quality
IMEx	Immersive Media Experiences
IMT	Immersive Media Content
IMT	Immersive Media Forms

IMT	Immersive Media Technologies
INV	Involvement
IPQ	Igroup Presence Questionnaire
ITC-SOPI	Independent Television Commission-Sense of Presence Inventory
ITU	International Telecommunication Union
ITU-T	ITU - Telecommunication Standardization Sector
IV	Independent Variables
IVE	Immersive Virtual Environment
IW	Interactive Walkthrough
IW	Passive Walkthrough
LC	Left-Click Frequency
MANCOVA	Multivariate Analysis of Covariance
MANOVA	Multivariate Analysis of Variance
MOS	Mean Opinion Score
MR	Mixed Reality
NASA-TLX	NASA Task Load Index
NE	Negative Effects
NV	Ecological Validity
OP	Object-Pick
PBR	Physically Based Rendering
PQ	Pragmatic Quality
QoE	Quality of Experience
QoS	Quality of Service
RC	Right-Click Frequency
REAL	Realness

RSME	Rating Scale Mental Effort
SP	Spatial Presence
VE	Virtual Environment
VR	Virtual Reality
XR	Extended Reality

Chapter 1

Introduction

Immersive media are not new. They are simply a part of an evolutionary push towards media becoming ever more immersive and exploring different ways of conveying the human experience. A podcast, for example, develops a sense of connection for the listener due to the medium's intimacy. Fully immersive media, like Virtual Reality (VR), transported our interactions with a medium from the confines of a rectangular screen to a 360-degree space. It is all around. This illusion of surroundness is a distinctive immersive media experience (IMEx) of VR, using physical actions as inputs to receive outputs in a perceived 3D space [6], dynamically altering the user experience inside mediated digital realities where data is spatial. From its early stages of bulky equipment tethered to one place, innovative solutions are making them lighter and creating new, fully immersive environments outside the limitations of a single room, with increasing accessibility for all. Our understanding of this new form of media must acknowledge that we can not reduce such mediated experiences in terms of their constituent technologies alone. In that, their psychological and phenomenological dimension are equally, if not more, important. IMEx, particularly VR experiences, are not the presence or absence of any requisite technology but the very particular psychological experiences they invoke in the user. Similarly, assessments of experiential quality within immersive media like VR are markedly different than the user-perceived assessments of traditional telecommunication media [7]. Primarily because the spatial sense of "being inside" or "being there" is not only a unique new aspect of telecommunication media altogether but, in fact, foundational to assessing such new multimedia experiences.

The premise above sits at the heart of this research endeavor. This dissertation delves into assessments of immersive new media, explicitly emphasizing quality assessments of VR Experiences. The challenge lies in comprehending the intricate

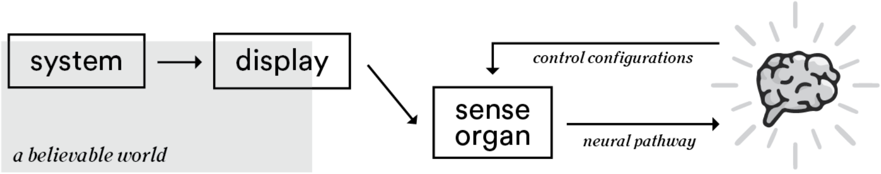


Figure 1.1: An immersive system simulates a reality that suspends our disbelief and convinces us into believing that the virtual world may be real, adapted from [1].

factors that contribute to a truly immersive experience in a field known for its preoccupation with technology. While previous research has made substantial strides in understanding specific facets of VR, this dissertation posits that a comprehensive evaluation demands an integrated perspective anchored at the intersection of Quality of Experience (QoE) and User Experience (UX) research within the context of Virtual Reality (VR) technology. Unlike traditional approaches that often separate these domains, our work recognizes the interconnectedness of technological attributes, psychological states, and experiential qualities in evaluating VR. It draws upon established theories in human-computer interaction, cognitive psychology, and media studies while integrating novel concepts to address the complex and multifaceted nature of immersive VR experiences. By bridging the gap, we aim to present a more nuanced understanding of the immersive qualities that define VR experiences.

A fundamental contribution of this dissertation is developing a holistic framework across disciplinary boundaries to assess the multifaceted dimensions of VR experiences systematically. A taxonomy is presented by disentangling essential quality aspects, which may enable a more nuanced examination of the interdisciplinary factors shaping VR encounters. Moreover, this work challenges the prevailing notion of presence as the sole psychological aspect of relevance in VR. At its core, IMEx relies on the subjective sense of being physically present within a virtual environment (VE). These environments could range from full-blown virtual realities (VR) to mixed realities (MR). In the context of fully immersive VR, real-world stimuli are replaced with computer-generated inputs to deceive the brain into perceiving the virtual world as reality – place illusion [8] (see Fig1.1). The possibility of such experiences requires an artful integration of various stimuli, manipulating audiovisual and sensory feedback, and defying a user's rational knowledge and prior beliefs to create a compelling mental model, which the user accepts as a reality [9] [10] [11] [12]. However, this work looks beyond the mere generation and manipulation of stimuli, extending the focus to its credibility instead [13]. This dissertation introduces the concept of authenticity as a pivotal dimension influencing the perceived quality of virtual worlds. The investigation into authenticity

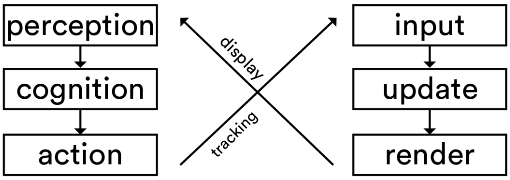


Figure 1.2: Action is the explicit response to the possibility for action presented by the system within the VE, adapted from [2].

explores users' conscious evaluations of how credible a virtual world appears, considering alignment with internal logic and conformance to user expectations. In other words, the coherent organization of stimuli (synchronicity or congruence), signs, and markers within the mediated environment enforces authenticity and elicits realistic responses from the user [14].

One characteristic of authenticity is how action possibilities, or *perceived actions*, play out within a given virtual world. The potential for action influences how a user may interact and engage with the VE and is crucial for curating highly immersive and intuitive VR experiences [15] [16] [17]. Such possibilities emerge from within the environment and are accentuated if congruity is found between the VE's behavior and the user's preexisting knowledge (Fig1.2). Users instinctively respond to the perceived actions that encompass the possibilities for exploration, interaction, and achieving goals. A meaningful VR experience materializes when the user's expectations, attitudes, and attention harmoniously converge with the VR encounter while simultaneously fueling their imagination to complete the immersive journey [18]. Secondly, advances in interactivity have paved the way for instinctual actions within extensive virtual settings. Users are no longer just observers but active participants, responding to various visual, auditory, and tactile stimuli [19][20]. This potential has been realized not only for entertainment but has also found practical applications in various fields, ranging from simulation-based training to rehabilitation and manufacturing [21] [22] [23] [24]. Perceived action and its embodiment in user movements can be nuanced due to various human- and system-related factors [25] [26]. Movements within the VR space are explicit responses to action possibilities and echo the respective intentions of the users. A VR system facilitating movement allows users to express themselves through deliberate and purposeful interactions and movements that may align with their intentions. This fosters authenticity in the simulated world by finding congruence between their intended actions and the behavior of the world [27] [28].

This dissertation further extends its contributions through a series of empirical studies to assess the influence of action possibilities on the perceived quality of VR experiences. Understanding the impact of perceived actions requires holistic methodologies that are not innately system-centric (focusing more on the technology than on the user) but shift focus to a more human-centric paradigm. These investigations, spanning user behavior in virtual and physical environments to the comparative analysis of input modalities like hand-tracking, employ naturalistic observational techniques and mixed-methods approaches. The exploration of VR interaction quality, user expectations, and the influence of input modalities enriches our understanding of experiential quality within immersive environments.

1.1 Scope and Objectives

The main focus of this research encompasses the technical, experiential, and perceptual aspects of VR. It recognizes that quality assessments in VR are multifaceted and outlines various factors and facets that influence them. More specifically, this study examines the potential implications of mismatches between users' expectations and actions supported by the VR system. The research seeks to elucidate how such disparities may impact a user's sense of authenticity and presence and, consequently, the overall quality of the immersive experience within VR. The scope of the research is not limited to the use of state-of-the-art techniques and technologies but also the psychological and perceptual dimensions they engage, inviting us to question: (i) How do we determine the authenticity of VR experiences? (ii) How do coherence and congruence issues impact VR interactions? (iii) How do mismatches affect quality assessments in VR about various psychological, emotional, and performance factors? These questions lead us to the following research objectives on which this thesis focuses:

O1: To explicate upon quality aspects of IMEx, in specific VR, emphasizing a multifaceted and holistic approach to quality assessments.

O2: To identify and attempt new QoE methodologies to understand human factors inside VR.

O3: To perform QoE assessments on the comparability of VR interaction quality, action possibilities of the VE, and the user's expectation.

1.2 Contributions

The central contribution of this thesis lies in its development of a comprehensive theoretical framework, subsequently validated through a series of selected experiments. The thesis initiates by formulating an overarching quality taxonomy and proceeds to delve into a more specific exploration of the perception and response to action possibilities within VEs. This is followed by identifying methodologies suited to the quality assessment of such VR interactions. The contributions can be briefly summarized as follows:

(C1) The first contribution of this thesis applies to our understanding of VR-based IMEx, which is covered in publications **I**, **III** and **IV**. This contribution addressed the research concern identified in **O1**. In **III** and **IV**, we determine the need for a holistic QoE framework aimed at comprehensively assessing and understanding the quality of user experience in VR environments considering the interconnectedness, interdependencies, and interactions of various components that make up immersive media and factors that influence immersive media, rather than focusing on system attributes in isolation. In short, the quality of the user's experience remains a complex and multidimensional concern. In response to this challenge, we devise a comprehensive taxonomy in **I** to help dissect and understand multiple aspects influencing VR experiences.

(C2) The second contribution of this thesis is the Presence-Authenticity Dyad, in response to **O1**. In **II**, we build upon existing scholarship to define authenticity as a critical dimension of quality perception complementary to the feeling of presence. We refer to the credibility of the virtual world, its alignment with internal logic, and its conformance to user expectations. Authenticity in VR is a user's conscious evaluation of how credible a virtual world appears. Notwithstanding nomenclature, we identify authenticity as experienced quality involving higher-order assessment of the VR experience. It's not just the initial perception of the virtual world but also about deliberate reflection and evaluation of that perception.

(C3) The third contribution of this thesis addresses research objectives **O2** and **O3**. We introduce an observational methodology in **V** to assess if perceived action inside a VE translated into overt behavior of the user on the outside. We apply quantitative behavioral observation to analyze user behavior, identifying the potential for using cross-disciplinary tools for QoE assessments. In **VI**, we expand our understanding of perceived action under the rubric of affordances and apply subjective measures to cross-examine user behavior described in **V**. In another empirical study, available in **VII** and **VIII**, we build upon **O2** and **O3** by evaluating whether the devices and/or interfaces impact user responses to perceived actions in a simple motor task. In this case, performance metrics are used (**VII**) in combination with subjective and usability measures (**VIII**) to assess the overall QoE.

1.3 Theoretical Framework

As VR technologies advance, it becomes imperative to assess the various complex factors that contribute to them. Towards this end, this work presents an interdisciplinary framework. It conducts empirical investigations of critical technological and human-centric elements, amongst others, that shape user perceptions of presence, authenticity, and overall quality in VR contexts. A holistic understanding of VR experiences must include a user's subjective, emotional, and overall experience with the VR, going beyond quality perceptions. A five-dimensional taxonomy is presented to disentangle and redefine often-conflated concepts. Below is a brief summation of the detailed account that follows in **Chapter 2**.

- ***Immersivity*** captures the subjective sense of being transported into a VE, characterized by a feeling of "being there." and made possible by various system aspects.
- ***Interactivity*** refers to the ability of users to interact with the VE and influence their experience, resulting in a sense of control and engagement.
- ***Explorability*** captures the ease and degree of freedom users can navigate and discover new elements within the VE.
- ***Plausibility*** refers to the congruence of the VE, its rules and interactions, aligning with user expectations and cognitive models.
- ***Believability*** refers to the extent to which the VE appears realistic and coherent, giving it a sense of authenticity and acceptance.

The theoretical underpinnings for this dissertation rest on the understanding that the quality of a VR experience necessitates an integrated approach that includes learnings from across disciplines.

Presence Theory: This theory, as exemplified by Sheridan's six degrees of presence [29] and Lee's later explications [30], focuses on the user's subjective sense of being physically present within the virtual environment. This dissertation expands the concept by introducing "authenticity" as a complementary dimension, recognizing users' critical assessment of the virtual world's credibility and alignment with expectations.

Media Immersion Theory: Works like Brenda Laurel's "Computers as Theatre," [31] explains how immersive media can create powerful experiences by engaging users on multiple sensory and cognitive levels. It extends the understanding of how VR leverages immersion to transport users into virtual worlds not just on a sensory level but equally via narrative and dramaturgical elements, enhancing the overall immersive experience.

User Experience Assessments: Quality of Experience (QoE) and User Experience (UX) are both related concepts that seek to understand and improve the user's experience. However, they differ significantly in their origins, theoretical foundations, and evaluation approaches [32]. With its roots in the telecommunications industry and technology, QoE was developed to evaluate technical system performance and quality features. They often use quantitative metrics such as Mean Opinion Score (MOS). In contrast, UX has its basis in human-computer interaction (HCI) and is known for its multidisciplinary approaches. It draws on psychological research to prioritize understanding human needs, emotions, and subjective experiences. The theoretical foundations of both fields are also reflective of their differences. UX benefits from well-established theories across various disciplines, while QoE's theoretical framework is gradually evolving to incorporate perspectives beyond technical quality assessment. The differences between QoE and UX are most pronounced in their evaluation methods. QoE relies heavily on standardized quantitative scales such as MOS, which are conducted in controlled lab settings to measure perceived quality of service and performance factors. Meanwhile, UX employs various quantitative and qualitative methods from multiple disciplines to capture the nuanced, contextual, and subjective aspects of the user experience in real-world settings. However, both fields recognize the importance of individual differences and the context in which experiences occur. As QoE research evolves to adopt more experiential and UX-inspired evaluation approaches, the lines between QoE and UX may become less distinct, leading to a more comprehensive understanding of user experiences that encompasses both technical quality and human-centric perspectives.

This thesis operates at the intersection of these established theories, creating a novel framework that considers

- **Technological aspects:** How hardware, software, and network capabilities provide the foundation for VR experiences.

- **Psychological factors:** How users perceive, interact with, and respond to the VE, encompassing presence, authenticity, and other cognitive processes.

- **Experiential qualities:** How VR evokes emotions, enjoyment, and meaning-making for users, including both pragmatic and hedonic aspects.

1.3.1 Situating the Research

As outlined, this work is situated within broader scholarly conversations about evaluating and optimizing VR experiences. The research is grounded in theoretical perspectives that view technology as extending fundamental human experiences and capabilities. The conceptual framework draws upon phenomenological philosophies, embodied cognition theories, and ecological psychology approaches that construe technologies as mediating our lived experiences and relationships with the world. In this view, fully realizing VR's potential requires recognizing it as more than just a tool but an experiential medium affording extended sensory, cognitive, and behavioral possibilities.

This integrative approach is reflected in the synthesized quality framework suggested in this dissertation, which expands the research agenda from the purely technical, predominantly positivist view in the field that emphasizes objective aspects above all else. This dichotomous departure recognizes that while the positivist view provides a solid foundation for evaluations in VR, it's essential to acknowledge that not all aspects of VR experiences may be easily quantifiable. This dissertation highlights the profound interrelation between technologies and human experiences. Notably, the work includes Norman's theory of *affordances* based on Gibson's ideas of information pick-up to theorize interactivity within VR [33] [34]. Supporting this conceptual framework, authenticity is explored as a subjective credibility assessment beyond just presence, adding nuance to understanding users' conscious, reflective evaluations of VR environments. This conceptual expansion views such assessments in terms of the internal logic of the virtual world and its alignment with user expectations, which resonates in Heidegger's notions of *'ready-to-hand* and *'present-at-hand* technologies [35]. The perspective posits VR as facilitating experiential possibilities beyond traditional psychological dimensions that reconfigure our embodied cognition, behaviors, and impressions. Separately, *storiness* is included in the conversation for its influence on shaping experiences. MR Ryan [36] referred to VR as a metaphor while exploring the relationship between traditional literary narratives and the new genres made possible by interactive media like VR. The role of narrative immersion is regularly included in evaluations of interactive storytelling [37] [38]. This work recognizes the importance of understanding the phenomenological experience of engaging with different narrative and dramaturgical elements of a VR experience as dietetic assessments of its credibility.

1.4 Critical Review & Synthesis

Extensive research was conducted across multiple relevant disciplines to develop a holistic framework with a comprehensive taxonomy capturing the key dimensions shaping quality VR experiences. Source materials from human-computer interaction, virtual environment design, game studies, media psychology, and communication/media studies were systematically reviewed. The review identified recurrent themes and salient concepts related to factors influencing VR-based user experiences. These diverse factors were synthesized into higher-level and low-order dimensions. This analytically derived taxonomy aimed to disentangle the core aspects contributing to immersive, interactive, believable, explorable, and plausible virtual environments.

Particular emphasis was placed on analyzing literature examining psychological processes and subjective states induced by VR systems. Prior research on presence, engagement, enjoyment, and perceptions of realism were closely examined. Analyzing existing scholarship revealed authenticity as an underexplored yet potentially pivotal dimension complementary to presence in shaping VR quality of experience. Perspectives across multiple philosophical traditions on the nature of authentic experiences and human credibility assessments were studied to establish authenticity as a critical facet. Theories from phenomenology, psychology, and aesthetics were evaluated to conceptualize authenticity grounded in how users perceive, interpret, and evaluate VR based on the compliance of internal schemas and expectations.

1.5 Methodology

Below, are details of the selected methods & measures for VR Assessments and their alignment with the research goals and philosophical underpinnings.

1.5.1 Empirical Study

An empirical methodology involves systematic data collection and analysis to understand users' subjective experiences and perceptions of the quality of a specific product, service, or system. This approach relies on empirical evidence gathered through user studies, surveys, experiments, or other research methods [39].

Our empirical QoE studies employed subjective rating scales, performance metrics, and behavioral measures to gather empirical data on users' experiences and perceptions. These methods capture both objective and subjective aspects of QoE, including usability, satisfaction, perceived quality, and emotional responses [40].

Users engaged with the system under evaluation, providing feedback and completing surveys after specific tasks or scenarios. The collected data was then subjected to quantitative analysis. Repeated measures analyses were used to identify relationships, trends, and statistical significance for factors contributing to positive or negative experiences.

1.5.2 Observational Methodology

We applied an observational methodology for assessing QoE. It involves systematic observation and recording of user behaviors, interactions, and expressions during their interaction with the system to understand their subjective experiences [41]). This approach captures real-time user data in natural settings, providing valuable insights into users' actual experiences [42].

Observations were conducted in appropriate laboratory settings of simulated scenarios. Researchers define specific behaviors, interactions, and expressions relevant to the assessed QoE, such as user engagement, frustration, satisfaction, or enjoyment. Data collection included video recordings, a structured coding system, and structured checklists to capture users' actions and non-verbal cues during interaction. These methods ensure systematic and unbiased recording of observational data. The collected data was analyzed to identify patterns and trends related to the observed behaviors and interactions. The findings from the data analysis are interpreted to conclude the QoE aspects being investigated, linking the observed behaviors and interactions to users' subjective experiences [43]. More details in **Chapter 3**.

1.5.3 Performance metrics

We used performance metrics for our study described in **Chapter 4**. In-game analytics are metrics encompassing quantitative measures utilized to assess and gauge the performance and effectiveness of a VR task [44]. These metrics capture a range of player-related aspects, including behavior, engagement, progression, and conducting user testing to measure performance indicators such as task completion time and error rates [45]. Such data serves not only to assess user performance but also to gain a comprehensive understanding of learning and training environments, facilitating optimization efforts [46].

1.5.4 Subjective Measures

Igroup Presence Questionnaire (IPQ)

A widely utilized psychological assessment tool designed to measure users' sense of presence in virtual environments [47]. It is comprised of a self-report questionnaire. The IPQ collects subjective data from users regarding their experience in a VE. We utilized the IPQ for our study discussed in **Chapter 4**. This is because our study was specifically interested in investigating the user's sense of presence, involvement in the task, and perception of realism – also the subscales of IPQ. Users respond to various questions on a Likert scale, indicating their level of agreement or disagreement. These statements cover diverse aspects of presence, such as feeling genuinely present, having a solid sense of being in the virtual environment, and perceiving the virtual environment as a plausible place to visit. The IPQ has found applications across multiple fields, including gaming, simulation, training, psychology, and human-computer interaction. Its validity and reliability have been established through numerous studies (Cronbach's $\alpha = 0.87$) [27] [47] [48], solidifying its status as a widely accepted tool for assessing presence in virtual environments.

ITC-Sense of Presence Inventory (ITC-SOPI)

The ITC-SOPI is a relatively new questionnaire designed to measure the subjective sense of presence experienced by individuals in various media contexts [49]. The questionnaire draws on previous research on presence determinants and existing self-report measures, focusing solely on users' experiences without considering objective system parameters. The measure identifies four primary factors: Sense of Physical Space, Engagement, Ecological Validity, and Negative Effects. The ITC-SOPI is reliable and valid, and its psychometric properties have been tested[50]. The ITC-SOPI was chosen for the study reported in **Chapter 3** because it caters to discomfort and adverse effects, which were a focus of our work.

NASA Task Load Index (NASA-TLX)

This is a well-established assessment tool that was utilized in our research mentioned in **Chapter 4** to measure perceived workload and evaluate the effectiveness of tasks and systems [51]. It has widespread adoption across diverse domains, showcasing its significant impact on human factors research [52]. As a subjective self-reporting measure, it relies on participants' workload evaluations. It provides multidimensional scores that capture various aspects of perceived workload: mental-, physical-, temporal-, effort-, performance-, and frustration Levels. NASA-TLX is not designed to measure a task objectively. It should be used with objective metrics that consider factors such as task completion speed and accuracy.

Rating Scale Mental Effort (RSME)

A lesser-known subjective uni-dimensional measure used for evaluating the perceived mental effort or cognitive load experienced by individuals during a task or activity. It finds common application in cognitive psychology, human factors research, and usability testing [53]. The RSME employs a numerical rating scale, ranging from 0 to 100, where participants are requested to rate the mental effort they invested while engaging in a specific task. This prompt rating was used inside VR to recorded participants' subjective scores. It is described in **Chapter 4**. Higher RSME ratings show more significant perceived mental effort, suggesting that the task necessitates more cognitive resources. Conversely, lower RSME ratings indicate a lower perceived mental effort and reduced cognitive demands.

AttrakDiff

This questionnaire is designed to evaluate the hedonic and pragmatic attributes of interactive products and systems [54]. It is a measurement tool for assessing interactive technology's overall user experience and attractiveness. It was used to evaluate two distinct input modalities in our study (see **Chapter 4**). The questionnaire comprises a series of paired adjectives representing different aspects of the user experience, including attractiveness, hedonic, and pragmatic qualities. Participants utilize these pairs to rate their subjective impressions of a product or system. The questionnaire aims to capture the emotional appeal and the practical utility of the interactive technology being evaluated. AttrakDiff has gained significant utilization in human-computer interaction and user experience research [55] [56]. It provides a standardized and reliable method for assessing users' perceptions of interactive products [57], facilitating comparisons between different designs and enabling their impact on user experience.

1.6 Limitations

Although mixed methods align with recent trends toward ecological validity and leverage new measurement capabilities, limitations provide useful lessons.

Task Specificity & Complexity: The specific set of actions and their complexity used in the studies here may not capture the full range of interactions users engage in within VEs. A narrow focus limits the applicability of the findings to broader VR experiences.

Device Dependency: Each study was limited in testing to a specific headset. The findings may not be generalizable to other VR systems or immersive technologies, restricting the external validity of the research. Comparing everyday consumer VR systems creates more ecological validity than specialized equipment but constrains the technological factors examined.

Ecological Validity: One of the studies may lack ecological validity as the VE might not fully replicate the complexities and nuances of the physical world. Participants may behave differently in virtual spaces compared to real-world scenarios, leading to limitations in generalizing the findings to real-life situations.

Control Group: The lack of a control group prevented a more apparent distinction between the impact of virtual actions and general environmental influences.

Sample Diversity: The study's generalizability may be limited if the sample is not diverse enough. If participants share similar backgrounds, preferences, or experiences, it might not adequately represent the broader population, affecting the study's external validity.

1.6.1 Methodological Reflections

The empirical work employs mixed methods and naturalistic observational techniques, integrating systematic behavioral codings, objective performance indicators, and subjective self-report instruments. These methodological choices contribute to the studies' overall ecological validity. Reflections on potential alternative methodologies could further enrich the research. For example:

Sample Size: Consider exploring larger sample sizes in future studies to enhance the statistical power and generalizability of findings. Although the stipulated ITU thresholds were met, increasing the number of participants could strengthen the conclusiveness of the results and lend them greater external validity.

Replication of Studies: Replicating the studies in different contexts and with diverse subject pools would have made the conclusions robust. Systematically replicating the findings across varying VR experiences and user demographics strengthens the overall confidence in the research and broadens its applicability.

Qualitative Analysis: Additional qualitative techniques like thematic analysis and grounded theory would have extracted even richer insights from subjective data and uncovered more nuanced patterns.

Cross-Device Comparison: Including multiple VR devices with different hand-tracking technologies could be used in future studies. Comparing findings across various devices will help identify whether the observed effects are device-specific or generalizable to other platforms.

Longitudinal Design: Longitudinal studies provide detailed insights into the long-term effects of virtual interactions. Observations over an extended period to identify lasting changes can be applied when time and budget permits.

1.7 Ethical Considerations

The experiments conducted for this user study involved human participants and the collection of potentially sensitive data, raising critical ethical considerations. Every effort was made to protect participants' rights and welfare at all stages of this study. Privacy and informed consent were prioritized, with the ethics review board oversight. All data collection, storage, and publication were handled to maximize confidentiality. No major ethical issues arose.

Ethical Oversight : All procedures performed for this research were by the ethical standards of the Institutional Review Board (or Ethics Committee) and the 1964 Helsinki Declaration and its later amendments. Permissions were sought prior to the commencement of each study. All experiment procedures and data collection standards adhered strictly to established guidelines described in ITU-T Recommendations P.809 [58], P.911 [59], and P.919 [60]. Any unanticipated ethical concerns arising during the study had to be reported to the ethics board.

Participant Recruitment and Consent: Participants were recruited voluntarily through word-of-mouth, email lists, pinboard postings, and via an online volunteer portal of the university open to the general public. People from diverse backgrounds signed up for the studies. The study's general purpose and essential procedures were explained during recruitment, but specific experimental details were withheld to avoid biased results. All participants were required to provide written informed consent before participation. Consent forms, experiment briefs, etc., followed the NTNU Sense-It templates used by all PhDs in the group and developed over many years of experimental research. The consent form outlined the study purpose, procedures, risks, benefits, compensation, data collection/use, confidentiality measures, and the voluntary nature of participation. Participants were made aware that they could withdraw at any time without penalty.

Participant Privacy: Protecting participant privacy has been paramount. All personal information and data collected were anonymized and assigned a random ID number. Data and video recordings were securely stored behind encrypted folders maintained by the university, accessible only to the research team. Published results will contain no identifying participant information. Information on research data management is followed in 1.8.1.

Potential Risks and Safeguards: There were no significant anticipated risks outside those encountered in everyday life. Mild fatigue or discomfort was possible during lengthy VR sessions. Participants were fully briefed on all procedures beforehand and allowed to withdraw if they felt uncomfortable or tired. The primary researcher personally monitored participants during sessions for any signs of distress. Sessions would be ended immediately if any participant wished to stop. The study was conducted during the second COVID-19 wave. Strict hygiene protocols were followed during test sessions based on the specified SARS-CoV-2 guidelines for educational institutions.

Benefits and Compensation: There were no direct benefits for participation beyond contributing to scientific knowledge. A modest compensation in the form of a gift card was provided to participants for their time commitment. Participants were debriefed on the full details of the study upon completion.

1.7.1 Use of AI-assisted technologies

This dissertation utilized AI technologies, specifically Grammarly® AI Writing Assistance, to enhance the quality and precision of the written content. NTNU IE Faculty policy currently includes a "Declaration of AI aids and tools" form for bachelor and master's students, but no such official declaration is available at the Ph.D. level yet. Therefore, tools were applied under established guidelines for using AI in scientific papers set by reputable publications in the field, e.g., Elsevier, Frontiers, etc. AI technologies were incorporated to optimize language clarity, improve sentence structure and grammatical correctness, and summarize existing arguments and overall coherence. Grammarly® proved valuable in streamlining tedious tasks, enhancing clarity, and structuring arguments. It was solely used for a meticulous review, improving the readability and language of the work and refining overall writing mechanics. Some regular prompts used in this dissertation are "improve it," "make it clearer," "paraphrase it," "sound fluent," "rewrite it," "make it objective," "make it detailed," "clean up notes," amongst others.

AI was not used to replace vital authoring tasks such as producing scientific, pedagogic, or research insights, drawing scientific conclusions, or providing clinical recommendations. All generated work has been carefully reviewed to avoid output that can be incorrect, incomplete, or biased. The author is ultimately responsible and accountable for the contents of the work.

1.8 Advancement of Knowledge

This work stands at the crossroads of HCI, VR research, QoE assessments, and UX evaluations. While existing research delves into various aspects of VR experiences, This research operationalizes holistic assessments of immersive VR that unite diverse fields and address the intricate interplay between technical, psychological, and experiential factors; the thesis offers a comprehensive framework for evaluating VR experiences, setting it apart from previous efforts.

- **Integration of QoE and UX Dimensions:** The primary contribution of this research lies in bridging the gap between QoE and UX assessments. The proposed holistic approach integrates technological attributes, psychological states, and experiential qualities, providing a comprehensive evaluation model that captures the intricacies of immersive VR experiences.
- **Multidimensional Framework & Novel Typology:** This research responds to the multifaceted nature of VR by developing a dynamic taxonomy that attempts to bridge the gap between systems-oriented, human-centric, and narrative-driven perspectives. This allows for a more granular and insightful evaluation compared to existing frameworks.

- **Presence-Authenticity Dyad for VR:** Exploring authenticity as a distinct dimension complements the concept of presence, offering a novel contribution to the field. The presence-authenticity dyad sheds light on assessing the credibility of the virtual world.

- **Methodological Contributions:** This work employs mixed-methods approaches to attempt nuanced assessments of VR experiences instead of solely quantitative or qualitative. This methodological contribution sets a valuable precedent for future research in this field. This approach to empirical investigations and analysis methods provides valuable insights into the relationship between VR interactions, user behavior, and overall experience.

1.8.1 Open Science Practices

All findings in this research work have been forwarded for open peer review and open access publications, which ensures broader dissemination and facilitates knowledge exchange. This allows research to reach a wider audience and contribute to advancing knowledge more collaboratively and inclusively. Further, research data, protocols, analysis scripts, and other research materials are readily available to fellow researchers to scrutinize methods, replicate findings, and build upon the body of work. This has been in line with the Policy for Open Science at the Norwegian University of Science and Technology, per the recommendations of the Research Council of Norway and the Ministry of Education, Norway.

Data Management

Research data was managed per NTNU's policy, guidelines, and other relevant laws, requirements, and principles. The researcher obtained supervised approval for the data management plan prepared for the research work and for data to be archived and made available in compliance with current guidelines. Individual data from research, such as measurements, facts, and information, usually are not protected by copyright as per Open Science policy. However, data collected for this research is not publicly available for security and privacy reasons.

1.9 Thesis Outline

This section is followed by **Chapter 2**, which provides the background and motivation for understanding IMEx (particularly VR) from a QoE viewpoint. The chapter compiles Publications **I**, **III**, and **IV** into discussions on the compositional structure of IMEx. It combines this with **II** to understand quality in terms of its distinct facets and the elements and features that define it. The chapter also discusses authenticity and presence as quality judgments from Publication **II**. It concludes by describing the methods and measures observed in this thesis. **Chapter 3** presents the work published in **V** and **VI**, examining how users respond to perceived action inside VR and the overt behavior they exhibit. In **Chapter 4**, an examination of the input modalities outlined **VII** and **VIII** is presented. The ensuing results are dedicated to scrutinizing the potential impact of these input modalities on performance metrics and the subjective perceptions held by users regarding their virtual reality encounters. Subsequently, **Chapter 5** encapsulates a summation of our primary contributions, discusses the implications derived from our research, and delineates prospects for future extensions.

Chapter 2

Background

Before one gets too lost in how immersive media stand in contrast to preceding media forms, it is essential to emphasize that no medium or single media event does its techno-cultural work in isolation from other media, any more than it works in isolation from social and economic forces. What is new about IMEx comes from how a medium refashions its predecessors and other contemporary media or how older media have refashioned themselves to answer the challenges of new media – that is, a desire to disappear. To lose opacity and become transparent. A similar willingness was apparent when videoconferencing was celebrated as a livelier and more realistic communication substitute for the telephone call. Or when CGI replaces actors, reimagines cities, and creates a compelling assemblage on the face of reality. It is apparent again in the triumph of the graphical user interface for computers and mobile devices. In this sense, a transparent interface would erase itself so that the user is no longer aware of confronting a medium but instead stands in an immediate relationship with the contents of that medium.

This chapter caters to **O1** and introduces our theoretical framework. Specifically, our holistic taxonomy for VR (in publication **I**), our reformulation of authenticity as judged quality (in publication **II**), and the necessary methodologies and tools to achieve our protocols. We begin this chapter by presenting core concepts from relevant literature vis-a-vis phenomena that drive the QoE of immersive media. We then provide a holistic view of the technical and contextual frameworks that make IMEx. Later on, we present a quality taxonomy for fully immersive VR experiences.

2.1 Immersive Technologies & Experiences

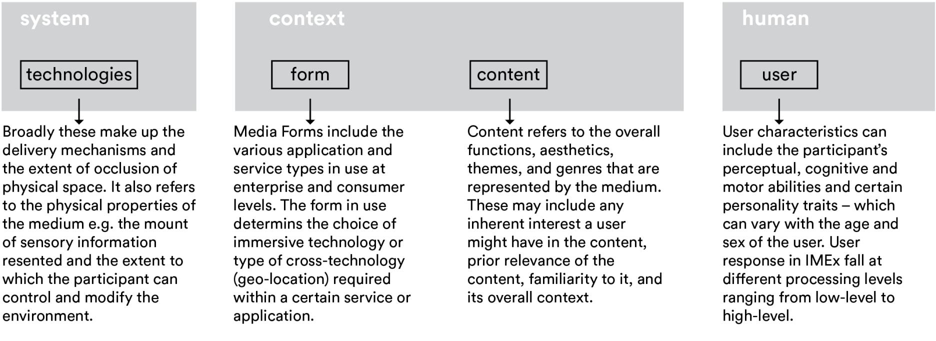
Immersive media are denoted not only emerging technologies but also the unique human experiences they provide. They encompass three main elements (Fig 2.1):

Immersive Media Technologies refer to various emerging technologies that generate omnidirectional experiences where users can perceive content in any direction. These technologies either overlay digital images and information on the physical context or create a new reality by completely occluding the natural context. They range from non-interactive 360-degree viewing to interactive extended realities (XR) along Milgram's virtuality continuum [61], encompassing Mixed Reality (MR), Augmented Reality (AR), Augmented Virtuality (AV), and fully immersive Virtual Reality (VR).

Immersive Media Forms or Domains refer to the various applications of immersive technologies across diverse industries, shaping the structure and mode of interaction with the media. They include video games, interactive digital stories, immersive Cinema, omnidirectional content, training & learning simulations for health and education, and rehabilitation activities like performing motor and cognitive tasks.

Immersive Media Content encapsulates the ideas, information, and experiences of immersive applications. They can be classified into content that caters to:

- seclusion – user isolation from the real environment
- navigation – user movement within the environment
- interaction – user's impact on the environment
- and modeling – environment creation method



Broadly these make up the delivery mechanisms and the extent of occlusion of physical space. It also refers to the physical properties of the medium e.g. the mount of sensory information resented and the extent to which the participant can control and modify the environment.

Media Forms include the various application and service types in use at enterprise and consumer levels. The form in use determins the choice of immersive technology or type of cross-technology (geo-location) required within a certain service or application.

Content refers to the overall functions, aesthetics, themes, and genres that are represented by the medium. These may include any inherent interest a user might have in the content, prior relevance of the content, familiarity to it, and its overall context.

User characteristics can include the participant's perceptual, cognitive and motor abilities and certain personality traits – which can vary with the age and sex of the user. User response in IMEx fall at different processing levels ranging from low-level to high-level.

Figure 2.1: A holistic framework that considers the varying technological, contextual, and compositional facets of IMEx.

The above classification aligns with various applications (passive storytelling, interactive games, solitary exploration, etc.) and helps select appropriate immersive technologies for the intended purpose. Beyond functionality, content also includes eudemonic aspects like aesthetics, themes, and genres, which can elicit various behavioral, cognitive, and affective user responses.

Nevertheless, IMEx is not one-size-fits-all. Each VR application and its experience offer a unique blend of the message it conveys, the reason for reporting it, and its delivery method. For example, technologies with limited interaction and navigation might be more suitable to narrate a specific story. In contrast, we may favor isolation, freedom, and interaction for an experiential simulation. However, such a mix might not be ideal for a particular educational experience. There are no absolute right or wrong choices in immersive media; each combination fulfills a specific purpose. Presenting content with the correct combination of technologies and matched to the expectations and skills of the user will provide a positive experience in the intended media form.

2.2 Experiencing VR

Regarding new multimedia information and communication technologies, the concept of quality has evolved to encompass a diverse array of perspectives that have expanded its scope. Traditionally, quality primarily defined a product or service's ability to meet predefined requirements and specifications [62]. The concept of quality in the context of immersive media, particularly VR systems, requires an optimization of both the objective attributes of the system and the subjective perceptions of users within their specific usage contexts. This integrated perspective on quality evaluation combines elements of engineering, human-centered design, and contextual awareness. It acknowledges the coexistence of quantifiable technical attributes of the system and the complex interplay of individual characteristics, contextual factors, and content interactions [63]. Quality is thus the "degree of delight or annoyance experienced by an individual" while engaging with a VR application or service, contingent upon fulfilling their expectations and needs [7].

Immersive VR has moved our media consumption from the confines of rectangular screens to interactive 360-degree environments that surround us. With VR, the once passive viewer has become an active participant who can navigate, interact, and respond to data in three-dimensional spatialities. This results in a more engaging and immersive experience. The initial bulky and wired prototypes have been replaced by lighter, more innovative, and standalone VR systems no longer tied to a single room. At the same time, the rise of immersive VR has also blurred the boundaries between conventional telecommunications and HCI systems, refining how we approach media consumption and use technology [6]. We believe that to evaluate immersive VR experiences properly, one must include assessing both the Quality of Experience (QoE) and User Experience (UX) elements. QoE frameworks use metrics such as display fidelity, tracking accuracy, and other technology-related aspects. They rely on objective metrics and quantitative methods to measure various attributes. On the other hand, UX evaluations are performed based on the quality of the design, focusing on ergonomics, interaction mechanics, UI/UX design, content creation tools, etc. Qualitative heuristic evaluation and usability testing methods are applied to evaluate UX qualities such as usability, comfort, and learnability.

This convergence offers a distinct set of opportunities and challenges for both fields. In particular, this convergence is most pronounced in its implications for evaluation and assessment methodologies. Assessing the quality of immersive media experiences presents new challenges due to the complexity of interwoven technological, content-driven, psychological, and user-centric factors, which makes standard assessment methods used for media delivery inapplicable [64]. Approaches segregating technical evaluations from users' experience risk creating fractured interpretations. Quality assessments in VR involve evaluating the system, the product (content or application), and the overall user experience [65]. Furthermore, these assessments are intertwined with the psychological states of the user that are at the core of a VR experience. We therefore look at how various sensorimotor contingencies of the system create a perceptual *immersion* [66] [48], thereby emplacing its audience in *immediacy* [16] [67] with the communication medium whereby they experience a feeling of *presence* [68] [9]. The core concepts that describe this unique experiential position are discussed below:

Presence, drawn initially from ecological psychology, refers to the user's experience of being situated in a virtual environment [69] [11] [68]. This experience operates on three levels [30]: (1) a sense of being within a simulated or distant place presented by the media, creating a solid illusion of place, (2) a sense of self, control, and "embodiment" that a user feels when situated and active within the media's simulated reality, (3) an awareness of other sentient beings within the same environment, prompting engagement with them.

Immersion in a technical sense describes how fully a system can present a vivid synthetic environment [38] [70]. This can be influenced by sensory technology, display proximity, and enhanced haptics [71]. Critical features of immersion include the degree of vividness, extensiveness, surroundness, and interactivity [8]. On semantic and diegetic levels, challenge-based or narrative-based immersions are also considered for their influence on experiential quality [37] [38] [72].

Immediacy refers to the media's pursuit of becoming "transparent" or "natural" to users [73]. Immersive technologies offer direct interactions with virtual objects compared to their indirect counterparts, thereby fostering a sense of proximity and psychological closeness [74] [75].

2.3 Quality Perception in VR

Quality involves perceptual, cognitive, emotional, and evaluative processes determining one's conscious perception of things. The study of qualia, or the qualitative and phenomenal aspects of consciousness and subjective experience, is closely related to quality [76]. Qualia encompasses various mental images, feelings, and sensations that reflect what it feels like to experience something. Qualia capture impressions such as goodness, beauty, desirability, and virtue that arise in consciousness when encountering something [77]. According to Jekosch's early description, "experienced quality" refers to the mental evaluation of how something's actual composition compares to someone's expected or ideal composition [78]. In other words, experienced quality is a subjective judgment of how well an entity's perceived composition aligns with desired expectations [79]. The term "entity" refers to any object or event that becomes an object of perception, material or immaterial.

In the physical realm, entities have objective physical attributes (or *quality elements*) that can be measured, like display resolution. But in the subjective realm of perception, entities exhibit psychological features (or *quality features*) like vividness and richness. The subject perceives an entity's quality features and compares these to internal ideals and expectations, which shapes a conscious impression of the entity's overall quality [80]. Möller [81] categorizes quality elements and features into complementary factors and aspects. Technical factors like throughput and jitter can affect perceptual aspects like immersion and embodiment. All contribute to the emergent qualia of overall VR quality experienced in the subject's mind. Quality does not reside in objects but in the subjective impressions they evoke. The sensorimotor contingencies of a VR system can only optimize objective quality elements, while the qualitative experience depends on the meaning and values the users ascribe.

Jekosch's reformulations [78] make a fundamental distinction between two types of quality perception: *perceived quality* and *judged quality* [80]. Perceived quality is similar to low-level thinking, referring to the immediate impression formed upon encountering a stimulus. According to Damasio et al., [82], swift quality assessment upon meeting a natural environment or a technological stimulus need not require deep cognitive processing but results from integrating basic perceptual features into an abstract evaluation. For instance, a user may perceive the quality of a depicted scene as poor upon first seeing it without consciously analyzing why. However, perceived quality can still be intentionally contemplated and judged after the fact. This evaluative process creates a quality judgment that reflects cognitive analysis. The subjective experience of this quality judgment is referred to as *judged quality*. Judged quality encompasses richer perceptual, conceptual, and affective content than perceived quality since cognitive evaluations activate complex associations and interpretations. Furthermore, it is influenced by conscious analysis and reflection, not just direct perception.

The difference between perceived quality and judged quality pertains to the idea of experience, which can be defined as the "stream of perceptions (of feelings, sensory percepts, and concepts)" that arise in a given situation [64]. Perceived quality, therefore, is aligned with the immediate impressions and sensations one experiences in the VR world. It is an intuitive and phenomenological aspect of the experience of that world. However, judged quality goes beyond mere experience. It necessitates additional cognitive processing to assess and consciously judge the quality of the depicted VR world. Hence, while perceived quality is embedded in the immediate experience, judged quality arises from reflective analysis and interpretation of the VR experience. The former is immediate, while the latter requires additional mental effort to arrive at an overall quality assessment.

2.4 The Presence-Authenticity Dyad

Achieving high-quality VR experiences involves optimizing low-level processing and higher-level functions of congruency and coherence to collectively assess stimuli, behaviors, and events within the VE. A user who remains immersed in the VR environment will transition from only perceiving visual and other sensory information to actively seeking possibilities for action and making swift evaluations of the logical sequence of event chains (Fig 2.2). If the notion of presence constitutes an immediate subjective sensation contingent upon the immersive attributes of a system, then how faithfully that system replicates behaviors, relationships, and rules consistent within its purported context are aspects of its credibility (Fig 2.3). Here, a contention emerges for a parallel description of the credibility engendered by the VR system. We refer to it as the *authenticity* of the VR experience.

Authenticity is the degree to which a VR environment reflects the regularities of the world it is trying to represent, according to Gilbert et al. [83] and Bowman [84]. As users spend more time in a VR environment, their sense of wonder may give way to a heightened awareness of its authenticity. They may begin to notice contradictions such as the inability to intuitively interact with virtual objects [85], non-responsive non-player characters [86], or a mismatch between their avatar and the world they inhabit [87]. A virtual world and its content will appear and feel

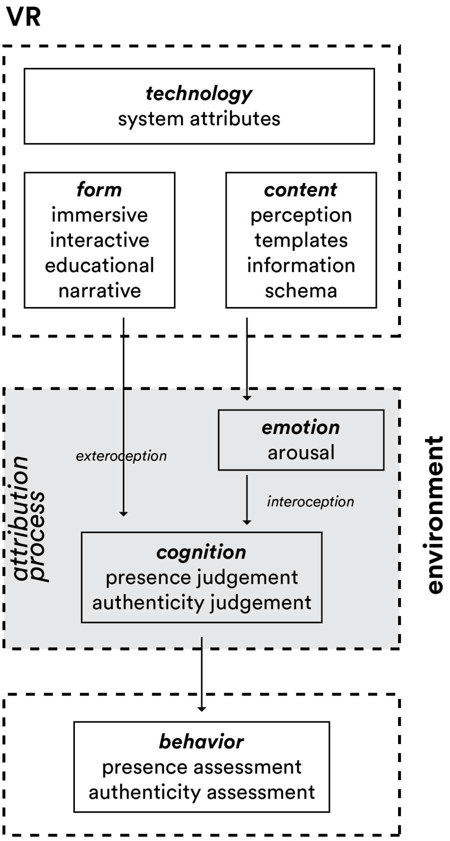


Figure 2.2: The attribution process leading to judgment inside VR, adapted from [3].

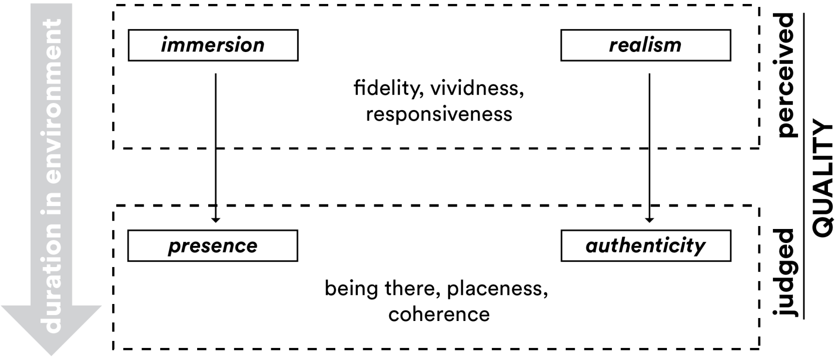


Figure 2.3: Quality judgment process between perceived quality and judged quality

authentic if its behaviors, relationships, and rules remain consistent within that internal context and align with the user's expectations, prior knowledge, preferences, and reciprocity [88] [89]. This shift highlights that while a robust place illusion is necessary, it may prove shallow and lose its spell if the virtual world gives the impression of being inauthentic.

This research, therefore, introduces the "Presence-Authenticity Dyad," recognizing authenticity as a complementary dimension to presence and crucial for evaluating the quality of VR experiences. In similar vocabulary to that which characterizes presence as a feeling of 'being there,' it is proposed that

Authenticity is *"a sense of 'trueness and genuineness'"* felt in a virtual place.

In agreement with Lee [90], we see it as users' judgment of the virtual world's trueness & genuineness regarding its stimuli, content, and behavior. Both trueness and genuineness refer to related but distinct characteristics. Both terms suggest that the user is deliberately judging the depicted place. This contrasts their immediate intuitive impression of the VR setting, which is based on various sensory inputs and atmospherics. When users first encounter the VR space, they rely on sensory percepts to discern the available visuals, sounds, and vectors. However, their initial quality perceptions simultaneously evolve as they compare them to their internally desired quality features (expectations). It is argued that even in the earliest stages of their embodied encounter, the experienced quality of the virtual world is enough to create a sense of presence (though ephemeral). As visitors spend more time in the immersive VR world, they become increasingly aware of the lifelikeness, interactive intuitiveness, audio-visual synchronicity, and other aspects of the VR setting. Quality judgments regarding *trueness* and *genuineness* require conscious assessments of the virtual world's congruence and coherence, reflecting a heightened state of intentional and reflective cognitive processing. Both quality descriptions go beyond initial impressions to include complex and nuanced evaluations of whether the VE maintains its integrity and credibility.

The definition of trueness involves conforming to reality and actuality or agreeing to fact and reality, focusing on the precision of the information in accordance with reality, and reflecting on facts. On the other hand, genuineness is about being honest and sincere, as well as being accurate and appearing as such. It surpasses mere accuracy and delves into sincerity, encompassing the quality of being real without pretense. When determining the experiential quality in VR, authenticity must be comprehended as the combination of the observed factual accuracy of the world (veridicality) and the sincerity of its self-expression (verisimilitude). While its trueness is based on evidence (objective), its genuineness is driven internally (subjective).

2.5 A Holistic Framework for VR Experience

To reiterate, VR is not a result of system capabilities and product qualities alone. It is intrinsically linked to the psycho-phenomenological dimensions it elicits – namely *presence* & authenticity, *immersion*, and *immediacy*. Without them, there is no VR. Therefore, cross-examining various interconnected factors is essential to evaluate the overall user experience arising from technological capabilities and the psychological effects of a VR application. Over the years, there have been numerous prolific frameworks providing a foundation for our understanding of VR [8] [66] [11] [30]. The picture that has emerged over time is of a technology that delivers synthetic media experiences to its users, which are unique in that: they are *immersive* [38] [48] not just like a story [91] [36] but in life-likeness; they are *interactive* and responsive [92] [93] but also *explorable* [94] [95]; and finally, they are *believable* [96] [97] in their appearance and *plausible* in the way they behave [13] [28]. In addition, several self-report inventories and questionnaires have become standard practice for subjective assessments of presence and user experience. However, we find an over-emphasis on either psychological constructs of presence and immersion [98][10] [48] [49] [27] or on emotional responses related to the user experience (engagement, enjoyment, usability, challenge, etc.) [99] [100] [101] [102] [103] [104] [105], none of which represent a complete picture at both systemic and user levels. Synthesis is therefore required for a blended methodology that can holistically combine assessments of system performance and experiential qualities alike [6] [106].

A five-dimensional taxonomy is proposed to address this gap and provide a holistic approach to evaluating VR experiences. The value of a broad taxonomy is also found in its ability to connect theoretical constructs with practical implications. The five abstractions disentangle and redefine often-conflated system-level factors to human-centric features. They include *immersivity*, *interactivity*, *believability*, *explorability*, and *plausibility*.

2.5.1 Factors Influencing VR Quality

It's important to highlight that the five abstractions identified in this taxonomy are significantly influenced by high-level factors, namely system, user, and context. The relative contribution of each abstraction to the overall quality of an application may vary depending on the specific application. Therefore, it's crucial to understand how these factors interact to evaluate the overall quality of a VR experience. It's worth noting that these abstractions are composite qualities that are simultaneously defined and characterized by different elements[78], as illustrated in Fig 2.4. The system's hardware and software components impact the visual, haptic, audio, and other sensory elements, contributing to various psycho-phenomenological effects. User characteristics, such as demographics, physiology, and psychology, can affect their susceptibility to motion sickness, ability to adapt to the virtual environment, and overall engagement. Finally, the context of VR use, including the environment and the specific task or activity, can also influence the user's perception of the experience.

- **System IF:** A VR system's performance is heavily dependent on its hardware capabilities, such as display and graphics, as well as tracking accuracy, latency issues, network connection efficiency, and the system's overall stability and reliability [107]. Regarding multiplayer contexts, operating system and software compatibility, interoperability, and scalability issues can negatively affect the VR experience [108]. Additionally, high-quality content, rendering, and audio are critical to creating immersive experiences [109]. The design and usability of interfaces and devices, as well as their comfort, safety, and ease of use, also significantly shape the overall experience.

- **Human IF:** We must consider various user-related factors such as age, gender, and physical abilities [2]. For example, older users might prefer simpler interfaces or experiences with a slower pace, whereas experienced gamers may want more complex and challenging content [110]. These factors include dynamic ones like experience-of-use, linked to individual differences such as spatial ability, familiarity with VR technologies, and gaming experience. They can influence engagement, presence, dissociation, and learnability [100]. Physiological factors like susceptibility to motion sickness and visual acuity can also significantly impact the VR experience. Additionally, a user's mood, stress level, and mental fatigue can affect their perception of the VR experience, and their imaginative faculties can influence their response to the various illusions within the medium.

- **Context IF:** Environmental conditions impact the performance of VR hardware and the user's comfort. The VR experience can also be affected by the user's surroundings, such as the size of the room, obstacles, ambient light, noise, etc. These factors can influence tracking accuracy, immersion, and safety. Comfort and safety are reported to be affected by the duration and frequency of use [111] [112]. The purpose of using VR, such as entertainment, education, training, therapeutic purposes, etc., is a critical factor that determines the composition of the application. For instance, a VR game would prioritize low latency and high graphical fidelity, while a VR therapy session would focus on user comfort and safety.

2.5.2 Assessments of VR Quality

Various measures and instruments are employed to assess different aspects of the VR experience, ranging from self-reported ratings to objective physiological data. A brief overview thus follows:

- **Self-Reported Measures:** involve asking users to provide subjective feedback on their VR experiences. These assessments often utilize questionnaires or surveys to gauge user perceptions and emotions regarding presence, involvement, engrossment, realism, etc. Some widely used self-reported measures include the MEC-SPQ [113], PQ-ITQ [48], TPI [10], ITC-SOPI [49], among many. Subjective measures that capture users' enjoyment and satisfaction [114] are also regularly employed.

- **Physiological Measures:** are objective measures that capture users' physical reactions to the VR experience, providing insights into their engagement and potential discomfort. These include Electroencephalography (EEG) for brain activity [115], Electrooculography (EOG) for eye movements, Electrodermal Activity (EDA) for skin conductance, Heart Rate Variability (HRV) [116] [117], and more recently, fMRI scans have gained popularity [118] [119]. Eye tracking systems are incorporated to monitor eye movements, revealing user focus, attention patterns, and gaze fixations within the VR environment [120] [121].

- **Performance Measures:** evaluate users' ability to complete tasks or achieve goals within the VR environment. These metrics can be objective or subjective, depending on the specific task or activity [122] [123].

- **Usability & Task Load Measures:** assess how easy and intuitive a VR experience is to use, considering factors like task completion, error rates, time on task, and user satisfaction. Standard usability metrics include System Usability Scale (SUS) [124], NASA-TLX [125], and After Scenario Questionnaire (ASQ) [126], among others.

2.5.3 A Quality Taxonomy for VR

In this section, we elaborate upon each of the five abstractions noted in the proposed taxonomy, breaking it into technical factors responsible for achieving it and the human-centric factors that describe its effects. A detailed version of the taxonomy that includes assessment methodologies for each abstraction can be found in publication **I**.

Immersivity

The extent to which a user feels surrounded by and present inside a virtual environment. It is critical in drawing users into the virtual world, creating a sense of presence and engagement. Several factors affect immersivity, including the sensory modalities available, their fidelity, and their vividness. Representational fidelity involves conveying a sense of place through sensory and symbolic cues.

Technical Factors

1. **Visual fidelity:** VR headsets provide a high-resolution display and a wide field of view for realistic and immersive experiences [8] [1]. The more the field of view (FoV), the better the feeling of being surrounded [17] [11].
2. **Tracking:** VR headsets can track your position and orientation, allowing synchronized movements in the real and virtual worlds. Accurate tracking, gesture recognition, and auditory inputs provide an embodied immersion [38] [127]. Full-body tracking yields the maximum possible embodiment.
3. **Persistence, Latency, & Refresh Rates:** VR headsets with higher frame rates, lower latency, and persistence cause less motion blur and lag, leading to a more comfortable and less nauseous experience [112] [128] – all of which can result in break-in-presence [129].
4. **Audio fidelity:** VR headsets with surround sound and binaural audio enhance the experience of being in virtual worlds [130]. Adding sound effects and ambient sounds further improves the user's sense of envelopment [131].
5. **Headset Types:** Different VR headsets serve different purposes. Some are tethered for better performance, while others are wearable for everyday use. Heavy devices, hanging wires, and loose fits may lead to distraction and discomfort. Lightweight headsets reduce discomfort and motion sickness [48] [132].

Human-Centric Factors

1. **Immersion:** The system's ability to stimulate the user's senses through visual, auditory, and haptic stimuli contributes to a stronger subjective feeling of "being there," or presence, in the VE [91].
2. **Attention:** Factors such as engagement level, environment novelty, and distractions can affect the user's ability to focus on the VE and block out the real world [133].

3. **Embodiment:** The accuracy of head tracking, body tracking, and motion capture systems in representing the user's movements and position in the virtual world determines the user's sense of embodiment [9].

Interactivity

The degree to which the user can interact with the virtual world and influence its events. Controlling the VR experience is essential for fostering engagement and active participation in the virtual world. Intuitive and responsive input devices that enable various actions for a natural and enjoyable experience also enhance a sense of agency or the feeling of being in control of the experience.

Technical Factors

1. **Intuitiveness & Responsiveness:** Using the input devices should be easy and responsive so that users can interact naturally and perform different actions in the virtual world. Slow response times or inefficient data exchange can negatively affect the user experience and satisfaction [134] [135]. Responsive inputs, user-friendly interfaces, and interactive features that meet or exceed the user's expectations yield a positive experience. [136].

2. **Input Modality:** Different input methods, such as gaze, laser, and hand-tracking, offer different ways to interact with the virtual environment [137]. The choice of modality, the task at hand, the challenge level, and user-centered factors can affect the user's experience. [138] [139]. Integrating natural gestures and movements to interact with the virtual environment can enhance immersion and engagement.

3. **Device and Interface Appropriateness:** Usability, aesthetics, utility, and other factors highlight the importance of the ergonomics and functionality of devices and user interfaces in influencing user experience [99] [140]. An intuitive control scheme that matches the user's expectations based on real-world knowledge can ensure the naturalness of interactions. Poor interface quality, mismatches, and unfamiliarity can adversely affect performance and lead to unfulfillment and dissatisfaction.

Human-Centric Factors

1. **User agency and control** When designing virtual reality experiences, it's important to consider the user's sense of control and ease of interaction. Giving users control over their actions and outcomes in the virtual world helps them feel more engaged. [29]. Thus improving the overall sense of agency [127].

2. **Ease of Interaction:** This can be achieved by providing clear feedback and intuitive controls that are easy to use. It's also essential to balance the complexity of interactions with the user's capabilities so they don't become overwhelmed or exhausted. The user interface should be consistent and accessible for extended VR sessions, considering physical constraints and learning abilities.

2.5.4 Explorability

This refers to the ease and freedom of movement with which users can navigate and discover new content within the virtual environment. This aspect is influenced by factors such as degrees of freedom (DoF) for movement, locomotion techniques, and wayfinding or pathfinding options. Other factors like map design, level of complexity, and the overall layout of the virtual world also provide users with opportunities to explore and discover the virtual world.

Technical Factors

1. **Degrees of Freedom (DoF):** The more DoF, the better the experience. It lets you move more naturally and feel immersed in the environment. It also reduces disorientation and motion sickness [84] [141].

2. **Spatial Resolution & Loading Times:** The detail and size of the explorable environment and the time it takes to load new areas or features are essential to meeting the user's natural desire to explore and discover the environment. The more details and features there are, the more fulfilling the experience.

3. **Navigation:** Good navigation tools can help you explore a virtual environment more easily [142]. Wayfinding (the mental component) and travel (the motoric component) are both important for this [143] . They help you understand where you are and how to get where you want to go [119].

4. **Locomotion:** Techniques currently used are motion-based, room-scale-based, slide-teleportation, and arm-swinging [25] [144]. Continuous, unhindered exploration allows you to move continuously and naturally without causing motion sickness [145] [146]. It's important to have good tracking to avoid discrepancies between what you see and what you feel [147].

Human-Centric Factors

1. **Sense of expansiveness:** The virtual environment's ability to provide spatial exploration and free navigation can encourage a sense of discovery and unraveling of the unknown [8].
2. **Spatial awareness and understanding:** Users who possess a clear understanding of the virtual environment's layout and landmarks are more likely to feel confident and motivated to explore further [148] [149].
3. **Curiosity and intrigue:** A virtual environment that actively fosters a sense of wonder and curiosity through encouraging exploration and accessibility to all corners can keep users engaged [150].

2.5.5 Plausibility

The extent to which a VR system can logically explain and remain consistent with real-world principles. It refers to the degree to which the VR environment and its contents exhibit logical congruence, follow common sense, and align with user expectations. Plausibility operates at the syntactic level and reflects in logical consistency, adheres to real-world principles, and feels rational & explainable.

Technical Factors

1. **Perceptual Constancy:** It is crucial to maintain consistency in object appearance despite changing environmental and contextual conditions [2] [151]. This consistency can be achieved through stable geometries and optimized models, which can help create an overall positive experience [49] [27].
2. **Aliasing & Sampling:** Reducing visual artifacts, such as jagged edges or pixelated textures, can help improve the visual continuity of the VR experience. Higher fidelity in geometry, audio, and interaction mechanics is also essential to improve visual realism [33].
3. **Audio Synchronization:** Accurate sound rendering based on virtual distance and location can enhance the aural authenticity of the experience [152].
4. **Physics consistency:** Emulating real-world scenarios or fictional scenarios in real-world settings requires the physics engine to behave realistically regarding gravity, collisions, kinematics, and materials. Consistency of simulated physical interactions using realistic physics engines can enhance authenticity. At the same time, uniformity in rules and logical cause-and-effect chains across the virtual environment can improve the feeling of presence [153].

Human-Centric Factors

1. **Perceived congruence:** Objects and behaviors within the virtual world should be consistent and logical to the user and match the real-world principles [13]. Incongruent features and erratic behaviors can disrupt the authenticity of the virtual world.
2. **Alignment and prior knowledge:** User's personal experiences and understanding of the world shape their perception of what's plausible. Even if not identical to the real world, systems that align with user expectations reinforce a sense of alignment [154]. Incongruence and mismatches may lead to a loss of spatial awareness, feel jarring, and break plausibility [6] [148].
3. **Cognitive dissonance:** Inconsistencies or dissonance between expectations and the virtual world can create discomfort and undermine plausibility [155].

2.5.6 Believability

A primarily user-centric aspect refers to the extent to which a user successfully perceives virtual events and experiences as accurate despite knowing they are not. The extent to which a VR system can deliver an experience with the realism and internal coherence required to make it feel believable for the user. It goes beyond mere visual fidelity and taps into the user's emotions, senses, and overall engagement with the virtual world. It reflects the genuineness of the depicted world in its subtle details and nuances that mimic reality and support a "suspension of disbelief," ensuring that users accept the virtual world as a reality.

Technical Factors

1. **Realism and fidelity:** High-fidelity stimuli, such as realistic render quality, control mapping, physics engine, and spatial audio, can help create a believable world [13] [156]. Physically based rendering, materials, and textures are also essential. All elements within the virtual world, from physics and interactions to character behaviors and story logic, should be consistent and make sense within the established setting and rules. Narrative and stylistic cohesion should also be present. Inconsistencies in cause-and-effect relationships or illogical elements can damage believability [157].
2. **Atmospherics and randomness:** Details that reflect real-world experiences, like environmental imperfections, object interactions, nuanced reactions, and character animations, can enhance the feeling of naturalness within the environment [158]. Attention to detail within virtual worlds can spark curiosity and motivate users to seek new things.

Human-Centric Factors

1. **Suspension of disbelief:** Users are willing to temporarily accept the virtual world as real despite knowing it's not. Engaging storytelling and immersive visuals and audio enhance believability [159].
2. **Scenario Logic:** Speaks to the narrative- and challenge-based immersion within a virtual world. Whether the complexity and realism of scripted events or narratives in the virtual world are logically consistent, it is reflected in how reasonably and predictably the world behaves [160] [161]. This includes characters' behavior, stories, and situations within the VR world [72]. Also, the extent to which the virtual world engages the user's reasoning, skills, and decision-making can heighten their cognitive absorption and make time fly by [133]. The sense that actions and experiences within the virtual world have significance adds to their meaningfulness [162] [37].
3. **Prior VR experiences:** Users with extensive VR experience may have higher expectations for these factors than novices. The literature suggests that individuals with vivid imaginations and susceptibility to suggestion may be open to fantastical elements and more accepting of realistic and fictional VR experiences[83].

Table 2.1: Matrix of QoS, QoE, and Inter-Relational Factors

	Immersivity	***Interactivity***	***Explorability***	***Plausibility***	***Believability***
QoS	The technical fidelity of the VR system.	The responsiveness and accuracy of the system to user inputs.	The system supports various navigation and locomotion techniques.	The system can consistently respond to user interactions and provide expected outcomes.	The system's ability to create a logically consistent virtual environment.
QoE	The subjective feeling of "being there" in the virtual environment.	The user's perception of their ability to interact with the virtual environment.	The user's perception of their ability to explore the virtual environment.	The user's perception of the system's consistency and realism.	The user's perception of the virtual environment's authenticity and predictability.
***Inter-relational**	Drives Interactivity & Facilitates Explorability.	Enhances Presence & Contributes to Believability.	Influences Plausibility & Reinforces Believability.	Affects Presence & Complements Believability.	Strengthens Presence & Interrelates with Interactivity.

The taxonomy provides a comprehensive framework for analyzing and evaluating the quality of VR experiences. VR experiences are complex and multifaceted, and a single metric or criterion cannot fully capture the richness and nuance of these experiences. Differentiating between the five abstractions allows for a more granular understanding of each aspect and its contribution to the overall quality of the experience. The taxonomy is versatile and can be adapted to various cases. For instance, a virtual museum tour might require high immersivity and interactivity but lower explorability [163] [164]. In contrast, human resource skills training might require high plausibility and immersivity but relatively lower interactivity [165] [166]. We can generate better-quality models by differentiating between the five fundamental abstractions (see Table 2.1).

Believability vs. Interactivity: A VR experience can be visually realistic but lack overall realism if the user cannot interact with the virtual world meaningfully. For example, a virtual setting with stunning graphics and textural detail may not remain believable if the user cannot pick up objects or interact with other characters [167].

Immersivity vs. Plausibility: A fully immersive experience may lack plausibility if the user encounters inconsistencies within the virtual world. For example, a user with a high-performance HMD may feel fully immersed in a VE, but if the world's physics are unrealistic or the characters behave in ways that are not consistent with human behavior, it may render the experience inauthentic [86].

Explorability vs. Plausibility: A VR experience can be highly explorable but lacks plausibility if the user cannot predict the consequences of their actions. For example, a virtual world may be large and expansive. Still, suppose the user cannot understand how the world works or predict how their actions will affect the world. In that case, they may not explore it effectively and find little motivation to discover hidden elements or influence the narrative [168].

Immersivity vs. Interactivity: Not all VR experiences involve high interactivity or explorability levels. 360° videos are VR experiences where you remain a passive observer, which are excellent examples of this distinction. This difference is crucial when evaluating VR experiences and justifies differentiating between purely immersive (passive or active) and interactive (minimal or high) experiences within the proposed taxonomy.

The taxonomy emphasizes studying factor interdependencies and their collective impact on quality perceptions. As applications expand, the ability to deconstruct experiences and understand key drivers of positive experiences within VR become valuable. This taxonomy gives a structure for systematically evaluating quality facets and their relationships.

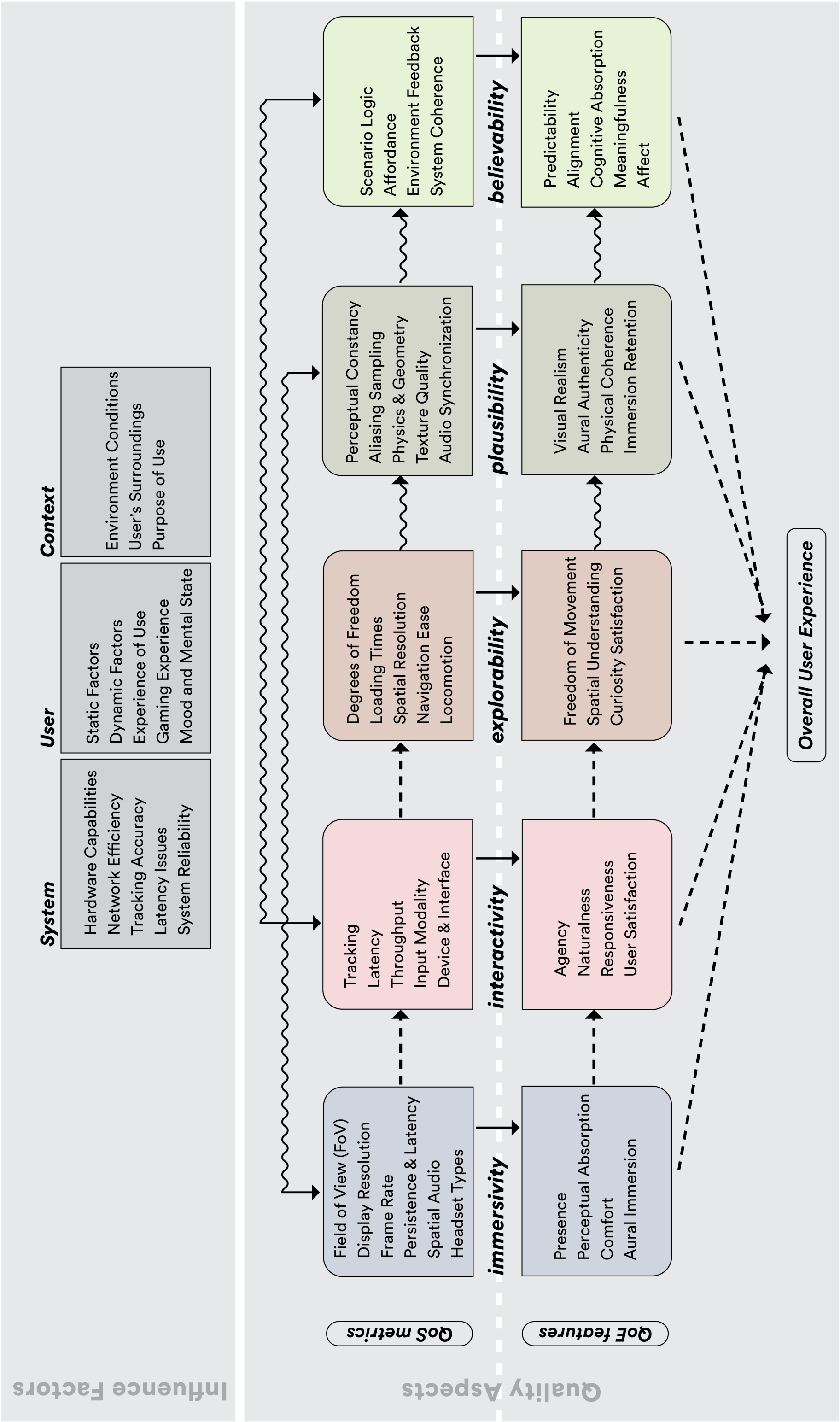


Figure 2.4: A taxonomy for QoS and QoE and inter-relational factors of IMEx

Chapter 3

Correlating Perceived-Actions in VR to Overt-Behavior outside

This chapter presents the results of publication **V** and its extension **VI**, which concentrate on objectives **O2** and **O3**. We define perceived actions as the user's interpretation of stimuli and cues within the VE, coupled with the subjective conviction of their capacity to interact and dynamically shape the contents of the VE [169] [170]. It encompasses the notion that users have control and agency in shaping the virtual world through intentional movements and interactions. It is crucial in enhancing authenticity and presence within the VR context [16].

The basic premise remains that (1) when users feel the capacity to engage with and affect the virtual world, mirroring their interactions in the real world, (2) it heightens their sense of agency, and (3) in turn, amplifies both the authenticity, by closely mimicking real-world experiences, (4) and presence, by fostering an immersive sense of engagement and interaction within the VE.

Users engaged in a VR experience construct mental models of the mediated environment in their minds. These models include understanding the possibilities for action within the virtual space, such as exploring, interacting, and accomplishing goals [66] [171]. Understanding how users perceive the potential for action in VR can be challenging. One possible window could be the decoding of their physical movements outside. If perceived action within VR is intense, then it will likely reflect in their overt behavior. It can be assumed that if a user's observed physical behavior appears purposeful and deliberate, then it's most probably an intended response to the stimulus on offer in the VE. However, the perception of action possibilities and their manifestation in user movements can vary depending on various factors ranging from the design of the VE to the available interactions and mapping capability of the system. A system and its interface may aid or impede users' natural and intuitive behavior. For instance, locomotion techniques that closely mimic natural walking or input modalities that allow direct grabbing both provide the opportunity for an experience that is natural, fluid, and closely aligned with a user's intentions [25] [144].

In publication **V**, we look for correlations between the action possibilities (on offer in the VE) and the overt behavior of the users – do they align? In **VI**, we investigate if the inclusion of action possibilities in the VE correlated to the general sense of presence for users. We hypothesize that (1) responses to action possibilities will also be reflected in overt behavior and that (2) positive correlations between the two will also positively influence the user's subjective VR experience.

3.1 Related Works

3.1.1 Plausibility in Virtual Environments

Real-time rendering has made the application of VR more attractive. It is increasingly seen as a visualization tool that can create realistic prototypes for learning, training, collaborative exercise, and testing diverse physical dynamics and performances [172] [173]. The effectiveness of immersive VEs stems from their close resemblance to our natural navigation, mapping, and manipulation techniques in the physical world. Our natural and intuitive response to perceived actions carries over into these virtual spaces, leveraging the opportunities they present [174]. Just like presence, the concept of plausibility illusion (Psi), the illusion that a virtual event is genuinely happening, is crucial for virtual reality research. This concept underscores the logical and consistent behaviors and events occurring within a specific virtual environment [11]. Psi aligns well with the understanding of quality as a cognitive assessment. Skarbez et al. conducted a study to test this, allowing participants to transition from low-coherence to high-coherence scenarios and modify their virtual avatars' attributes and behaviors to mirror their real-world selves. The result indicated a higher degree of plausibility in the highest-coherence scenario, where users most identified with the best-behaved avatar [13] [175].

For the user, quality is a judgment that distinguishes between perceived quality and expected quality. Previous studies [176] [166] have investigated the impact of plausibility discrepancies on forming a comprehensive sense of presence, emphasizing the necessity to devise protocols for evaluating coherence factors and their consistencies. This study observed specific affordances within an IVE and assessed their influence on user behavior and perceived experiential quality.

3.1.2 Affordances and Perceived Quality in Virtual Environments

We interpret environmental features or items not objectively but in terms of what they enable us to do [177]. This action-centric perception process leverages the opportunities or affordances offered to an entity by its surroundings [143]. Although the concept has a long-standing presence in academic discourse, it's beyond the scope of this work to delve into the details. Nevertheless, it's crucial to stress that affordances aren't purely objective or subjective; they bridge the divide between the two, possessing physical and mental attributes but not being confined to either category [34]. Affordances can thus be interpreted in two manners: (a) as characteristics of the environment or (b) as relationships between an entity and its surroundings. Based on Hassenzahl's [43] hedonic and pragmatic model, we identify four unique types of affordances that range from immediate functional objectives to deeper biological or psychological needs [178], which are:

1. ***Manipulation Affordances***: are directly discerned affordances highlighting the physical/sensory compatibility between the user and the object.
2. ***Effect Affordance***: refers to the operational aspect of an object based on manipulation, which is directly perceived based on the user's cause-and-effect understanding.
3. ***Use Affordance***: is connected with the physical and mental skills of the user employing the appropriate cognitive or usage plans.
4. ***Experience Affordances***: are linked to the user's psychological and biological requirements and are only perceived with accurate knowledge and usage patterns.

Manipulation affordances represent the baseline, denoted by motor actions intended to achieve certain goals (i.e., effect and use affordances). At the pinnacle are be-goals (experience affordances) that inspire actions towards certain ends [178]. Collectively, they illustrate the how, what, and why of potential interactions. For instance, a VE might allow a user to press a button (motor goal). The environment's effect affordance can link the pressing of a button to the activation of a light source—establishing a cause-effect relationship. This sequence of actions could serve a purpose, such as lighting up a scene, and the fulfillment or failure of this purpose results in emotional outcomes, like satisfaction or annoyance [179]. Referencing Hassenzahl again [43], these emotional (hedonic) facets contribute to the user's experience affordance, facilitating the accomplishment of be-goals. The pragmatic aspects of a user's experience arise from the compatibility between the user's abilities and the environment's potential for manipulation, the effect, and use affordances.

In real-world sequences, users expect a light to turn on when they press a button. When this action results in lighting up a scene, it satisfies the intended use. If these actions align with the user's abilities, their successful execution within VEs can elicit joy, while failure might lead to irritation. The various interactions available within a VE and their subsequent outcomes provide users with experiences. Steffen et al. [180]investigated how affordances in VR applications provide them a competitive edge over physical reality in certain scenarios, like simulation-based training. Meanwhile, other research has demonstrated how real-world affordances – such as texture, gradient, handle size, hand size, etc. – can influence user choices and emotional states within VEs [181]. We focus on the psychological aspects of affordances within VEs, particularly their impact on perceived quality and the sense of presence and plausibility.

3.1.3 Evaluating User Behavior in Immersive Environments

For a user, the state of presence, or immersion, is characterized by a shift in attention away from the physical environment and towards the mediated environment, resulting in a behavioral response to the physical and symbolic affordances of the environment [182] [48]. A user can be assumed to be immersed when they react to the affordances or action possibilities of the mediated environment with specific behavioral responses.

The ways users utilize the available tools or interfaces, interact with the environment, and moderate their actions are all essential behavioral aspects that help us enhance the overall QoE in immersive applications. We find Kahneman's [183] dual system of thought useful here: System 1 (fast, intuitive, emotional) and System 2 (slow, deliberate, logical). Surveys and questionnaires [48] [10] [49] [184] have often been used as subjective measures to capture self-reported System 2 reflective processes, which involve skills, mental or emotional states, and other elements that typically require conscious mental effort. However, these questionnaires only offer a fleeting look at a person's actions, thoughts, and emotions. Recently, physiological measures [185] have also gained popularity for recording reflexive System 1 processes, capturing covert, subconscious behaviors that are more intuitive and automatic. Since behavior within IVEs often mirrors that in the real world, there is a strong case for an observational methodology to study subject behavior. We assert that observation methodologies can provide an effective means of measuring the QoE.

3.2 Motivation

Advances in commercially available VR technology in tracking and mapping have improved locomotion and navigation techniques within VEs. Such advancements open up avenues to understand the relationship between perceived action, user behavior, and the overall user experience. Numerous studies have explored locomotion techniques, evaluated user perception, and employed subjective and physiological assessments to understand the mind and body interactions in mixed and immersive media from a QoE perspective. However, there remains a crucial gap in research regarding the comparative analysis of user behavior between virtual and physical environments. We assume the presence of a dynamic between users' perception of action and its influence on their overt movements. We attempt to investigate this empirically, employing behavioral observation methods, video-based analysis, and surveys.

3.3 Methods

Motivated by the potential of VR in behavioral studies [10], this research investigates two distinct virtual architectural environments: a "passive walkthrough" and an "interactive walkthrough." Both environments offer similar spatial experiences, with the latter introducing additional interactivity features. The objective was to investigate whether adding affordances—and interactivity features within IVEs affected human behavior outside to assess any correlation between executed virtual actions and observed physical movements. Further, if any behavior changes also corresponded with the experiential performance of the IVEs.

We used video-based observation and time-use surveys to analyze user behavior in the two environments. Profile surveys were collected for subjective assessments, and self-reported pre/post-experience questionnaires were used.

3.3.1 Design

An empirical, comparative study with mixed methodologies assessed two IVE types. Two visually identical virtual models were used, manipulating only the affordances of the environment. A repeated measures design was designed with a single categorical group having two conditions:

a. ***Passive Walkthrough (PW)***, an immersive environment with navigation affordances but no interactive features;

b. ***Interactive Walkthrough (IW)***, an immersive-interactive environment with navigation affordances and a few manipulation and effect affordances.

For the two conditions, user behavioral patterns were analyzed against the subjective experiential scores reported by each participant to verify whether *manipulation and effect affordances in the immersive-interactive (IW) scenario would result in higher perceived experiential quality and higher behavioral activity compared to the non-interactive scenario (PW).*

Participant behavior was recorded into a single behavioral category of "locomotion", and subdivided into two types:

1. ***Durational Events***: Sit, Stride, Sit
2. ***Non-Durational Events***: Point-and-Click, Turn, Bend, Extend, Shrink

3.3.2 Participants

N=34 participants (18 male, 16 female, µ= 26.7 ± 6.7) participated in the study. They were recruited via mailing lists and an online registration form. People from diverse backgrounds and mixed demographics signed up for the study. Participants received gift vouchers for their participation.

All participants were active users of multimedia technologies. Most had prior experience with head-mounted displays but no previous experience using VR in a lab setting. Ten participants reported no competence in VR, 15 participants had basic competence, and 9 said intermediate competence in using VR applications.

68 total experiences (N = 34 × 2 scenarios/subject) were recorded 3.1. Participants tried both scenarios in a randomized order to cope with carryover effects. Out of 34, two sessions (for subjects S4 and S10) were excluded because of incomplete video data.

3.3.3 Stimuli

The model for the virtual environment was created with Trimple Sketch-Up Pro and enhanced with texturing, lighting, and interactivity elements in UnReal Engine (UE4). The environment represented the interior of a one-bedroom apartment with specific dimensions of 32 feet on each side. The layout featured an open-plan design where the kitchen extends into the living room, and there was a balcony accessible from both the living room and the bedroom. Additionally, there was a bathroom and storage area within the VE. Fig 3.2 shows an image of the VE.

Both scenarios supported natural walking movements within the VE. They utilized a point-and-teleport technique for navigation. IW incorporated manipulation affordances and effect affordances in addition to navigation affordances. All doors within the virtual environment are initially open, allowing users to move throughout the interior space freely.

Passive Walkthrough (PW)

The model used high-poly geometry and was prepared using datasmith in Unreal Engine 4, with a focus on simulating real-world materials through physically based rendering (PBR) texturing and realistic lighting. To enhance immersion, a spatial soundscape was implemented. The environment was optimized for use with HTC Vive Pro, and both handheld controllers can be used for point-and-teleport navigation. Hidden collider components were also applied to surfaces to avoid unrealistic perforation effects and users teleporting or walking through walls.

Interactive Walkthrough (PW)

This model has all the features from the PW. In addition, the IW scenario also uses additional manipulation and effect affordances. (Figure 3.3 shows a participant engaged with the environment. While one handheld controller was used for point-and-teleport, the second was used for the interactivity features that include:

a. Two light toggles around the average eye level: A familiar design feature, a button (with a light-bulb icon) provided the required cognitive affordance and the opportunity for manipulation using the handheld controller. A laser pointer (similar to real-life pointers) could be directed at the button to toggle on–or off. The explicit manipulation affordance was immediately satisfied with the effect affordance, as the user should notice their actions, resulting in additional scene lighting.

Figure 3.1: A user explores the immersive virtual environment using a tethered VR headset inside our lab.

Figure 3.2: A view of the virtual living room inside the virtual architectural interior model (Unreal Engine © Epic Games).

b. Six operable doors around the average waist height: Interaction with doors was communicated via metaphorical affordance, i.e., the imitation of real-life door handles. However, the familiar and explicit affordance of hold-and-twist was not present; instead, their manipulation was possible through a hidden affordance activated when a user clicked the handheld controller closer to the handle. The effect and pattern affordances were revealed to the user through successive movements, resulting in learning how to open a virtual door.

c. Six cabinets and drawers at various heights: Different height levels were used to assess the naturalness of the user's behavioral response. The cabinets used the same manipulation and effect affordances as the doors. All doors in this scenario were closed by default, so users had to open them using the handheld controllers to access different spaces.

3.3.4 Premises & Apparatus

The experiment occurred within the department's dedicated VR laboratory, which occupies an area of approximately 16 × 19 feet. This laboratory is specifically de-

Figure 3.3: A participant in the IW condition bends down to explore interactivity options for a kitchen appliance.

signed to facilitate both subjective and physiological assessments. The VR simulation was executed on a 64-bit Windows 10 Pro desktop PC. The PC was equipped with an Intel Core i7 7700 processor operating at 3.6 GHz, 32 GB DDR4 SDRAM with a frequency of 2800 MHz, and a single 3 GB NVIDIA GeForce GTX 1060 graphics card.

To explore the virtual environment, participants utilized the HTC Vive Pro Head-Mounted Display (HMD), which supports six degrees of freedom (6DOF) and motion tracking. The HMD offers a total resolution of 1440 × 1600 pixels per eye, with a refresh rate of 90 Hz. It boasts a field of view (FoV) of 110 degrees and provides 3D spatial audio capabilities. Participants were designated a fixed play area measuring 10 × 14 feet in the laboratory for moving around.

To monitor the participants' activity and identify any undesired artifacts or malfunctions in the graphics or interactivity, the VR experience was externally displayed on a 65-inch Samsung Full-HD TV. This display served as an observation tool during the experiment.

3.3.5 Procedure

Experimental sessions were scheduled in advance using Google Forms, each accommodating only one participant at a time. Each time slot allotted 60 minutes, which included testing both scenarios and completing the corresponding questionnaires. Upon arrival, participants were greeted by the moderator and asked to fill out a 10-item background information survey. Following that, participants were introduced to the HTC Vive Pro controllers and instructed on using them through a brief tutorial within the SteamVR Home space. Subsequently, participants received a set of instructions detailing the experimental procedure.

To ensure their informed consent, all participants were required to sign a consent form indicating their willingness to participate. The experiment consisted of Passive Walkthrough (PW) and Interactive Walkthrough (IW). The order in which participants engaged in these scenarios was intentionally randomized for each individual to prevent any potential carryover effects. Participants were allowed to spend as much time as they desired in each scenario.

After each scenario, participants completed a post-experience survey, assessing their experiential VE evaluation. Once they had experienced both scenarios and completed the questionnaires, participants were thanked for their participation and compensated for their time.

3.3.6 Materials

Subjective Measures

For this experiment, we utilized the ITC-SOPI (see 1.5.4). The experimental protocol began by collecting background information from participants, including demographics, digital proficiency, and their level of competence in using VR. Afterward, participants completed a post-experience questionnaire following each scenario. In addition to rating their experiences, participants could provide additional comments at the end of the rating process. The ITC-SOPI questionnaire, encompassing these four aspects, allowed researchers to gather comprehensive data on participants' experiences within the virtual environment.

Time Log

The application for each use created a run-time log. A combined run-time of just over 10 hours was recorded for all participants in both experiences.

Video Data

Participant activity was video recorded while allowing for experimental control. Over 10 hours of video data was collected. The choice of video-based observation allowed subjects to express themselves unobtrusively and facilitate natural behavior. Behaviors of interest were annotated to perform analysis.

Behavioral Observation

The data were post-processed for analysis and observation coding inside open-source event logging software, BORIS (Behavioral Observation Research Interactive Software) [42]. All behaviors were coded based on manual video analysis by a single person to ensure reliability. Codings were done in an ethogram (details follow in the next subsection). Observations were coded for each subject in each scenario. Two main types were determined:

1. ***State Events***: durational events that have beginnings and ends.
    - Still: The subject remains stationary in one position in the physical space. They might sway, bend, or rotate while on the same point without an intentional step.
    - Stride: The subject moves intentionally forward or backward from their stationary position. This can be one complete stride or more.
    - Sit: The subject assumes a sitting position.
2. ***Point Events***: non-durational events that are generally momentary only.
    - Click or Point: The subject clicks the controller in space either close to the body or away from it.
    - Turn: The subject rotates 90 to 180 degrees about an axis in a stationary position or during movement.
    - Bend: The subject bends forward or backward.
    - Extend: The subject extends their limbs or part of their body outward to touch, kick, or peek at objects inside the virtual environment.
    - Shrink: The subject draws their limbs or part of their body inwards as a gesture of cautiousness or alertness inside the virtual environment.

Since active run times for participants varied considerably, a uniform 3-minute observation time was used. A 3-minute interval/slice was randomly selected from the functional run-time sequence of each user.

3.4 Results

An experiment was designed with a single definite group at PW and IW levels. Four dimensions from the ITC-SOPI, namely, SP, EN, NV, and NE, comprise our study's measured quantitative variables. The manually coded participant behavior types from video analysis made the second set of dependent variables. Thirty-four participants evaluated two virtual scenarios, out of which two sessions were excluded due to incomplete data. Sixty-four data entries (2 per subject X 32) were received and analyzed. Observations were conducted at three-minute intervals per condition across all subjects to maintain uniformity since the run times varied significantly among participants. Below, we discuss the results.

3.4.1 Observed Behavior and Time-Use

To determine how users behaved, we begin with the collective means for run-time and activities of all participants in Fig 3.6. In PW, the total run-time was 284 minutes and 25 seconds (or 17,065 seconds), while for IW, it was 337 minutes and 38 seconds (or 20,258 seconds). Although the mean run-time for IW (µ= 596, SD = 241) was higher than that of PW (µ= 502, SD = 232), an ANOVA analysis revealed no statistically significant differences in run-time between the two conditions: $F(1,62) = 2.64$, $\underline{p} = 0.109$, $\eta^2 p = 0.038$. Refer to Fig 3.4.

A time-use assessment compared the two dominant events, Still and Stride states, in both test conditions. Fig 3.5 displays their distribution in each condition. Still-to-Stride ratio was 75:25 for PW and 79:21 for IW, indicating a marginal difference. In the Stride state, the mean duration was higher in PW (µ= 44.92, SD = 29.8) than IW (µ= 36.2, SD = 23.03). However, no significant difference was found between the two conditions: $F(1,62)=1.713$, $\underline{p} = 0.195$, $\eta^2 p = 0.027$. For the Still state, the mean duration was slightly higher in IW (µ= 139.48, SD = 26.83) compared to PW (µ= 133.6, SD = 32.8), but again, no significant difference was observed between the two conditions: $F(1,62)=0.621$, $\underline{p} = 0.434$, $\eta^2 p = 0.010$.

We looked at the frequency of point events (Click, Turn, and Bend). There were no significant differences between the IW & PW for Turn and Bend events: $F(1,62) = 2.209$, $\underline{p} = 0.142$, $\eta^2 p = 0.034$, and $F(1,62) = 2.009$, $\underline{p} = 0.161$, $\eta^2 p = 0.031$, respectively. However, a significant difference was observed for Click events: $F(1,62) = 4.771$, $\underline{p} = 0.03$, $\eta^2 p = 0.071$. Participants recorded a higher frequency of Click events in IW (µ= 30.5, SD = 14.3) compared to PW (µ= 23.3, SD = 11.9).

3.4.2 Subjective Experience Reported

The questionnaire data were analyzed by calculating the mean opinion score (MOS), representing all participants' average judgment for one scenario. To examine the potential effects of active run-time (the duration of time spent within the virtual environment) on the four dimensions of the ITC-SOPI, a multivariate analysis of covariance (MANCOVA) was conducted, with active run-time as a covariate. Statistical significance was set at $p = 0.05$.

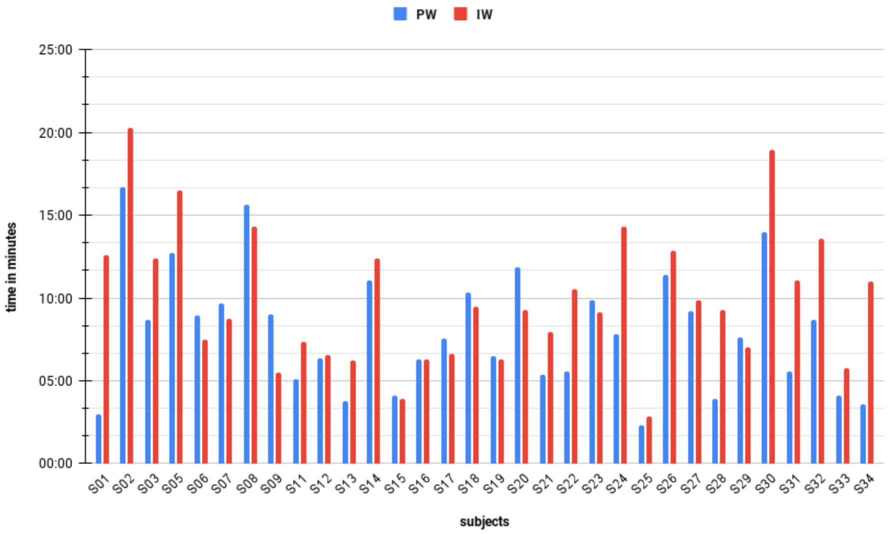


Figure 3.4: Run-time activity for participants in conditions IW and PW..

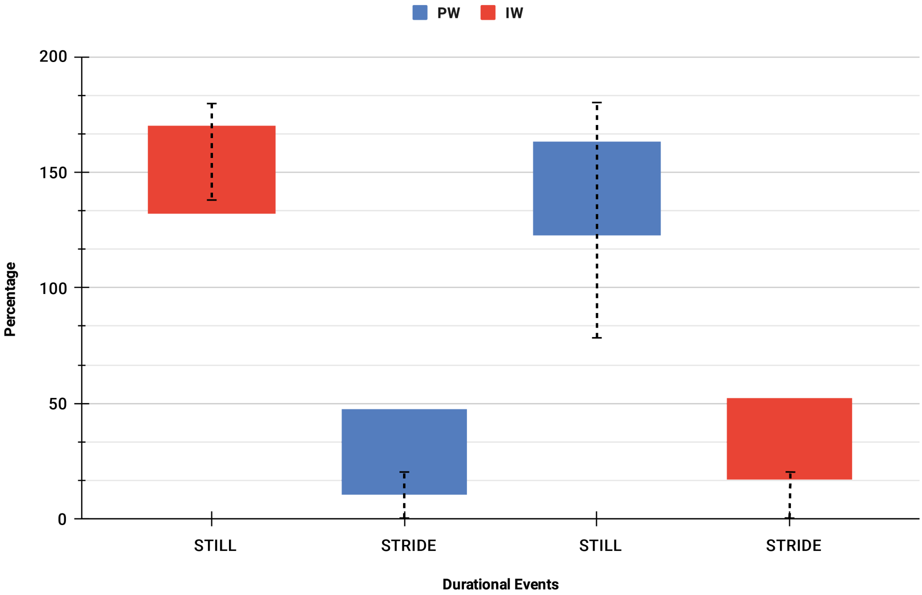


Figure 3.5: Still vs. Stride events in condition PW & IW.

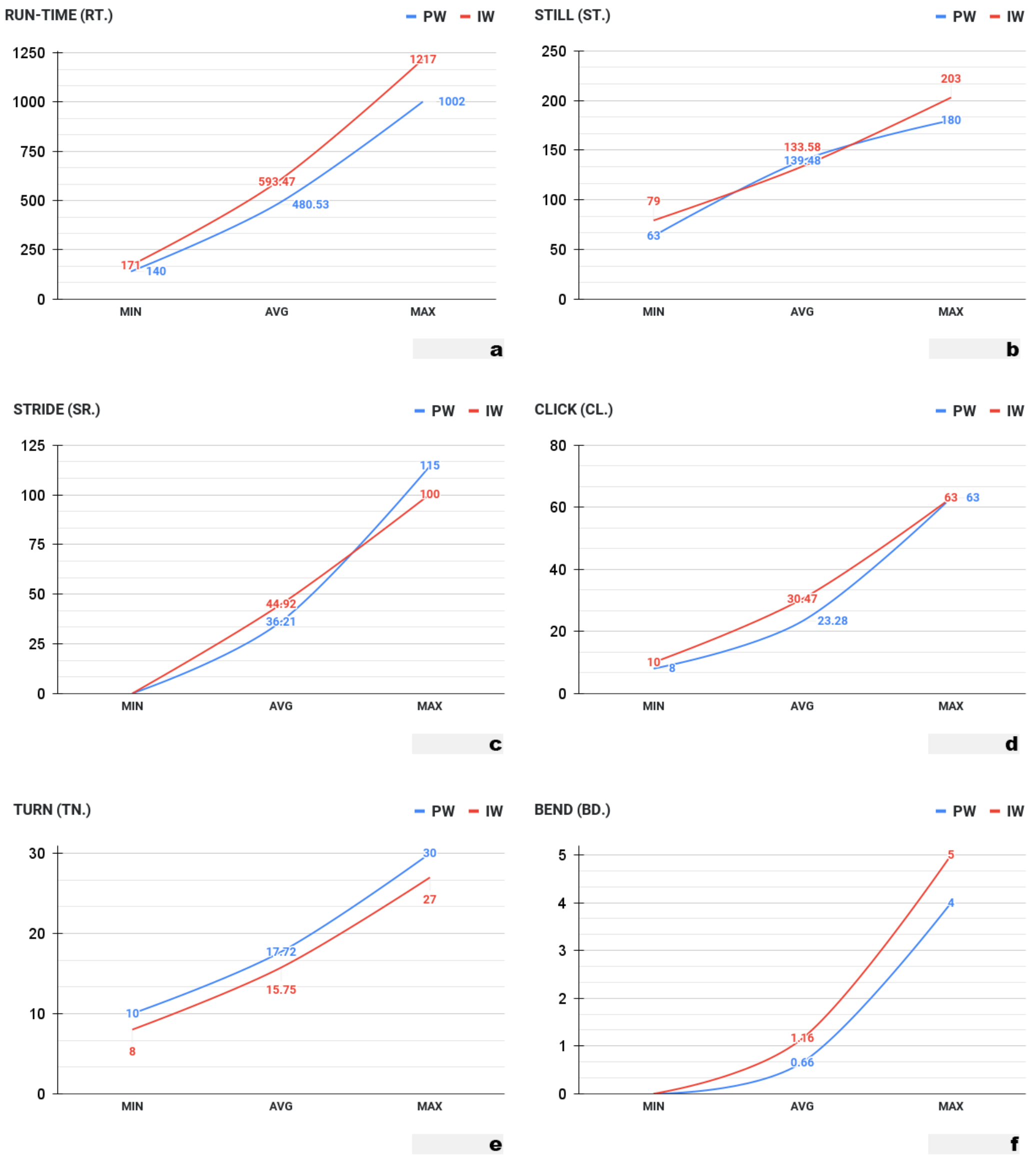


Figure 3.6: Collective Min & Max for user behavior under PW & IW: (a) Run-Time, (b) Still state, (c) Stride state, (d) Click event, (e) Turn event, and (f) Bend event

The results indicated a statistically significant difference between the two categorical scenarios (PW and IW) in terms of the combined dependent variables: spatial presence (SP), engagement (EN), naturalness (NV), and negative effects (NE), even after controlling for active run-time. The MANCOVA yielded the following results: $F(4, 59) = 4.662$, $\underline{\mathrm{p}} = 0.02$, Wilk's $\lambda = 0.76$, $\eta^2 p = 0.24$.

Table 3.1: Collective Means for ITC-SOPI items under both conditions. Means were calculated for all given questions under each item for all participants.

Item	Condition	Mean (µ)	Std.Dev
Spatial Presence (SP)	PW	3.18	0.61
	IW	3.65	0.50
Engagement (EN)	PW	3.56	0.57
	IW	3.89	0.50
Naturalness (NV)	PW	3.66	0.73
	IW	3.84	0.64
Negative Effects (NE)	PW	2.13	0.91
	IW	1.94	0.79

To further explore the differences between the scenarios, separate analyses of variances (ANOVAs) were conducted for each dependent variable (SP, EN, NV, NE). The findings revealed significant differences between PW and IW for SP and EN. However, no statistically significant difference was found for NV or NE.

- Spatial presence (SP: $F(1, 62) = 43.50$, $\underline{p} = 0.001$, $\eta^2 p = 0.166$.
- Engagement (EN): $F(1, 62) = 26.00$, $\underline{p} = 0.017$, $\eta^2 p = 0.089$.
- Naturalness (NV): $F(1, 62) = 133.0$, $\underline{p} = 0.279$, $\eta^2 p = 0.019$.
- Negative Effects (NE): $F(1, 62) = 75.40$, $\underline{p} = 0.383$, $\eta^2 p = 0.012$.

We find the comparison of the means of the three items under both scenarios in Table 3.1. The mean levels in the table indicate higher values for the IW scenario in at least two categories of SP and EN.

3.5 Discussion

The above results indicate that user-perceived experiential quality improved from IW to PW, while user behavior showed no significant difference across the sample. Notably, the number of Click events increased in IW, indicating greater action possibilities. The study found that the inclusion of even a few manipulation and effect affordances significantly enhanced the *place illusion* inside the IVEs. User-perceived spatial presence increased from PW to IW, with a $\underline{p}$-value of 0.001. Although PW offered multi-directional viewing, it remained passive, while the interactions in IW made the environment feel more active. Subjects did not feel immersed in a lifeless world but in one that responded to their actions. Some participants expressed this in their written feedback. The level of engagement or involvement also increased significantly in IW (p-value = 0.017), indicating the significance users attached to the stimuli or activities within the virtual environment. This emphasizes the importance of affordances in providing opportunities for action within IVEs. Both scenarios received high mean scores. However, the inclusion of manipulation and effect affordances did not significantly impact participants' perception of naturalness, life-likeness, or persuasiveness of the virtual environment (p-value = 0.279). Similarly, there was no notable difference in negative effects between the scenarios (p-value = 0.383). In this section, we expand the collective results with a by-subject comparison for further understanding. Observational assessments were conducted for individual subjects. We considered select subjects based on the criteria groupings of time spent, subjective feelings inside, and sickness due to activity transitions.

3.5.1 Perceived Experiential Quality Improves in IW

The study's most notable finding was improving user-perceived experiential quality from Passive Walkthrough (PW) to Interactive Walkthrough (IW). This shift in user perception can be attributed to several factors:

Increased Action Possibilities

One key factor contributing to the enhancement of experiential quality in IW is the more significant number of Click events observed (20 in PW, 43 in IW). These Click events represent user interactions with the virtual environment, such as selecting objects or triggering actions. In PW, the limited interactivity resulted in fewer Click events, while IW offered users more action possibilities. This increase in interactivity likely fostered a sense of agency and control, making the virtual environment more dynamic and engaging.

Enhanced Place Illusion

Including manipulation and effect affordances in IW they significantly enhanced the place illusion, as evidenced by the increased user-perceived spatial presence (SP) (SP in PW: 3.18, SP in IW: 3.65). Despite offering multi-directional viewing in PW, the environment remained passive, lacking responsiveness to user actions. In contrast, IW's interactive nature immersed users in a world that responded to their interactions. This shift from a lifeless world to one with perceived responsiveness contributed to the heightened spatial presence.

Greater Engagement and Involvement

The significant increase in user engagement and involvement in IW further supports the importance of affordances in shaping user experiences. Affordances, in this context, refer to the opportunities for action within the virtual environment. The fact that users attached greater significance to the stimuli and activities within IW highlights the role of interactivity in fostering engagement and immersion.

3.5.2 By-Subject Comparison for Time Spent

To better understand the observed behaviors, we conducted a by-subject comparison based on specific criteria, including time spent, subjective feelings inside the virtual environment, and any reported sickness due to activity transitions.

Subject S15

This participant exhibited a relatively short duration of engagement in both scenarios (see Fig 3.7), indicating that the transition from PW to IW did not significantly impact their time-use behavior. However, despite limited engagement, the participant reported higher scores in IW compared to PW for SP(PW: 2.72, IW: 3.5), EN(PW: 2.62, IW: 3.31), and naturalness NV(PW: 3.8, IW: 4.2). These increases in SP, EN, and NV suggest a noticeable difference in the still-to-stride ratio between the two conditions (PW: 52:48, IW: 76:22). In IW, the participant remained stationary for longer durations to interact with objects, as evidenced by the nearly doubled number of Click events (PW: 20, IW: 43). See Table 3.2 for all comparisons. This observation aligns with the overall trend of increased interactivity leading to higher experiential quality.

Subject S27

This participant consistently logged a median run-time for both scenarios (see Fig 3.8), indicating that the change in the scenario did not significantly affect their time-use behavior. The participant reported slight increases in scores for all three dimensions in IW: SP(PW: 3.1, IW: 3.7), EN(PW: 3.8, IW: 4.15), and NV(PW: 4.0, IW: 4.4) (Table 3.2). The similarities in behavior across the two scenarios suggest that this participant may have had a relatively stable perception of experiential quality, regardless of interactivity levels. However, the participant's subjective ratings improved in IW, indicating a preference for the more interactive environment.

Table 3.2: Durational and Non-Durational Data for S15, S27 and S05

Subject	Condition	Durational (%)		Non-Durational (no.)			
		still	stride	click	turn	bend	Run Time
S15	PW	52	48	20	26	1	4m 8s
	IW	76	22	43	26	3	3m 56s
S27	PW	68	29	11	14	0	9m 13s
	IW	74	26	10	18	0	9m 52s
S05	PW	99	0	25	20	2	12m 45s
	IW	99	0	51	10	0	16m 31s

Figure 3.7: Ethogram showing event log for S15 under both conditions PW & IW. The X-Axis shows time whereas the Y-axis shows durational and non-durational behavior.

Figure 3.8: Ethogram showing event log for S27 under both conditions PW & IW. The X-Axis shows time whereas the Y-axis shows durational and non-durational behavior.

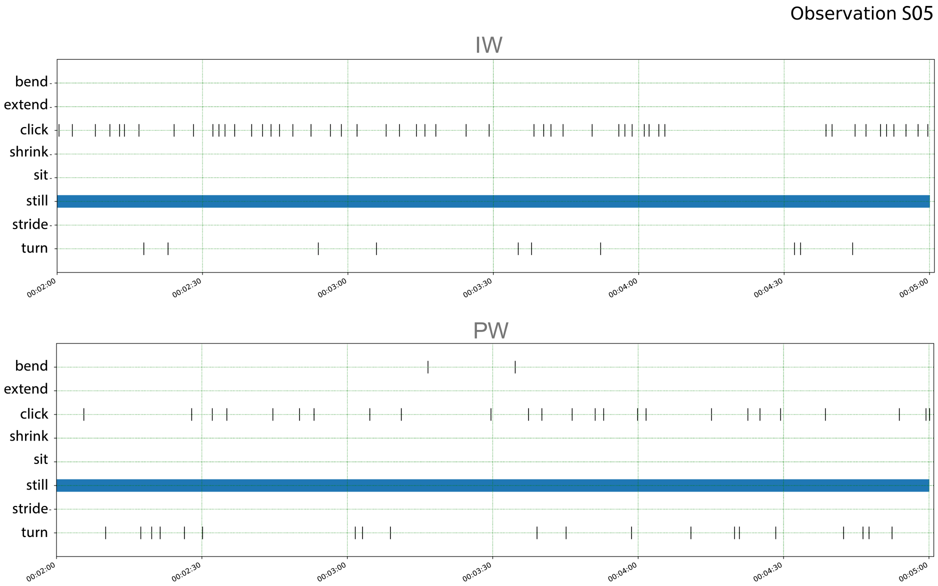


Figure 3.9: Ethogram showing event log for S05 under both conditions PW & IW. The X-axis shows time, whereas the Y-axis shows durational and non-durational behavior.

Subject S05

In contrast to the previous two participants, S05 (see Fig 3.9) consistently maintained long engagement durations in both scenarios. However, a closer examination of their behavior reveals a key distinction. Despite the increase in Click events from PW to IW, the SP score did not see a significant boost SP(PW: 3.4, IW: 3.5). This suggests that while interactivity was present and users engaged more actively with the environment, it did not uniformly translate into higher perceived spatial presence. Interestingly, S05 reported higher levels of engagement and naturalness in PW compared to IW (Table 3.2). This discrepancy could be due to the participant's preference for a more passive experience, even though they were physically engaged with the virtual environment.

3.5.3 Subjective Feelings and Presence

In this section, we explore the impact of subjective feelings on presence within the virtual environment. We consider four individual subjects: S14, S23, S24, and S26, based on their responses to Spatial Presence and Negative Effects questions in the ITC-SOPI questionnaire.

Table 3.3: Durational and Non-Durational Data

Subject	Condition	Durational (%)		Non-Durational (no.)				
		still	stride	sit	click	turn	bend	extend
S14	PW	97	3	0	21	14	0	0
	IW	79	21	0	55	23	0	1
S24	PW	51	49	0	17	19	0	0
	IW	79	21	0	20	8	0	0
S23	PW	57	43	0	12	19	0	0
	IW	70	30	0	18	21	0	0
S26	PW	81	19	0	29	11	1	0
	IW	88	12	0	22	24	5	1
S08	PW	51	46	0	24	4	2	0
	IW	59	41	0	20	14	5	6

Subject S14 – Highest Presence Score in PW

The experience of S14 is exciting as they achieved the highest mean score for presence in PW among all participants SP(PW: 4.44, IW: 4.39). During the observation, the participant mostly remained stationary in PW but displayed lateral movement and bodily reactions in IW (See Fig 3.10 and Table 3.3). Despite increased activity, the higher SP score in IW could be attributed to a combination of factors. First, the novelty of interactivity in IW may have positively influenced their perception of presence. Second, the sense of agency and control afforded by interactivity may have contributed to the enhanced spatial presence.

Subject S24 – Lowest Presence Score in PW

S24, on the other hand, reported the lowest SP score in PW. Their scores were SP(PW: 2.33, IW: 3.22). However, their behavior exhibited a relatively even distribution between still and stride events (See Fig 3.10 and Table 3.3). This suggests that other factors beyond physical activity influenced their perception of presence. It's possible that the content or design of the virtual environment in PW did not effectively engage this participant, leading to a lower SP score.

Subject S26 – Highest Presence Score in IW

S26 reported the highest SP score within the IW condition. With scores being SP(PW: 3.44, IW: 4.67). Their behavior leaned toward still events but included regular interspersed lateral movements and bodily gestures (See Fig 3.11 and Table 3.3). This combination of physical activity and perceived spatial presence in IW underscores the role of a dynamic and responsive virtual environment in enhancing presence. It's evident that the interactive elements in IW significantly contributed to this participant's heightened sense of presence.

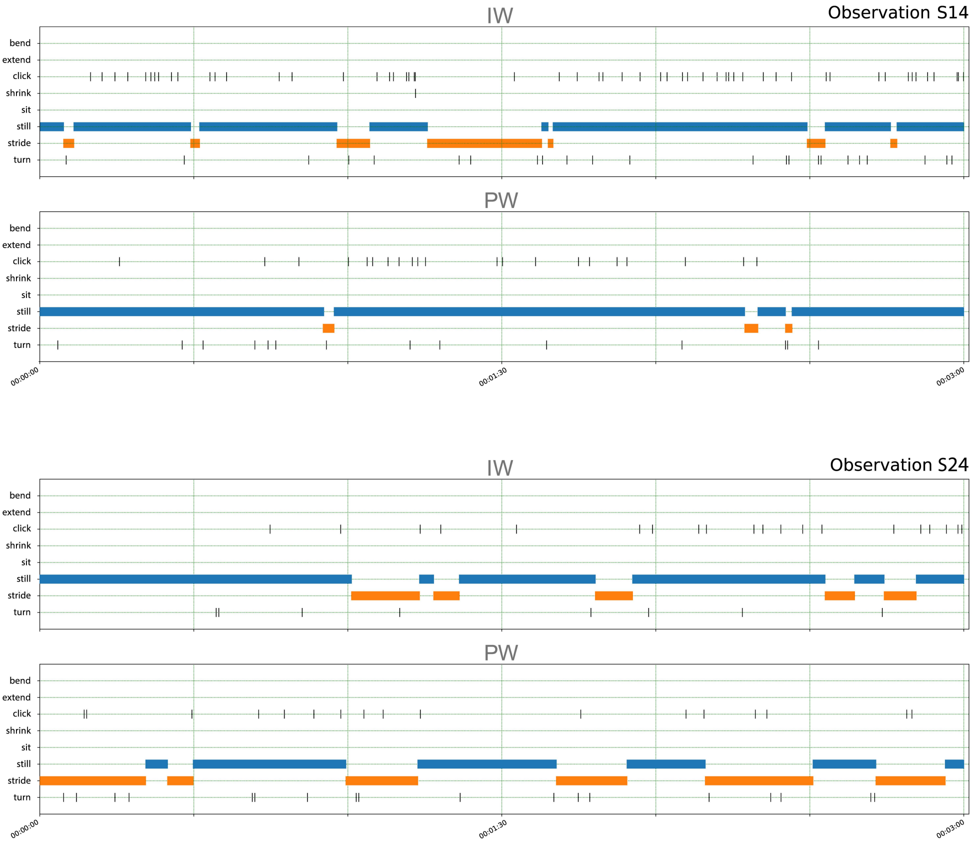


Figure 3.10: Ethogram showing event log for S14 and S24 under PW & IW. The X-axis shows time, whereas the Y-axis shows behavior.

Subject S23 – Lowest Presence Score in IW

S23 underwent PW first, followed by IW, and reported the lowest SP score within the sample in both conditions SP(PW: 2.89, IW: 2.0). Despite frequent transitions between still and stride states (See Fig 3.11 and Table 3.3), the interactive features of IW did not result in significant behavioral adjustments or a positive SP score for S23. This suggests that for some participants, interactivity alone may not be sufficient to enhance presence if other factors, such as content quality or personal preferences, are not adequately addressed.

3.5.4 Activity Transition and Sickness

The final aspect of the discussion pertains to the relationship between activity transitions and reported sickness within the virtual environment.

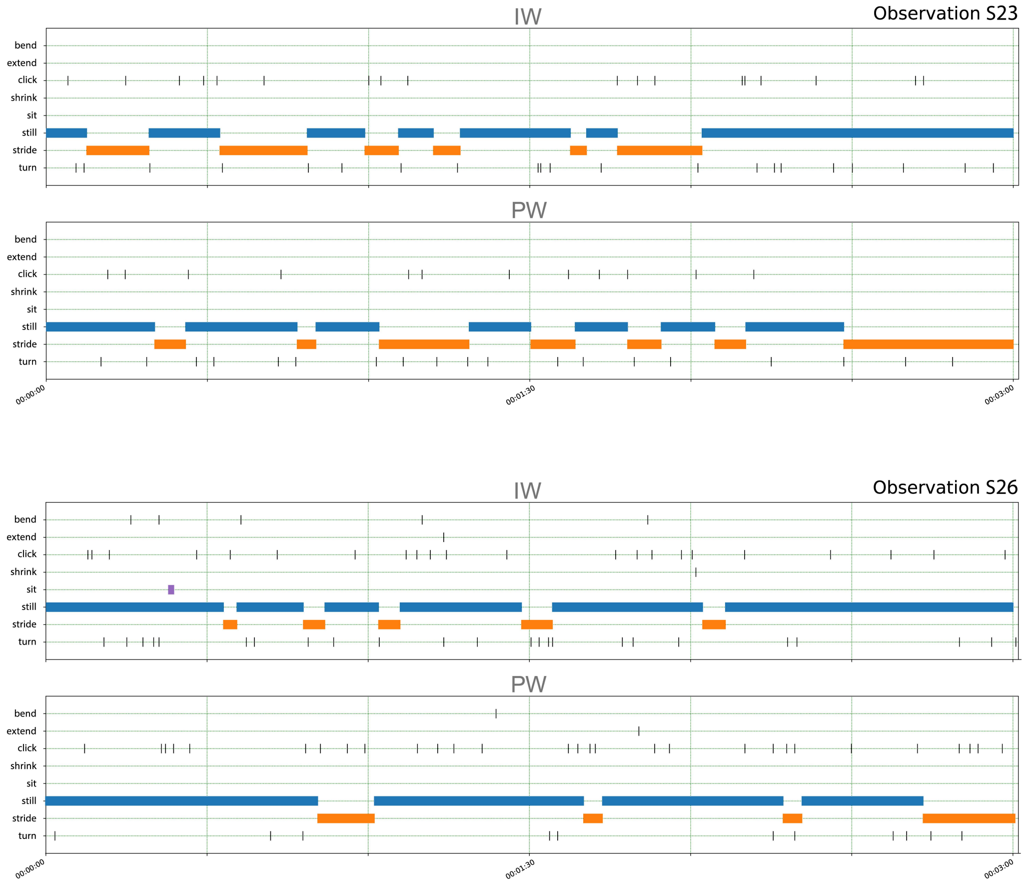


Figure 3.11: Ethogram showing event log for S23 and S26 under PW & IW. The X-axis shows time, whereas the Y-axis shows behavior.

Subject S08 – Diverse Behavior Patterns

S08 exhibited a diverse range of behaviors across both conditions, maintaining a relatively balanced still-to-stride state ratio (See Fig 3.12 and Table 3.3). This participant's expressive movements contrasted with the more restrained behavior of S14. Fig 3.13 and Fig 3.14 compare their actions presented as state and transition frequencies to illustrate their overt actions. With S08, we can see 31 total transitions between each respective durational state and all five non-durational states. For example, the subject transitioned at the same frequency (0.058) from a still position to a stride, then a stride and point, and back to being still again. In contrast, S14 made 13 total transitions between the two durational states to only two non-durational actions.

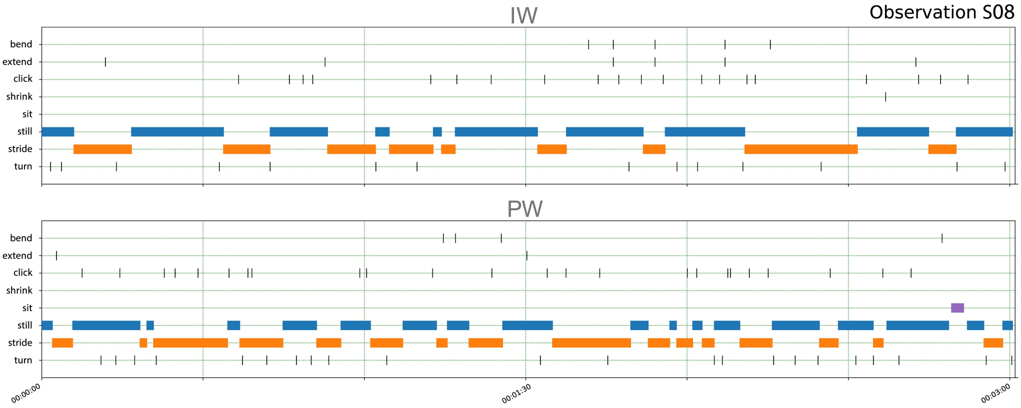


Figure 3.12: Ethogram showing event log for S08 under PW & IW. The X-axis shows time, whereas the Y-axis shows behavior.

Negative Effects (NE) and Activity

Notably, S08 consistently reported higher values for adverse effects in both scenarios, NE(PW: 3.6, IW: 3.2). Interestingly, this trend was also observed in S14, who reported higher NE for PW despite displaying more active behavior in IW. This suggests that negative feelings within the virtual environment cannot be solely attributed to increased physical activity. Other factors, such as the design of the virtual environment or individual differences in susceptibility to VR-induced discomfort, may contribute to the reported adverse effects.

3.6 Summary

This chapter focuses on the results of publications **V** and **VI**, which examine the effects of perceived actions within VEs on the overt behavior of users on the outside. Perceived actions refer to the users' perception of stimuli and cues within the VE and the subjective belief in their ability to engage and influence the contents of the VE actively. The chapter emphasizes the importance of understanding users' perception of action in VEs and proposes the examination of their physical movements outside as a manifestation of that perception. An empirical study examined the relationship between perceived action, user intentions, and observed physical movements, revealing insights into the impact on user experience. The study successfully applied an observational methodology to understand user-centric factors within the context of QoE for immersive media.

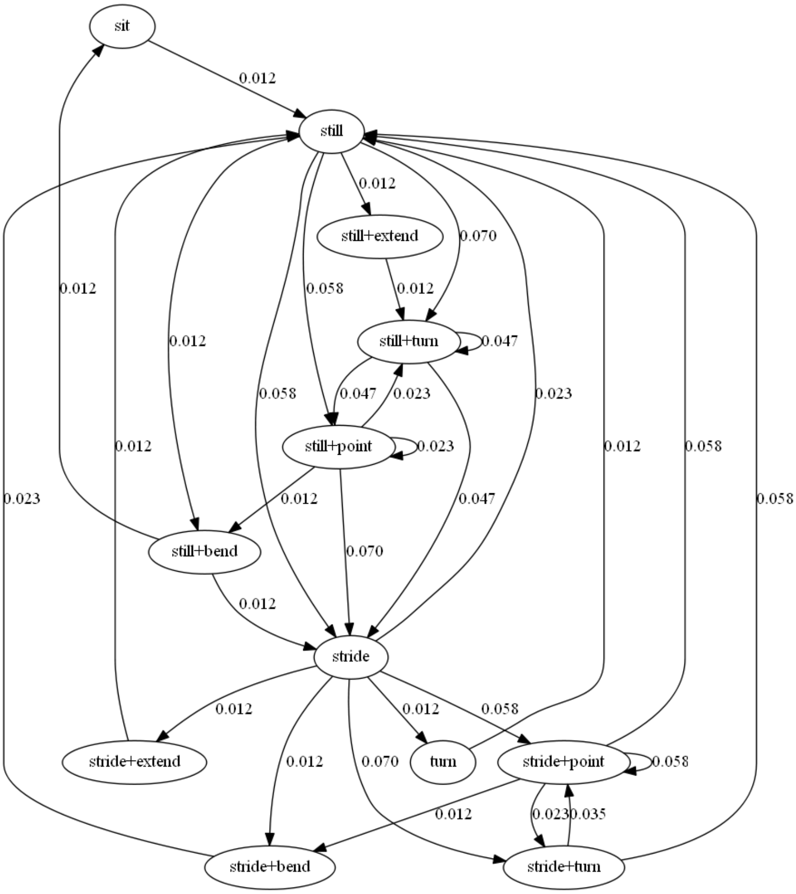


Figure 3.13: Event transitions for S8 in PW.

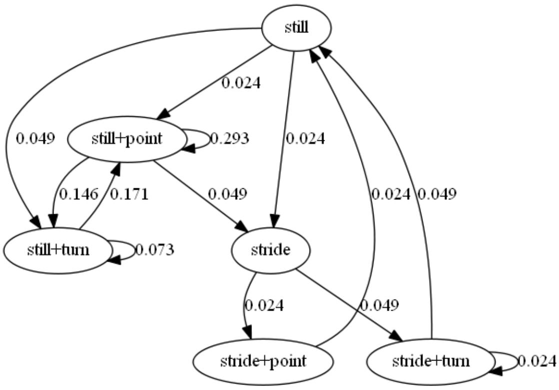


Figure 3.14: Event transitions for S14 in PW.

Chapter 4

Input Modality, Motor-Tasks, and User Performance in VR

This chapter presents the results of publication **VII** and its extension **VIII**, which concentrate on objectives **O3**, that is, on assessing the comparability of the VR interaction quality to the action possibilities on offer inside the VE.

The mass accessibility of VR has made technologies like instinctive interactions and direct handling of entities familiar at the consumer level. From a behavioral perspective, VR symbolizes an advanced interface between humans and computers where users can express themselves directly and freely within computer-created surroundings [28]. Users shift from passive spectators to active participants. Such technological advancements are akin to establishing experiential realities [17] [9]. Until recently, handheld controllers and head-mounted displays (HMDs) made up the predominant user-interaction paradigm for VR. However, with the recent improvements in consumer-grade VR devices, non-mediated natural interactions have emerged as a potential solution to augment realistic interactions in VEs [19]. The potential for this can be leveraged in a variety of applications. More notably, for simulation-based training across diverse fields such as motor learning, gaming, mining, and surgery [186] [187] [23] [188] [189].

In this regard, hand-tracking (or virtual hands) promises a naturalistic gesture-based modality closely mimicking real-world actions [190]. Hand-tracking solutions are, therefore, a promising step forward if a surge in public adoption of VR is intended primarily for entertainment and educational purposes. Devices like the Oculus Quest HMD, which support both handheld controllers and hand-tracking modalities, offer an opportunity to examine the implications of these interaction modes on user experience [191] [192]. But how well advances in modality types match the complexity and dexterity demands of content remains to be seen.

In this context, our integrated studies sought a careful comparison of two input modalities within VR. The primary focus was to assess their impact on motor tasks, such as a reach-pick-place operation. Both publications **VII** and **VIII** converge on the critical understanding that while hand-tracking presents an opportunity for more naturalistic interaction within VR environments, its impact on user experience, particularly in terms of cognitive workload and overall usability, needs careful consideration.

In publication **VII** we apply a performance assessment methodology based on in-game analytics to evaluate the two modalities – are virtual hands better?. In **VIII**, we corroborate this examination against the user-perceived mental workloads reported for each modality. We also evaluate the ease of use and desirability of both modalities and determine which contributed more to the subjective experience of presence.

4.1 Related Works

4.1.1 Authentic Interactions in VR

Interactivity within VR environments entails user responses and modifications to the components of the environment [193]. These changes rely on three essential elements: input devices that capture user actions (input modality); display devices that present the effect of these actions back to the user (HMDs); and transfer functions that map device movements into movements of controlled display or interface elements (mapping method) [193].

There is a continuum for our interactions within VEs, with one end signifying realistic interactions that faithfully mimic real-world engagements and the other showcasing nonrealistic or symbolic interactions [193]. The type of interaction employed depends on the task or challenge at hand [194]. Interaction fidelity is the extent to which real-world interactions can be replicated within VEs, is essential in this context [194]. It can also be comprehended as the level of congruence between the actions performed in a virtual task and those needed for a similar real-world task [2]. The degree of perceived authenticity (its likeness to the real world) is influenced by its interaction fidelity [193], which gives credence to the environment's subjective sense of reality and overall credibility [195].

The perceived authenticity of a VR world hinges on three factors: (1) the sensory experiences provided by the environment, (2) the credibility and plausibility of the scene or situation, and (3) the intuitiveness and ease of interaction within the VE [195]. User expectations towards realism may vary depending on the presented content. For instance, a fictitious VR experience might evoke different expectations compared to a VR learning environment [196]. The latter requires a high degree of interaction fidelity due to the need for precision and dexterity, and users might evaluate elements like the size and proportion of their virtual body about their real body [196] [182]. While a higher interaction fidelity is expected to enhance training efficiency, the required level can vary based on the task [197]. For instance, a high-fidelity interaction technique was suitable for a Virtual Biopsy Trainer. At the same time, a low-fidelity system was sufficient for a Laparoscopic Surgical Trainer [198] [199].

Interaction fidelity becomes a yardstick for how well a specific task feature (in terms of the action it requires) is aligned to the device and interface available (appropriateness and mapping) and if both the latter correspond to the skill and literacy of the user [194]. The perceived quality of the interaction depends on the congruence between the visually available virtual object, the interaction technique, and the respective system features that facilitate it. Expectations based on previous modalities can influence perceptual responses, and any incongruence between these expectations and the experience can impact perceived quality and trigger a range of user reactions, from surprise and delight to disapproval and disappointment [195].

4.1.2 Making Interactions Natural

Interactions within VR can be broken down into *selection* and *manipulation*. Entities in the VR environment can be selected through controller inputs, gestures, or gaze, and subsequently manipulated through resizing, reorientation, scaling, rotating, or moving [138]. VR systems deploy either direct or indirect methods of interaction. Indirect interaction implies interacting with virtual objects via a proxy like a controller and using symbolic references to establish a connection with the virtual entities, such as pressing a button to shift a box [138] [193]. Conversely, direct interaction methods use our bodies to engage with virtual entities directly. For instance, a virtual object can be grasped with virtual hands when it is within reach. When the object is too far for direct grasping, ray-casting or gaze-based approaches can be used.

Modern VR systems increasingly adopt interface paradigms that accommodate natural user interactions (NUI), supporting direct object manipulation through gestures or body movements without intermediary devices [200]. Theoretically, direct interactions foster a greater sense of psychological closeness or proximity due to using action cues, enhancing a user's perceived authenticity of the VE. That being said, direct manipulation methods are not without limitations. They can be impractical due to constraints on input devices like limited tracking range or limitations of the users such as physical challenges. Generally, the method that best matches the interaction modality and meets the task's requirements tends to yield superior performance. The effectiveness of interaction improves when the selection and manipulation methods of a virtual interface align with the system's speed, range, and tracking capabilities and are complemented by the user's understanding, skill, and learning capacity [201] [202]. Unrealistically complex interactions or mismatches in expectations can lead to adaptation problems and negative implications for human factors [203] [204]. However, whether this improves user experience is yet to be fully explored. For example, Voigt-Antons et al. [123] found that even when participants felt they had less control with hand-tracking, they still reported a positive user experience. Conversely, in the study by Masurovsky et al., [191], lower performance metrics were counterbalanced by high subjective scores given by users for controller-free interaction.

4.1.3 Performing Motor-Tasks in VR

Prehension, the capability to reach out and grasp an object, is an essential motor ability that allows us to handle items in our surroundings. It demands the synchronization of multiple muscles and joints in the hand, arm, and shoulder to arrange the hand and fingers around an object to secure it firmly [205]. It's vital for executing everyday tasks like using utensils, picking up and moving items, or handling tools and other objects. Over time, numerous categorizations of human grasp types have been proposed [206] [207] [208] [209]. These taxonomies divide grasp types into main categories like power, precision, and intermediate types. Specific subcategories are defined based on thumb positions and finger/palm contacts. Essentially, different prehension types depend on the distinct movement and positioning of the hand and fingers. In VR, prehension can refer to the user's ability to select and manipulate virtual objects using either:

- Handheld controllers that emulate the user's hand movement and position in the virtual environment (VE). These controllers may include buttons, triggers, and other inputs that allow the user to carry out a variety of actions, such as grabbing, releasing, and manipulating virtual objects or

- Hand gestures, also known as hand-tracking, enable the VR system to identify and interpret the user's hand movement and position in the physical world and project them into corresponding hand gestures in the VE. This can be accomplished using sensors on the user's hands or using cameras and other sensors to monitor the movement of the user's hands and fingers.

In theory, both these techniques can augment immersion in a VE and improve a user's perceived authenticity of the virtual world. However, there are apparent limitations. For instance, handheld controllers are technically exceptional as they facilitate real-time tracking of hand location in virtual space, but they still necessitate an intermediary device. Conversely, hand-tracking offers a direct approach but does not fully enable interaction with virtual objects. Despite being able to grasp them, squeezing or lifting them is not feasible because virtual objects lack weight, volume, or texture.

Therefore, it is meaningful to understand prehension in VR, given its applicability in an array of tasks, activities, and scenarios suitable for simulation-based virtual training, assembly, prototyping, etc. The majority of research concerning prehension in VR is focused on areas of motor therapy [210] [211] [212] and rehabilitation [213] [214] [215]. In a study involving 13 participants, Furmanek et al. [216] contrasted reach-to-grasp movement patterns within VE to those executed in the physical environment. The comparison was based on established kinematic variables and conducted in three stages: initiation, shaping, and closure. The study found that user performances were comparable in both settings, except for differences identified in the closure stage, which was extended in VE. In another experiment, participants carried out a reach-and-grasp task under monocular, motion parallax, and binocular viewing conditions using a telepresence system [217]. Although the study utilized a prehension-based activity, it focused only on depth and distance assessments.

4.1.4 Measuring Task Performance, Task loads, and Feelings

User performance and user experience in VR are intricately tied and mutually reinforcing. Users who can interact proficiently, execute actions, and accomplish their objectives will likely have a positive experience. Conversely, more engaged, engrossed, and content users are likely to perform more effectively. The possibility of action is generally perceived to affect VR experiences [9] [27]. Naturally, a genuine VR experience (realism) may captivate users. Still, their interest and motivation are contingent on the engagement/challenge provided by the content/task (involvement) and how effortlessly and instinctively the system/interface enables them to execute it (usability).

We consider users' performance metrics in this study, derived from game-based quantitative data (see 1.5.3), which is regarded as a good measure for evaluations [46]. Buttussi et al. [218] used the IPQ to examine the effects of three locomotion techniques (joystick, teleportation, and leaning) on participants. The IPQ is a reliable measure for determining the perceived interaction fidelity and visual render quality experienced by users [219] [220]. For cognitive and performance workloads, the NASA-TLX has been extensively used in VR based training material and learning [221] [45] [222]. Separately, [223] [224] applied the AttrakDiff to assess how the usability and desirability of the interface influenced tool effectiveness in VR learning applications.

A detailed description of the measures and tools applied in this study is available in Section 1.5 of this document.

4.2 Motivation

The transition from controller-based to hand-tracking technology in VR raises several interesting questions regarding their implications on user performance, perceived workloads, and overall QoE. While hand-tracking technology promises increased naturalness of interactions, it remains unclear if the transition to technology would be as natural. For instance, given their closer resemblance to real-world interaction dynamics, would virtual hands facilitate more efficient task execution than handheld controllers?

Furthermore, the relationship between interaction realism (the degree to which device interaction simulates real-world interaction) and the experience of naturalness is not linear, suggesting that increasing realism does not always translate to enhanced user experience. This highlights the need for empirical research to understand the implications of each modality. There is ground for investigating the comparative performance of handheld controllers and hand-tracking technology in executing basic reach-and-grasp tasks within a VE and its impact on user experience.

4.3 Methodology

We looked at the impact of input modality on user experience and performance in VR simulation-based reach-grab-place tasks. In addition, the study also considered whether visual factors (e.g., alterations in environmental color) impact users' perceptions while performing the task, as previously noted (Felton, 2021) (Billger et al., 2004). We assessed the collected results about users' subjective experience (e.g., sense of presence, perceived cognitive workload, and ease of use) and their game-based performance metrics.

4.3.1 Design

We used a repeated-measures 2 X 2 design to contrast two input modalities among participants executing a reach-grab-place task inside a VE displayed at two levels of representational realism. The input modality type (M) and representation level

(R) form our two independent variables (IV). Each had two levels: M, hand-tracking (M1) and handheld controller (M2); and R, saturated (R1) and grayscaled (R2) variants. We arrive at four variations, as seen in Fig 4.1:

- M1R1 - hand-tracking in a color VE
- M1R2 - hand-tracking in a grayscaled VE
- M2R1 - handheld-controller in a saturated VE
- M2R2 - handheld-controller in a grayscaled VE

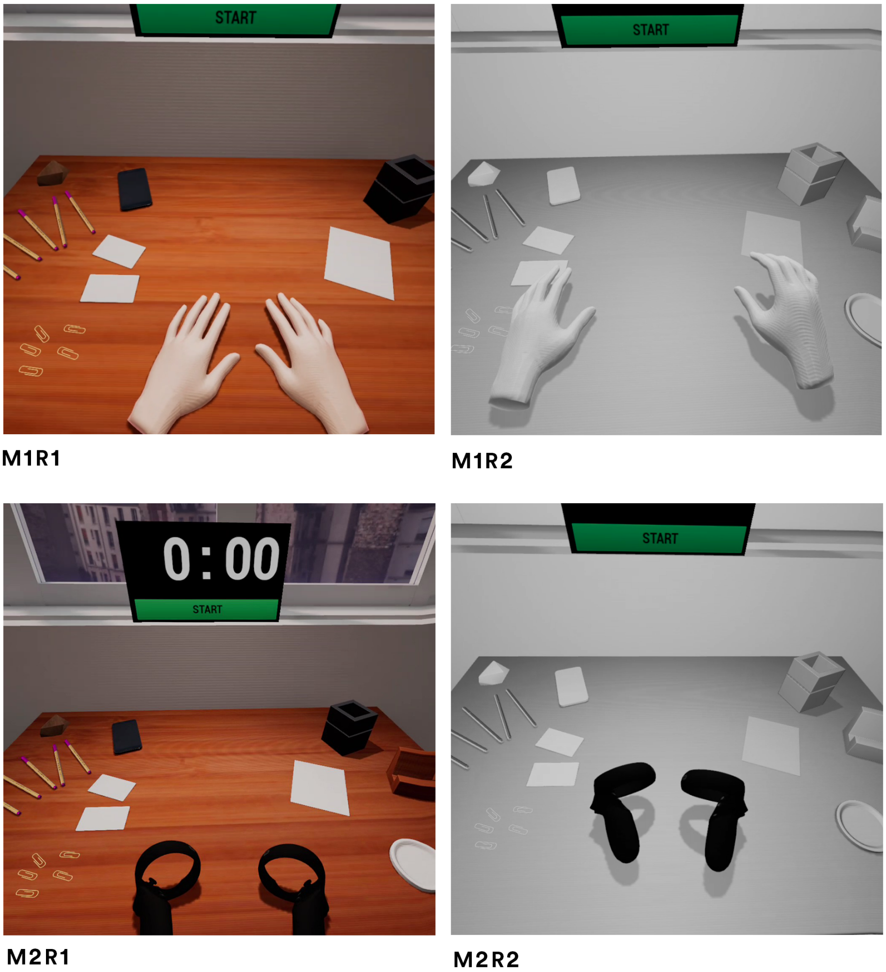


Figure 4.1: Four variations based on Modality Types (M1, M2) and Representation Levels (R1, R2).

Dependent variables (DV) included user performance metrics and scores for user-perceived overall presence, mental workload, and usability derived from self-reported questionnaires.

4.3.2 Participants

Our sample comprised a diverse demographic of N=33 participants (15 males, 18 females, average age=24.7 ±2.3). Participants were sourced from an online volunteer portal hosted by the university, which was accessible to the public at large. The study attracted individuals from various backgrounds who were given gift cards as a token of appreciation for their participation. The study occurred at the university's VR/XR labs, as shown in Fig 4.2. At the beginning of the test, participants provided their demographic information and their familiarity or inclination toward immersive technology through a pre-study survey administered via Google Forms. Regarding VR experience, the majority of participants (n = 22) reported having "no" experience; a few (n = 6) disclosed "some" prior experience; others (n = 4) had "intermediate/moderate" experience, and only one participant (n = 1) claimed to have "good" VR experience. Similarly, only one participant had

Figure 4.2: A participant using hand-controller modality interacts with the environment displayed inside the HMD.

participated in a VR lab study before. All participants (n = 32) reported having at least "some" experience with video game controllers, and a single participant (n = 1) had previously used hand-tracking. All participants had either normal visual acuity or vision corrected to normal.

4.3.3 Stimuli

The Virtual Environment (VE) included a virtual room with a table before a window. Participants were asked to execute a simple reach-grab-place task at this virtual table. The VE was designed with a 1:1 ratio to match the exact dimensions of the physical laboratory in which the experiment was conducted, measuring 5.4 m by 4.4 m.

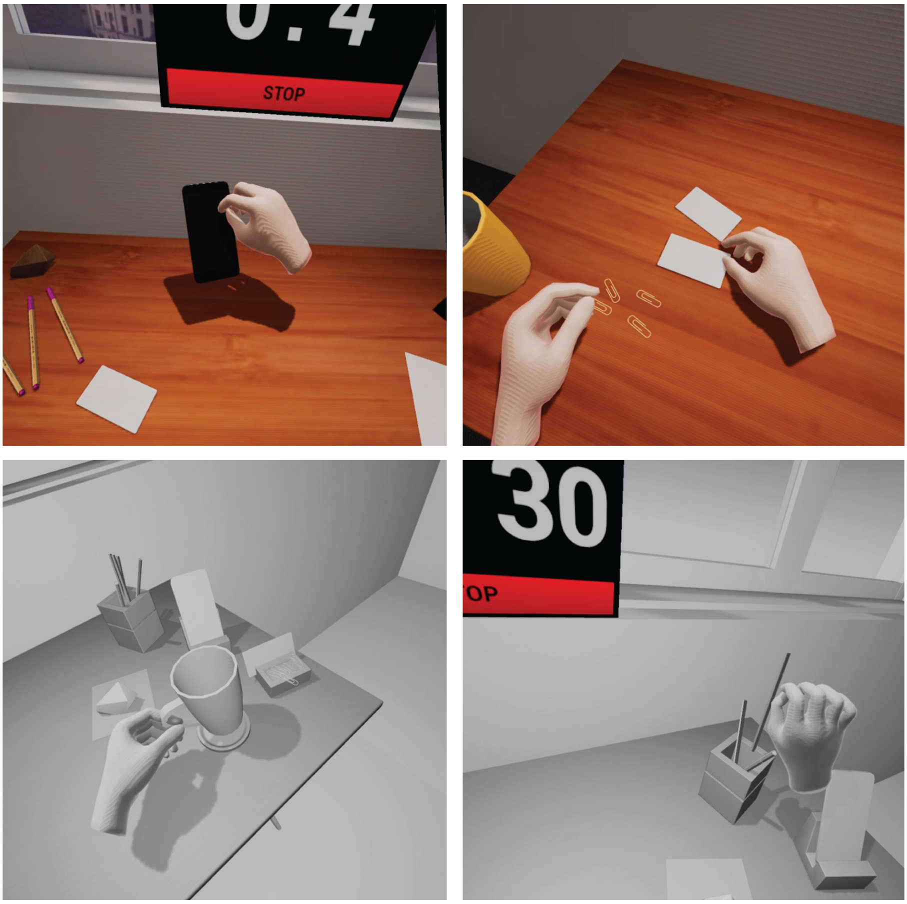


Figure 4.3: Top Left: mobile phone grabbed with virtual hands in M1R1. Top Right: user picks a business card with virtual hands in M1R1. Bottom Left: user places the coffee mug with virtual hands in M1R2. Bottom Right: A pen is placed using virtual hands in M1R2.

Figure 4.4: Top Left: coffee mug placed on the coaster using controllers in M2R1. Top Right: pen placed in the holder using controllers in M2R1. Bottom Left: the paper clip is picked using controllers in M2R2. Bottom Right: paper weight placed on paper in M2R2.

This model was created using Sketchup Pro (Version 20.2.172). The door and windows in the VE were designed to mirror those in the physical lab and the furniture placement and orientation aligned with the physical room. The virtual table was superimposed on the physical table, and their heights were aligned to give participants a tangible surface beneath their arms. At the same time, they were engaged at the virtual table. The textures, lighting, interactive features, and game-based elements were implemented using Unreal Engine (Version 4.26).

4.3.4 Task

The task in VR entailed participants reordering a variety of 15 items on a virtual table. Each variant would start with all 15 items on the virtual table's left side. Participants were then asked to pick up and relocate each of the 15 items individually to their assigned positions on the right side. The participants needed to place the objects accurately at their intended locations. Participants were allowed to adjust the items' orientation for convenience and comfort. Screenshots of both modalities at work can be seen in Fig 4.3 and Fig 4.4. The task involved grabbing and moving the following objects:

- 5 pens into a pen holder
- 5 paperclips into a clip saucer
- 2 business cards into a cardholder
- 1 mobile phone into a phone holder
- 1 paperweight onto a stack of paper
- 1 coffee mug onto a coaster

The six objects in the task each require a specific grip type (see Fig 4.5), outlined below [4] [209]:

a. *Terminal opposition* (Fig 4.5A) is a precision grip that enables the handling of thin or tiny objects like a needle or paperclip. The thumb and the tip of the index (or middle) finger meet when grasping delicate objects.

b. *Tetradigital grip* (Fig 4.5B) is used to hold larger objects firmly. This grip, involving pulp contact, can hold items such as a pencil, brush, or pen. The thumb's pulp firmly presses the object against the pulps of the index, middle, and ring fingers.

c. *Subterminal opposition* (Fig 4.5C) is a grip where the sides of the fingers and thumb make contact with the object instead of the fingertips or palms. It is commonly used for holding thin, elongated objects like paper.

d. *Tridigital grip* (Fig 4.5D) involves the thumb, index, and middle finger and is often used to bring food to the mouth. Subterminal tridigital prehension is used for small round or irregular objects.

e. *Panoramic pentadigital grip* (Fig 4.5E) is used to hold large, flat objects with the fingers widely spread out and the thumb positioned in maximal counter opposition. This grip securely holds such objects.

f. *Three-finger pinch-dorsal contact grip* (Fig 4.5F) is used to hold objects like cups. It involves placing the pad of the index finger at the middle phalanx level and the middle finger's radial side on the cup for balance and support.

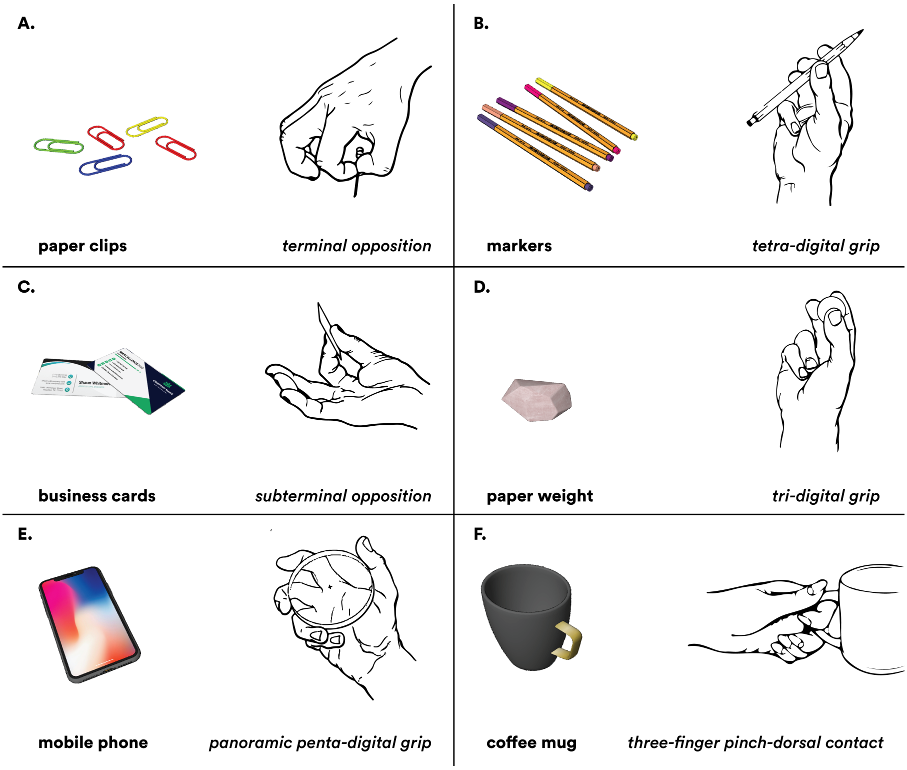


Figure 4.5: Six objects used in the reach-grab-place task. Each object required a specific static grip to be handled. (prehension illustrations by I.A. Kapandji [4])

A virtual timer tracked the time for each task. The timer could be halted if the user became bored, frustrated, or annoyed with the input modality or the task. In these instances, the result would be recorded as "incomplete."

4.3.5 Premises & Apparatus

The VR game was specifically designed for compatibility with the Oculus Quest VR headset (Oculus VR, Inc., 2020). This standalone device can wirelessly operate games and software on an Android-based OS. The headset offers a stereo viewing experience with an OLED display for each eye, each with a resolution of 1440 x 1600 and a refresh rate of 72 Hz. The headset also includes internal cameras that provide inside-out positional tracking, enabling six degrees of freedom (6DOF). The Oculus Quest operates with the Oculus Touch (hand controller) and supports controller-free gestures. The usage of both modalities is illustrated in Fig 4.6.

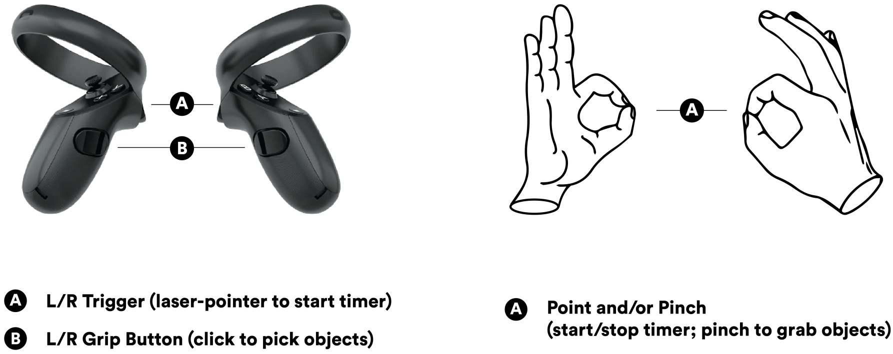


Figure 4.6: Two input modalities available on the Oculus Quest (Oculus Quest © Oculus VR, Inc.)

The Oculus hand-tracking system interprets distinct hand postures and tracks specific points on the hands, such as knuckles or fingertips, in real time. There is no officially published accuracy rate for the Oculus Quest. An analysis [225] found that the hand-tracking system had an average fingertip positional error of 1.1 cm, a finger joint angle error of 9.6 degrees, and a temporal delay of 38.0 ms.

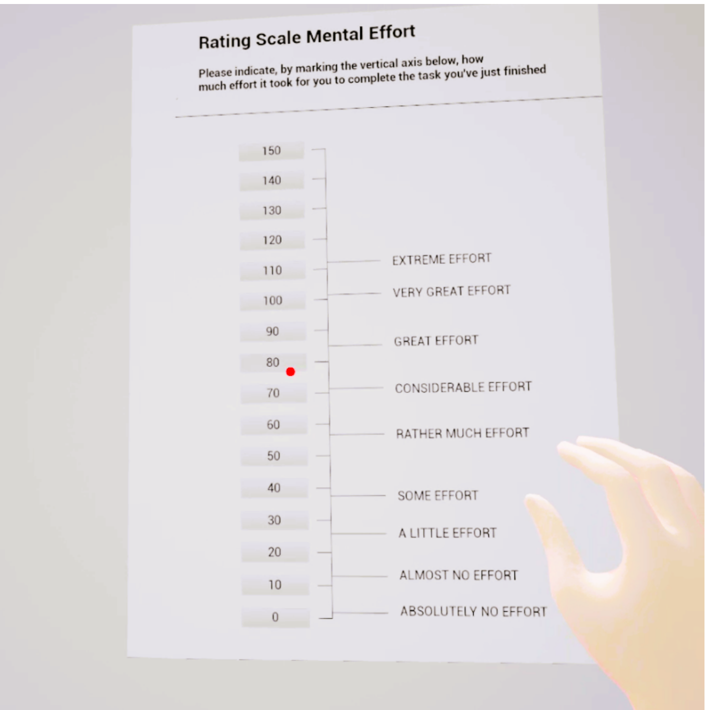
Regarding the handheld controller, participants could press the left or right grip button to lift any of the six objects. An object could be picked up by bringing the controller close. As for hand-tracking, the device recognizes basic gestures, including [226]:

- Pointing - users extend their index finger forward, which the system interprets as a selection gesture.
- Pinching - when users touch their thumb and index finger together, the system recognizes this as a gesture to grab items.

4.3.6 Procedure

Upon arrival, participants received an introduction kit, including a consent form. Seated at a table, the moderator briefed them on their task. The physical table was mapped 1:1 to the virtual table in the game. To familiarize themselves with the task, participants initially performed a physical version. This way, when performing the task within the virtual game, they could concentrate on the method of execution rather than the organizational aspect of the task. Each participant experienced all four game variants randomly to avoid sequence effects.

They were also instructed on how to use the two interaction techniques. Once they expressed confidence with the provided information, the experiment began. An



Rating Scale Mental Effort

Please indicate, by marking the vertical axis below, how much effort it took for you to complete the task you've just finished

150
140
130
120
110 EXTREME EFFORT
100 VERY GREAT EFFORT
90 GREAT EFFORT
80
70 CONSIDERABLE EFFORT
60 RATHER MUCH EFFORT
50
40 SOME EFFORT
30 A LITTLE EFFORT
20 ALMOST NO EFFORT
10
0 ABSOLUTELY NO EFFORT

Figure 4.7: RSME scale applied inside the VR environment.

average experiment session lasted 60 minutes, divided into four game trials. Each game variation followed a 5-7 minute break to allow participants to adjust from the virtual environment and prevent eye strain. Upon donning the HMD, the sequence for each game was as follows:

1. The task began by pressing the "start" button on a virtual timer.

2. They then proceeded to carry out the task of placing the 15 virtual objects in their designated positions.

3. After completing the task, they pressed the "stop" button on the timer.

4. Upon stopping the timer, a VR-based RSME scale appeared, where participants rated their perceived mental effort for the completed task (see Fig.3).

5. The Participant would take off the HMD, rest a few minutes, and then fill out the post-game questionnaires.

4.3.7 Materials

Performance Metrics

These are indicators identified in user behavior patterns and trends during task performance. User performance was assessed based on a game log created at the end of each task, which included:

1. Active-Time (AT): duration a user took to finish the task.
2. Object-Pick (OP): count objects picked up during the task.
3. Right-Click Frequency (RC): a tally of right-clicks made to grab the objects.
4. Left-Click Frequency (LC): a tally of left-clicks made to grab the objects.

**RC and LC are combined and denoted under Click Frequency (CF)*

Igroup Presence Questionnaire (IPQ)

The 14 items of the IPQ were rated on a seven-point Likert scale. The three subscales measured:

1. Spatial Presence (SP): sensation of physically existing in the VE
2. Involvement (INV): to gauge the attention given to the VE
3. Realness (REAL): the sense of reality attributed to the VE

Cognitive Load (NASA-TLX)

The index measured performance across six dimensions to establish an overall workload score:

1. Mental Demand (MD): cognitive, decision-making, or calculation needs
2. Physical Demand (PD): quantity and intensity of physical activity
3. Temporal Demand (TD): time pressure associated with task completion.
4. Effort (ED): difficulty of maintaining performance
5. Performance (RD): success level in task completion
6. Frustration Level (FD): feelings of security/insecurity or discouragement/contentment

UX Evaluation (AttrakDiff)

The AttrakDiff assesses UX-related quality perceptions under:

1. Pragmatic Quality (PQ): This describes a product's usability and the success users have in achieving their goals with it.
2. Hedonic Quality (HQ): This measures emotional responses of how stimulating/inspiring a product is and if users identify with it.
3. Attractiveness: This describes the overall value of the product based on quality perception, whether positive or negative.

4.4 Results

Our research incorporated a single virtual task (reach-pick-place) executed under four unique conditions. Each condition was a separate virtual game. We recorded performance metrics and subjective scores in the measures described earlier for every game. To investigate the impact of Modality (M) and Representation (R) on user performance, we conducted a two-way multivariate analysis of variance (MANOVA). The levels of the two categorical conditions, M (M1 & M2) and R (R1 & R2) were scrutinized to identify any significant distinction between the two unequivocal conditions when evaluated collectively on the the various dependent variables (DV) that included quantitative game metrics (x 3 DVs) and the different subscales for IPQ (x 4 DVs), NASA-TLX (x 6 DVs), and AttrakDiff (x 3 DVs).

Table 4.1: The mean and standard deviation for the three dependent variables, split by the independent variables for all *N=32* participants. Both original and log-adjusted values are shown.

Item	Condition	Mean(µ)		Std.Dev (SD.)		N
		org	*log*	*org*	*log*	
AT (sec)	M1R1	314s	2.45	169.43	0.21	32
	M2R1	152s	2.14	67.52	0.19	32
	M1R2	272s	2.40	116.58	0.19	32
	M2R2	154s	2.12	104.12	0.24	32
OP (no.)	M1R1	32	1.44	17.04	0.25	32
	M2R1	21	1.26	9.20	0.31	32
	M1R2	29	1.43	11.82	0.17	32
	M2R2	18	1.20	8.30	0.28	32
CF (no.)	M1R1	399	2.54	243.65	0.23	32
	M2R1	58	1.70	37.15	0.23	32
	M1R2	371	2.52	192.1	0.20	32
	M2R2	66	1.70	65.10	0.29	32

4.4.1 Performance Analysis

A two-way MANOVA examined the effects of M and R on user performance measures: AT, OP, and CF. The dataset for AT and OP had a positively skewed distribution, so logarithmic transformation was applied uniformly. Results of Box's M Test (Equality of Covariance Matrices) showed no violation of the assumption (p = 0.26). Significance values for Levene's Test (Equality of Error Variances) (P_{AT} = 0.73, OP_{OP} = 0.85, and P_{CF} = 0.37) exceeded 0.05.

MANOVA revealed a significant main effect of modality type, $F(3,\ 122) = 195.8$, p = $<$ 0.001, Wilk's $\lambda = 0.172$, $\eta^2 p = 0.83$. There was no significant main effect of representation level, $F(3,\ 122) = 0.85$, p = 0.47, Wilk's $\lambda = 0.98$, $\eta^2 p = 0.02$. Nor was there a significant interaction effect between modality and representation, $F(3,\ 122) = 0.20$, p = 0.90, Wilk's $\lambda = 0.99$, $\eta^2 p = 0.005$. Follow-up univariate analyses indicated that modality type produced significant differences across all three performance scores,

- AT: $F(1,\ 124) = 62.74$, p = $<$ 0.001, $\eta^2 p = 0.34$
- OP: $F(1,\ 124) = 20.1$, p = $<$ 0.001, $\eta^2 p = 0.14$
- CF: $F(1,\ 124) = 374.8$, p = $<$ 0.001, $\eta^2 p = 0.75$

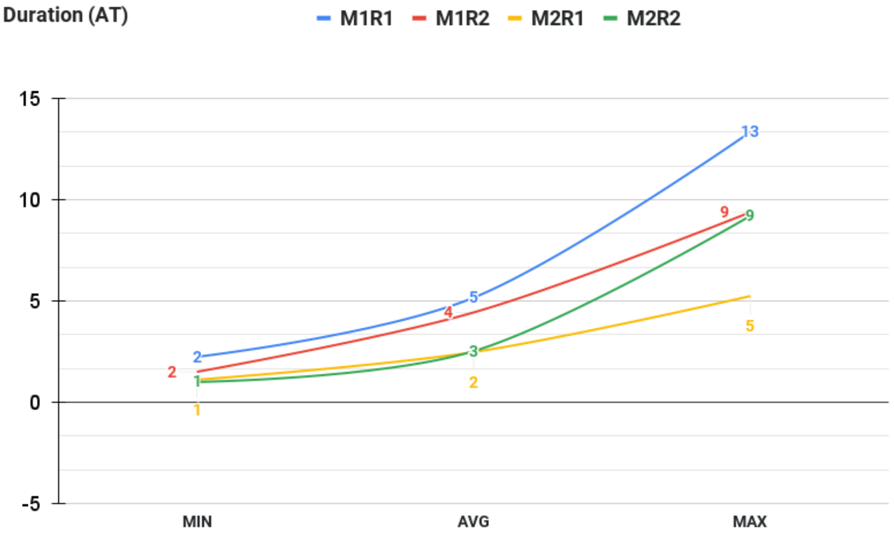


Figure 4.8: Min & Max plot for AT under the four conditions. The Y-axis shows the duration in minutes.

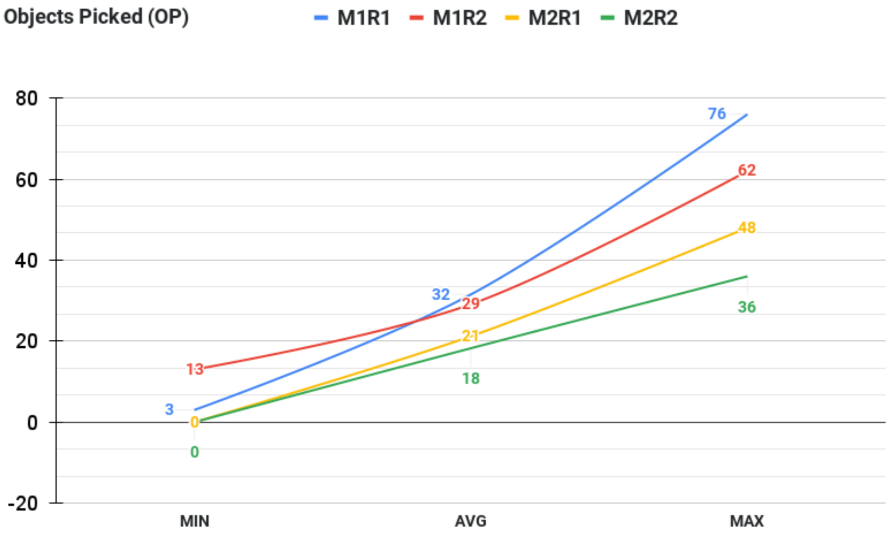


Figure 4.9: Min & Max plot for OP under the four conditions. The Y-axis shows a number of objects picked.

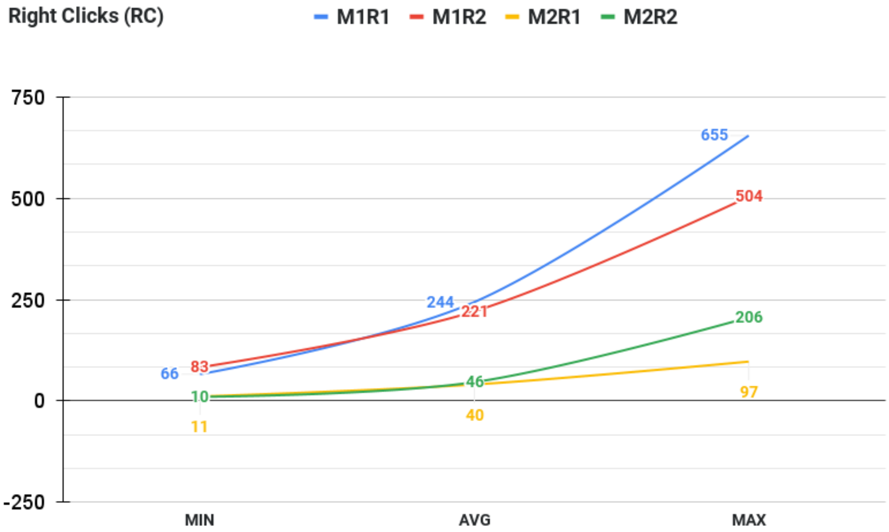


Figure 4.10: Min & Max plot for RC under the four conditions. The Y-axis shows a number of right clicks.

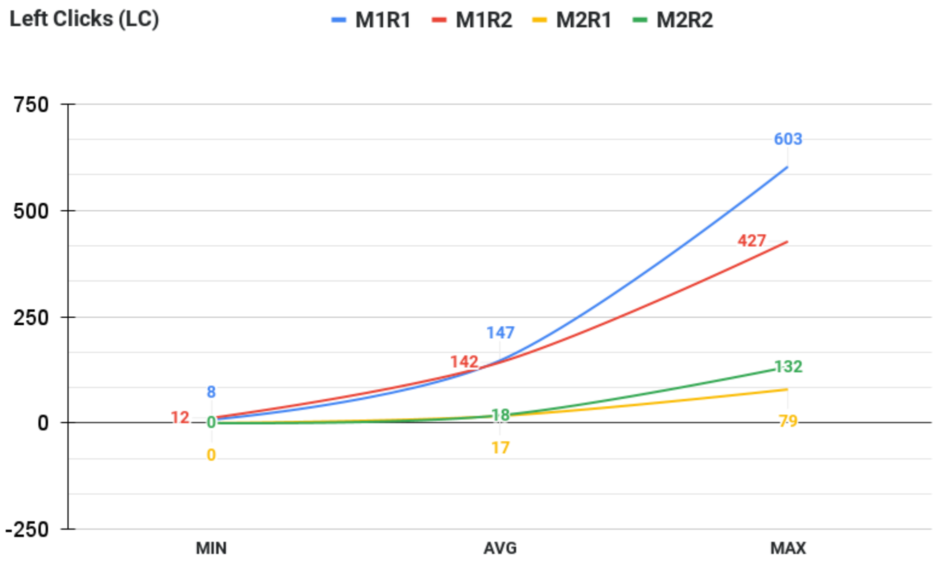


Figure 4.11: Min & Max plot for LC under the four conditions. The Y-axis shows a number of left clicks.

Descriptive statistics for the dependent variables, presented in Table 4.1, shows the number of observations(N), the mean(μ), and the standard deviation(SD.) in each group. The table shows both the log-transformed and the original data. As shown (Fig 4.8), users using hand-tracking took longer to complete the task (μ= 314s, SD. = 169.42; μ= 272s, SD. = 116.58) compared to those using handheld-controllers (μ= 152s, SD. = 67.52; μ= 154s, SD. = 104.12). There was no significant separation for the mean scores based on levels of representation. The mean score for the number of objects grabbed OP (Fig 4.9), was more in hand-tracking (μ= 32, SD. = 17.04; μ= 29, SD. = 11.82) compared to handheld-controller (μ= 21, SD. = 9.20; μ= 18, SD. = 8.3). Similarly, the number of attempted grabs, CF, was exponentially higher for hand-tracking (μ= 399, SD. = 243.65; μ= 371, SD. = 92.1). It indicated that subjects found it harder to grab objects using hand-tracking. This difficulty, therefore, can be one reasonable explanation for the longer completion times. The difference is pronounced for both Right Clicks (RC) and Left Clicks (LC). Mean scores for RC (Fig 4.10) are exponentially higher in hand-tracking (μ= 244, SD. = 146; μ= 221, SD. = 109) compared to handheld-controller (μ= 40, SD. = 22; μ= 46, SD. = 44). A similar trend is witnessed in the mean scores for LC (Fig 4.11), which remains high in hand-tracking (μ= 147, SD. = 121; μ= 143, SD. = 112) and lower in handheld-controllers (μ= 17, SD. = 21; μ= 18, SD. = 26).

4.4.2 Presence

A two-way MANOVA with a covariate was performed to account for AT, considering the potential influence that the duration spent within the VE could have on the four IPQ items. The maximum Mahalanobis distance value used for assumption testing was 15.55, falling below the critical value of 18.47 (df=4) needed for multivariate normality (refer to Table 2). The results from Box's M Test of Equality of Covariance Matrices demonstrated no violation of the assumption ($\underline{p}$ = 0.59). All significance values ($\underline{P}_{GP}$ = 0.64, $\underline{P}_{SP}$ = 0.59, $\underline{P}_{INV}$ = 0.47, and $\underline{P}_{REAL}$ = 0.55) in Levene's Test of Equality of Error Variances exceeded 0.05. The MANOVA results did not yield any statistically significant effects. There was no difference in the means of the four dependent variables of IPQ (GP, SP, INV, and REAL).

Firstly, there was a non-significant effect of modality type on IPQ scores, M: F(4, 120) = 0.734, $\underline{p}$ = 0.57, Wilk's λ = 0.98, $\eta^2 p$ = 0.024. Secondly, there was a non-significant effect of representation level on IPQ scores, R: F(4, 120) = 0.670, $\underline{p}$ = 0.61, Wilk's λ = 0.98, $\eta^2 p$ = 0.022. Finally, there was a non-significant interaction effect between modality and representation on IPQ scores, M x R: F(4, 120) = 0.158, $\underline{p}$ = 0.96, Wilk's λ = 0.995, $\eta^2 p$ = 0.005.

From our analysis, we fail to reject the null hypothesis, concluding that neither input modality nor visual representation had a significant impact on the user-reported IPQ scores when considered together. Table 4.3 shows Means (µ) and SDs of the IPQ sub-scales. All three subscales and the general presence category show little

Table 4.2: The critical chi-square values for evaluating Mahalanobis Distance at a critical alpha of 0.001 are shown below [5]. Values are shown from 2 to 10 degrees of freedom.

df	critical value	*df*	critical value	*df*	critical value
2	13.82	5	20.52	8	26.13
3	16.27	6	22.46	9	27.88
4	18.47	7	24.32	10	29.59

Table 4.3: The mean(µ) and standard deviation (SD.) for the four different dependent variables of IPQ, which have been split by the independent variables for all *N=32* participants.

	M1R1		*M1R2*		*M2R1*		*M2R2*	
IPQ	**µ**	**SD.**	**µ**	**SD.**	**µ**	**SD.**	**µ**	**SD.**
GP	4.63	1.62	4.31	1.67	4.93	1.54	4.77	1.62
SP	4.21	0.71	4.18	0.77	4.27	0.78	4.20	0.89
INV	4.50	1.37	4.37	1.21	4.81	1.20	4.52	1.23
REAL	3.61	0.65	3.42	0.70	3.49	0.79	3.35	0.63

movement as functions of M-type and R-Level.

4.4.3 Mental Workload

Two separate measures, RSME and NASA-TLX, were used to evaluate MWL for users. The unidimensional RSME measure was completed by participants within the VR environment immediately after they completed their virtual task. The NASA-TLX was administered after the participants had taken off their headsets.

NASA-TLX

A two-way MANOVA analyzed the impacts of modality and representation on users' NASA-TLX scores. To ensure the assumptions for MANOVA were met, we performed a logarithmic transformation so all data exhibited a normal distribution. The Box's M Test of Equality of Covariance Matrices result was $\underline{p} = 0.50$, and the Levene's Test of Equality of Error Variances presented values as $\underline{P}_{MD} = 0.31$, $\underline{P}_{PD} = 0.48$, $\underline{P}_{TD} = 0.53$, $\underline{P}_{RD} = 0.16$), $\underline{P}_{ED} = 0.42$), and $\underline{P}_{FD} = 0.05$.

All four conditions were assessed across the six sub-scales. No significant effects were detected for representation levels ($\underline{p}$ = = 0.99) and interaction ($\underline{p} = 0.66$). The MANOVA showed a significant effect of modality type on the NASA-TLX indices, M: $F(6, 119) = 8.374$, $\underline{p}$ = < 0.001, Wilk's $\lambda = 0.703$, $\eta^2 p = 0.30$. Univariate ANOVAs look at impacts of modality type on each dimension:

- MD: $F(1, 124) = 30.50$, $\underline{p}$ = < 0.001, $\eta^2 p = 0.20$
- PD: $F(1, 124) = 36.41$, $\underline{p}$ = < 0.001, $\eta^2 p = 0.23$
- TD: $F(1, 124) = 23.20$, $\underline{p}$ = < 0.001, $\eta^2 p = 0.16$
- RD: $F(1, 124) = 00.60$, $\underline{p} = 0.44$, $\eta^2 p = 0.005$
- ED: $F(1, 124) = 18.80$, $\underline{p}$ = < 0.001, $\eta^2 p = 0.132$ •

Table 4.4 displays both the log-adjusted and original means(µ) and standard deviations(SD) in each group. Significant differences were observed across all indices except RD ($\underline{p} = 0.44$). These results suggest that the overall perceived workload was higher for handtracking (M1) than hand-controllers (M2). Yet, there was a marginal difference between saturated (R1) and grayscaled (R2) representations. Subjects required less cognitive workload when using the handheld controller to complete the task. The handheld controllers were the least demanding as shown by the means for M2R2 across all indices: MD(µ= 2.38, SD = 1.24), PD(µ= 2.53, SD = 1.08), TD(µ= 2.78, SD = 1.48), RD(µ= 2.84, SD = 1.65), ED(µ= 2.47, SD = 1.63), and FD(µ= 3.06, SD = 1.30).

Table 4.4: The mean and standard deviation for the six different dependent variables of NASA-TLX, which have been split by the independent variables for all *N=32* participants.

Item	Condition	Mean(µ)		Std. Dev (SD.)		N
		org	*log*	*org*	*log*	
MD	M1R1	4.06	1.82	1.86	0.87	32
	M2R1	2.53	1.05	1.65	0.92	32
	M1R2	4.06	1.89	1.52	0.70	32
	M2R2	2.38	1.06	1.24	0.75	32
PD	M1R1	4.06	1.88	1.66	0.69	32
	M2R1	2.59	1.17	1.39	0.79	32
	M1R2	4.19	1.95	1.49	0.60	32
	M2R2	2.53	1.20	1.08	0.66	32
TD	M1R1	4.25	1.92	1.90	0.73	32
	M2R1	2.81	1.257	1.73	0.85	32
	M1R2	4.13	1.91	1.56	0.70	32
	M2R2	2.78	1.27	1.48	0.80	32
RD	M1R1	3.03	1.39	1.51	0.84	32
	M2R1	3.16	1.45	1.51	0.84	32
	M1R2	3.22	1.55	1.31	0.70	32
	M2R2	2.84	1.26	1.65	0.89	32
ED	M1R1	3.60	1.62	1.90	0.85	32
	M2R1	2.75	1.21	1.72	0.85	32
	M1R2	3.97	1.86	1.62	0.62	32
	M2R2	2.47	1.04	1.63	0.88	32
FD	M1R1	4.84	2.15	1.72	0.68	32
	M2R1	3.34	1.53	1.62	0.85	32
	M1R2	4.81	2.18	1.47	0.55	32
	M2R2	3.06	1.47	1.30	0.69	32

Table 4.5: Pearson correlation coefficient (PCC) and regression coefficient (RC) between RSME and the combined NASA-TLX score.

NASA-TLX	RSME	PCC	RC	*p* value
M1				
NASA-TLX	RSME	0.815	0.664	0.000
M2				
NASA-TLX	RSME	0.744	0.554	0.000

RSME

Results from RSME are consistent with those from NASA-TLX. Participants reported a much higher mental workload while using hand-tracking than controllers, regardless of the representation level of the game. The RSME scores decrease significantly between the two modality types. For M1 under R1 & R2, they were µ= 73.33 to µ= 72.42 respectively. And for M1 under R1 & R2, they were µ= 47.0 to µ= 44.24 respectively.

A two-way ANOVA showed a significant impact for only input modality on RSME scores: M: $F(1, 66) = 40.25$, $\underline{p} < 0.0001$, $\eta^2 p = 0.239$. Results for representation level and interaction effects were not significant.

A good correlation was observed between RSME and the combined NASA-TLX score under both modality types. According to the Pearson correlation test and regression coefficient: M1, (r = 0.815) (r2 = 0.664); M2, (r = 0.744) (r2 = 0.554), in Table 4.5.

4.4.4 Usability

The findings from the AttrakDiff were broken down into three facets: pragmatic quality (PQ), hedonic quality (HQ), and attractiveness (ATT). Before conducting a two-way MANOVA, the Mahalanobis distance was examined for assumption testing, with the highest value being 14.95, falling under the critical threshold of 16.27 (df = 3) needed for multivariate normality (see Table 2). The Box's M Test of Equality of Covariance Matrices showed no violation of assumption ($\underline{p} = 0.57$). The Levene's Test of Equality of Error Variances demonstrated significance values greater than 0.05 in all cases ($\underline{P}_{PQ} = 0.16$, $\underline{P}_{HQ} = 0.41$, and $\underline{P}_{ATT} = 0.15$).

No significant primary effects were found for R-Level: $F(3,\ 122) = 1.953$, $\underline{p} = 0.125$, Wilk's $\lambda = 0.95$, $\eta^2 p = 0.05$. Additionally, no interaction effects were detected for (M x R): $F(3,\ 122) = 0.335$, $\underline{p} = 0.80$, Wilk's $\lambda = 0.99$, $\eta^2 p = 0.08$. However, the two-way MANOVA did disclose statistically significant differences for the means of PQ, HQ, and ATT based on M-type: $F(3,\ 122) = 7.953$, $\underline{p} = < 0.001$, Wilk's $\lambda = 0.84$, $\eta^2 p = 0.16$.

Follow-up univariate ANOVAs probed individual differences for M and R on the three dependent variables. Modality type significantly affected usability (PQ: $\underline{p} = < 0.001$) and desirability (ATT: $\underline{p} = 0.033$) of the VR application, whereas representation levels impacted intrigue (HQ: $\underline{p} = 0.025$) and desirability (ATT: $\underline{p} = 0.038$).

Using the online AttrakDiff tool for further examination, we produced "Result Diagrams": portfolio-presentation, average value diagram, and word-pair descrip-

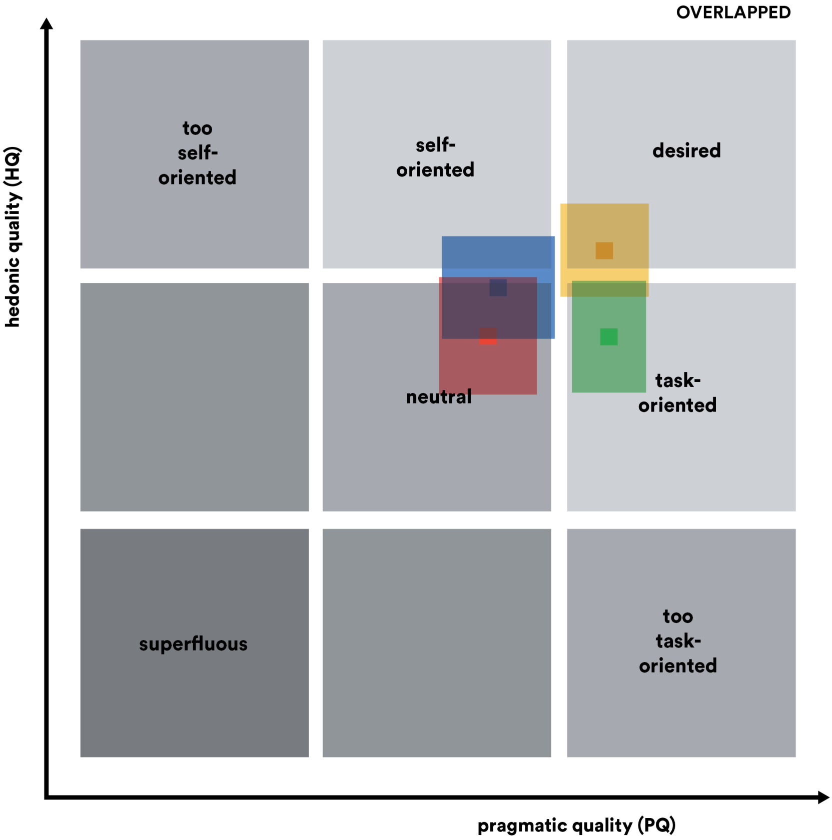


Figure 4.12: An overlapped Portfolio-Presentation for the four experimental conditions. Color legend: M1R1 (blue), M1R2 (red), M2R2 (green), M2R1 (yellow).

tion. Fig 4.12 shows the portfolio-presentation results for the four conditions. M1R1 leans towards "self-oriented" in the "neutral" zone; M1R2 is within "neutral"; M2R2 is "task-oriented"; and M2R1 leans towards "task-oriented" in the "desired" zone. All confidence intervals are comparable in size, indicating participant agreement about the hedonic and pragmatic qualities of the conditions. However, M2R1 has the smallest confidence rectangle, implying higher reliability.

Comparing the portfolio-presentation with the average value diagram (Fig 4.13), M2R1 generally performs better (PQ = 1.40, HQ = 1.24, ATT = 1.85), but M2R2 scores higher in PQ. M1R2 underperforms (PQ = 0.43, HQ = 0.53, ATT = 0.90). The scores for M2R2 range from a high for PQ to a low for HQ (PQ = 1.44, HQ = 0.86, ATT = 1.28). In all four conditions, perceived usability outperforms emotional response.

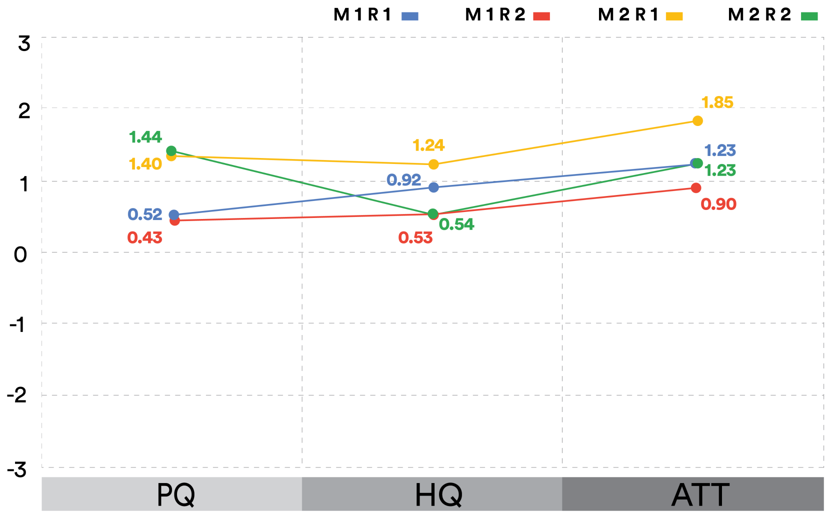


Figure 4.13: Diagram of Average Values for the four experimental conditions. Color legend: M1R1 (blue), M1R2 (red), M2R2 (green), M2R1 (yellow)

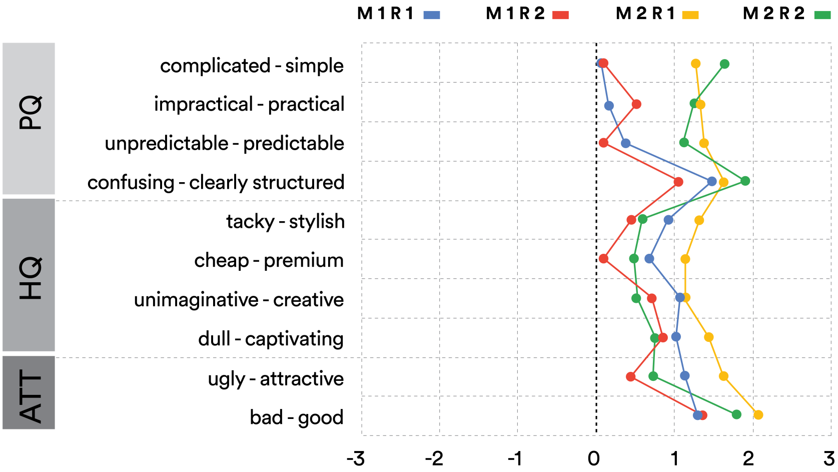


Figure 4.14: Description of Word-Pairs for the four experimental conditions.

Lastly, Fig 4.14 presents the results of the word-pair descriptions, with M2R1 scoring higher in almost all pairs, signifying superior user experience. All conditions fall within the positive user experience range, implying that hand-controllers offer higher perceived usability and visual realism induces a positive emotional response, most evident in M2R1, which also scored highest in quality perception.

4.5 Discussion

We comprehensively compared two input modalities—hand-tracking versus handheld-controller while conducting a reach-grab-place task in an immersive VR setting. This comparison was implemented across two levels of visual representation. We evaluated objective user performance metrics and subjective user experiences, considering perceived presence, cognitive workload, and user-friendliness. Hand-tracking, given its organic interaction, was hypothesized to score higher in intuitiveness and naturalness.

4.5.1 On performance

Our study revealed that the handheld-controller input method was more efficient for executing the reach-grab-place task, resulting in shorter completion times and superior object manipulation compared to hand-tracking. This was mirrored in the mental workload measurements, where participants reported lower burdens when using handheld-controllers, suggesting that this method may be less mentally taxing. Likewise, participants rated the handheld-controller modality higher for task-focused utility and overall attractiveness. However, no significant differences were identified regarding immersion, engagement, and realism between the two input methods within VR. This implies that while the handheld-controller may be more practical and favored for task performance, it doesn't necessarily enhance the user's subjective VR experience. Interestingly, the level of visual representation (from saturated to grayscale) didn't influence the user's VR experience either.

It's widely assumed that virtual experiences involving physical actions can boost spatial presence and cognitive immersion – mainly using virtual hands. Our results echo previous research that found hand-tracking modules to be perceived as more challenging for task execution. This may stem from a discrepancy between user expectations and the actual functionality of hand-tracking.

When evaluating hand-tracking versus handheld-controller input methods, it's crucial to consider the precision and reliability of tracking. Hand-tracking systems may struggle to capture user hand movements accurately, especially during quick or complex gestures. The Oculus Quest device we utilized does not specify an official hand-tracking accuracy rate, but previous studies deem it appropriate for various applications. Another common challenge with VR headsets offering inside-out tracking, like the Oculus Quest, is difficulty detecting hands due to self-occlusion. Some of our subjects encountered this issue, negatively impacting their task performance. Conversely, handheld-controllers provided a more accurate and dependable input method, as they do not depend on the headset's tracking capabilities.

4.5.2 Expectations

The kind of activities the user will engage in is a significant factor. In this experiment, we used a reach-grab-place task, where participants needed to apply different grips to handle particular items. We initially postulated that hand-tracking might be an ideal fit for the reach-grab-place task due to its incorporation of fine motor skills and precise hand movements, allowing users to use their own hands instead of a physical controller.

Revisiting the grips depicted in Figure 3 is crucial in this discussion. These illustrations present six unique ways to hold six different objects, contrasting with the singular "pinch" gesture presently supported by the VR system. Lifting paper clips using a terminal opposition grip aligned best with the VR headset's "pinch" gesture. The subterminal opposition grip for grabbing business cards was a close second. However, the participants had to resort to the same pinch gesture for picking up and moving the virtual mobile phone, which differs substantially from real-world practices. People usually apply a panoramic penta-digital grip for such shaped objects.

This issue is due to isomorphic mapping's limitations, which align elements in one system, such as physical actions, to corresponding elements in another, like virtual actions [200]. While isomorphic mapping can heighten immersion or presence in a virtual environment (VE), mismatches can lead to adverse effects [147], as we discovered. Technological constraints in accurately mapping real hand movements can result in lower scores. Grabbing a mug would elicit an immediate, automatic response in real-world situations. In contrast, the grip mismatch in VR made the same task require more cognitive effort. An action that should have been second nature suddenly needed conscious thought [35]. It's easy to see how such interactions might be seen as counter-intuitive, and these operational disparities could adversely impact the user's perceived authenticity or intuitiveness of the environment [195].

4.5.3 Usability

Considering user comfort and ease of use, our research revealed a notable disparity in usability between handheld-controllers and hand-tracking, with the former being more user-friendly. The controller-free interface consistently placed in the "neutral" category, which, while not discouraging, did not match our initial expectations of a higher user experience rating [191]. Of all the combinations, the controller in a realistic scenario (M2R1) consistently scored highest across AttrakDiff dimensions, emphasizing that user experience requires a delicate equilibrium of goal attainment, desire maintenance, and instinctual behavior.

The preference for the controller modality was significantly evident in NASA-TLX's mental workload indices, where participants reported an increased effort in cognitive information processing and individual responses when using virtual hands. Mental, physical, and temporal demands were nearly halved when participants transitioned from virtual hands to controller-based interactions. Contrary to our initial hypothesis, these findings suggest that controller-based interactions were more intuitive, and hand-tracking was perceived as noticeably frustrating and demanding. This could also be due to acquired digital literacy (Riecke et al., 2018), as all participants (N = 32) reported previous controller experience. Despite its closer-to-natural operation, hand-tracking needed an adjustment period before participants could achieve the necessary proficiency for the VR task. Therefore, a component of learnability could have influenced perception (Drew et al., 2018). On the other hand, handheld controllers presented a more familiar, steady, and ergonomic method of object interaction, and due to participants using controllers not expecting to interact with virtual objects using natural prehension, this is likely why the effects of expectation mismatch were absent. These findings suggest unpredictable input modality behavior can negatively affect user performance and overall experience.

4.5.4 Haptic Feedback

In our study comparing the two input modalities in VR, we observed notable differences in user performance and usability. While we did not explicitly include haptic feedback in our experimental setup, it is crucial to acknowledge its potential influence on the results.

Haptic feedback, which provides users with tactile sensations when interacting with virtual objects, can significantly impact user performance in VR. The absence of haptic feedback in the hand-tracking condition might also explain our study's longer completion times and less efficient object manipulation. Users may rely on haptic cues to gauge the force and precision of their interactions with virtual objects. The lack of this sensory feedback in hand-tracking may have led to increased cognitive workload and slower task completion.

Usability ratings, as measured by AttrakDiff and NASA-TLX, also warrant consideration in the context of haptic feedback. Handheld-controllers inherently provide haptic feedback when users interact with objects, which can enhance the sense of user-friendliness and intuitiveness. In contrast, the absence of haptic feedback in hand-tracking may have contributed to the perception of frustration and increased mental workload reported by participants.

4.6 Summary

This chapter focuses on the results of the comparative study of hand-tracking technology and handheld-controllers in VR, described in publications **VII** and **VIII**. We examine a simple motor activity, reach-grab-place task, to mimic real-world motor performance. We also investigated if enhancing the environment's visual realism (level of representation) alongside natural gesture-based interaction would improve subjective evaluations of presence, mental workload, and ease of use. Despite the increased naturalness of interactions promised by hand-tracking technology, the research found that it did not significantly enhance user performance or the subjective feelings of presence, naturalness, and engagement in the VR environment. The chapter recognizes that the relationship between interaction realism (how closely device interaction simulates real-world interaction) and the experience of naturalness was non-linear, indicating that increased realism does not necessarily improve QoE. Our findings contribute to research on virtual hands' potential and limitations for creating naturalistic VR experiences.

Chapter 5

Conclusions

This thesis aimed to advance scholarly understanding of IMEx by empirically investigating users' perception of action possibilities within VR environments and examining how such perceptions manifest in overt user behavior and subjective experiential responses. The overarching motivation stemmed from the need to conceptualize and validate comprehensive quality assessment frameworks that holistically account for the multifaceted and complex nature of VR experiences across its technological, compositional, contextual, and human-centric dimensions.

Through a series of mixed-methods studies utilizing behavioral observation techniques, performance metrics, and subjective self-report instruments, the research presented in this thesis offers substantial original empirical contributions illuminating the relationships between perceived action possibilities, overt user behavior, input modalities, and subjective impressions of presence, workload, pragmatics, and hedonics in VR contexts.

The background chapter provided a conceptual foundation by compiling integral insights from publications **I**, **III**, and **IV** to highlight the need for sophisticated IMEx evaluation frameworks that consider the intricate interconnections and interdependencies between influencing factors, experiential aspects, and user-reported features that collectively shape immersive quality perceptions. The proposed taxonomy of quality aspects emphasized a perspective beyond assessing technological capabilities in isolation to adopt a holistic stance weighing the confluence of system, user, and contextual factors on experiential facets. Publication **II** augmented this conceptualization by offering an original interpretation of authenticity as a cognitive judgment of quality in VR environments based on users' subjective expectations and priors, rather than purely technological fidelity. This aligned with the notion of quality as a reflective assessment, in contrast to an intuitive perception.

The first empirical investigation presented in publication **V** adopted a naturalistic observational methodology using systematic video analysis to examine potential correlations between perceived action possibilities afforded through virtual affordances and users' overt physical movements and behaviors. The findings revealed that while the inclusion of affordances and action possibilities significantly enhanced users' self-reported sense of spatial presence and engagement, no consistent predictive relationship could be established between perceived affordances and overt physical behavior across all participants. The study uncovered notable individual differences in how users perceive and physically respond to action possibilities based on factors like prior VR exposure, subjective expectations of system capabilities, and personal inclinations toward overt behavioral expressiveness.

Publication **VI** built on the findings of publication **V** by delving deeper into the concept of affordances and identifying four specific types ranging from immediate functional goals to higher psychological needs that collectively shape user experiences in immersive VR environments. The proposed taxonomy encompassing manipulation, effect, use, and experience affordances highlighted the multifaceted relationships between user abilities, environment features, and pragmatic and hedonic emotional outcomes. The study considered ergonomic and eudemonic affordance facets for optimizing holistic user experience.

Shifting the focus to input modalities, publication **VII** presented an empirical comparison of direct controller-free hand-tracking versus indirect handheld controllers for executing a motor task in VR using objective performance metrics. Despite the theoretical promises of more naturalistic interaction, hand-tracking did not demonstrate quantifiable advantages in terms of efficiency, accuracy, or dexterity compared to traditional handheld controllers. Publication **VIII** supplemented these performance-centric findings by incorporating assessments of perceived workload and pragmatic quality, reiterating that controller-free hand-tracking did not consistently enhance subjective feelings of presence, naturalness, or cognitive engagement relative to indirect controllers. The two studies underscored the need for careful technology-task-user fit assessments rather than assumptions of naturalness or fidelity.

Collectively, the series of empirical investigations presented in this thesis demonstrates the value of adopting naturalistic observational techniques and mixed-methods approaches that combine systematic behavioral codings, objective performance indicators, and subjective self-report instruments to derive comprehensive and ecologically valid insights into the Quality of Experience in immersive environments like VR. The studies offer empirical evidence illuminating the intricate relationships and gaps between technological capabilities, perceived affordances, input modalities, overt user behavior, and subjective user impressions of presence, cognitive workload, pragmatics, and hedonics in VR contexts.

Below we reiterate the standout limitations that provide opportunities for further work:

- Sample characteristics: Expanding the sample size and diversity could strengthen the generalizability of findings. Differences based on age, gender, culture, prior VR exposure, and domain experience could be investigated.
- Research Design: All studies were cross-sectional. Longitudinal studies could provide insights into changing perceptions with exposure and training over time.
- Objective Correlates: Physiological measures could be incorporated to strengthen empirical correlations between overt behavior and subjective states. This does present challenges especially when users are constantly moving.
- Static environments: The virtual test environments lacked dynamic elements. Dynamic environments with moving elements could reveal new interaction challenges.
- Limited Modalities: Only hand-tracking and controllers were compared. Other modalities like gaze or voice could be investigated. Incorporating haptic feedback could be worthwhile. Testing with lab-graded hardware could also be a direction.

Nevertheless, it is important to note that a cross-disciplinary perspective integrating concepts like affordances and naturalistic observational methodologies considerably expands the existing toolkit for human-centric VR evaluation by capturing subjective factors that may not be consciously accessible through self-reports. In addition to affirming relationships between system parameters and quality perceptions emphasized in earlier VR research, the studies demonstrate that user-centric factors like prior expectations, individual differences, technological acceptance, and digital literacy moderate the impact of technological capabilities on quality perceptions in complex ways that warrant deeper investigation.

5.0.1 Questions that Remain

There are several key unresolved issues and open questions that remain from this research:

How much technology is sufficient? The findings demonstrate the impact of technological capabilities, like object affordances, on user interactivity and experience; however, questions around determining sufficiency thresholds. Suppose relatively simpler features can engender strong feelings of presence and perceived credibility for users. In that case, it remains unclear how much value advanced features and interfaces, like hand-tracking, may contribute. Further work into affordances may help map the tradeoffs between technological complexity and experience optimization.

Should affordances only be pragmatic? The findings suggest simpler modalities enabled effective pragmatic task performance comparably to more naturalistic ones like hand-tracking. However, the relative importance of pragmatic usability and hedonic qualities like engagement, enjoyment, and meaning remains unclear across different scenarios. This raises relevant questions about the pragmatic vs. hedonic priorities of available object affordances, and warrants a further examination of their hedonic and aesthetic quality dimensions.

Shall VR experiences be more tailored? The research highlights potential gaps between an experience's technological possibilities and how users perceive and experience it based on their prior expectations and expertise levels. However, more direct investigation is needed into how baselines for quality assessments differ for expert/novice users or across different contexts of use (e.g., gaming, training, etc.). Users' mental models likely moderate perceived input-task fit, technology acceptance, and quality judgments. Understanding these influences of preconceptions and iterative expectation setting is key for experiences that align with and adapt to users.

Are subjective measure enough? The mixed-methods approach combining quantitative behavioral metrics with qualitative self-reports provided rich data, but some limitations remain. Self-reported measures of subjective psychological states like presence and authenticity have inherent challenges. Incorporating more direct psychophysiological measures (for cognitive/affective responses) could strengthen construct validity, but employing such tools in complex tasks has physical impediments and produces questionable data. Automated sensor-based tracking could enhance behavioral coding reliability beyond manual video observations alone. There is enormous research potential in advancing measurement approaches.

In conclusion, this dissertation provides significant and timely scholarly contributions toward advancing holistic Quality of Experience frameworks for understanding the multifaceted complexity of VR experiences. The substantial empirical findings provide insightful evidence highlighting gaps between technological capabilities, user-centered performance, and quality perceptions that illuminate promising directions for future research. Further studies can build on these findings to address unresolved questions about technology, cognition, behavior, and perception pertaining to user expectations, technology acceptance, input-task fit, presence, and optimizing naturalistic interactions and positive experiences in emerging immersive digital environments and applications. The research underscores the need to complement technological improvements with a greater understanding of cognitive and behavioral processes involved in immersive experiences.

Bibliography

[1] S. LaValle, “Virtual reality,” 2016.

[2] J. Jerald, The VR book: Human-centered design for virtual reality. Morgan & Claypool, 2015.

[3] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. Mühlberger, “The impact of perception and presence on emotional reactions: a review of research in virtual reality,” Frontiers in psychology, vol. 6, p. 26, 2015.

[4] I. A. Kapandji, “The physiology of the joint,” lower limb., vol. 2, 1987.

[5] J. Pallant, SPSS survival manual: A step by step guide to data analysis using IBM SPSS. Routledge, 2020.

[6] A. Perkis, C. Timmerer, S. Baraković, J. B. Husić, S. Bech, S. Bosse, J. Botev, K. Brunnström, L. Cruz, K. De Moor, et al., “Qualinet white paper on definitions of immersive media experience (imex),” arXiv preprint arXiv:2007.07032, 2020.

[7] K. Brunnström, S. A. Beker, K. De Moor, A. Dooms, S. Egger, M.-N. Garcia, T. Hossfeld, S. Jumisko-Pyykkö, C. Keimel, M.-C. Larabi, et al., “Qualinet white paper on definitions of quality of experience,” 2013.

[8] M. Slater and S. Wilbur, “A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments,” Presence: Teleoperators & Virtual Environments, vol. 6, no. 6, pp. 603–616, 1997.

[9] M. Slater, “Immersion and the illusion of presence in virtual reality,” British Journal of Psychology, vol. 109, no. 3, pp. 431–433, 2018.

[10] M. Lombard, T. B. Ditton, and L. Weinstein, “Measuring presence: the temple presence inventory,” in Proceedings of the 12th annual international workshop on presence, pp. 1–15, 2009.

[11] J. Steuer, “Defining virtual reality: Dimensions determining telepresence,” Journal of communication, vol. 42, no. 4, pp. 73–93, 1992.

[12] F. Biocca, T. Kim, and M. R. Levy, “The vision of virtual reality,” Communication in the age of virtual reality, pp. 3–14, 1995.

[13] R. Skarbez, “Plausibility illusion in virtual environments,” 2016.

[14] S. Hudson, S. Matson-Barkat, N. Pallamin, and G. Jegou, “With or without you? interaction and immersion in a virtual reality experience,” Journal of Business Research, vol. 100, pp. 459–468, 2019.

[15] G. C. Burdea and P. Coiffet, Virtual reality technology. John Wiley & Sons, 2003.

[16] R. P. McMahan, C. Lai, and S. K. Pal, “Interaction fidelity: the uncanny valley of virtual reality interactions,” in International Conference on Virtual, Augmented and Mixed Reality, pp. 59–70, Springer, 2016.

[17] T. B. Sheridan, “Recollections on presence beginnings, and some challenges for augmented and virtual reality,” Presence: Teleoperators and Virtual Environments, vol. 25, no. 1, pp. 75–77, 2016.

[18] R. Pals, L. Steg, J. Dontje, F. Siero, and K. van Der Zee, “Physical features, coherence and positive outcomes of person–environment interactions: A virtual reality study,” Journal of Environmental Psychology, vol. 40, pp. 108–116, 2014.

[19] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev, 3D user interfaces: theory and practice. Addison-Wesley Professional, 2017.

[20] K. Pfeuffer, B. Mayer, D. Mardanbegi, and H. Gellersen, “Gaze+ pinch interaction in virtual reality,” in Proceedings of the 5th Symposium on Spatial User Interaction, pp. 99–108, 2017.

[21] T. Rose, C. S. Nam, and K. B. Chen, “Immersion of virtual reality for rehabilitation-review,” Applied ergonomics, vol. 69, pp. 153–161, 2018.

[22] R. McGloin, K. Farrar, and M. Krcmar, “Video games, immersion, and cognitive aggression: does the controller matter?,” Media psychology, vol. 16, no. 1, pp. 65–87, 2013.

[23] H. Zhang, “Head-mounted display-based intuitive virtual reality training system for the mining industry,” International Journal of Mining Science and Technology, vol. 27, no. 4, pp. 717–722, 2017.

[24] W. S. Khor, B. Baker, K. Amin, A. Chan, K. Patel, and J. Wong, “Augmented and virtual reality in surgery—the digital surgical environment: applications, limitations and legal pitfalls,” Annals of translational medicine, vol. 4, no. 23, 2016.

[25] C. Boletsis, “The new era of virtual reality locomotion: A systematic literature review of techniques and a proposed typology,” Multimodal Technologies and Interaction, vol. 1, no. 4, p. 24, 2017.

[26] A. Ferracani, D. Pezzatini, J. Bianchini, G. Biscini, and A. Del Bimbo, “Locomotion by natural gestures for immersive virtual environments,” in Proceedings of the 1st international workshop on multimedia alternate realities, pp. 21–24, 2016.

[27] T. Schubert, F. Friedmann, and H. Regenbrecht, “The experience of presence: Factor analytic insights,” Presence: Teleoperators & Virtual Environments, vol. 10, no. 3, pp. 266–281, 2001.

[28] M. Slater, “Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments,” Philosophical Transactions of the Royal Society B: Biological Sciences, vol. 364, no. 1535, pp. 3549–3557, 2009.

[29] T. B. Sheridan et al., “Musings on telepresence and virtual presence.,” Presence Teleoperators Virtual Environ., vol. 1, no. 1, pp. 120–125, 1992.

[30] K. M. Lee, “Presence, explicated,” Communication theory, vol. 14, no. 1, pp. 27–50, 2004.

[31] B. Laurel, Computers as theatre. Addison-Wesley, 2013.

[32] I. Wechsung and K. De Moor, “Quality of experience versus user experience,” in Quality of Experience: Advanced Concepts, Applications and Methods, pp. 35–54, Springer, 2014.

[33] S. F. Gibson and B. Mirtich, “A survey of deformable modeling in computer graphics,” tech. rep., Technical report, Mitsubishi Electric Research Laboratories, 1997.

[34] J. Gibson, “The concept of affordances,” Perceiving, acting, and knowing, vol. 1, 1977.

[35] R. Coyne, “Heidegger and virtual reality: The implications of heidegger’s thinking for computer representations,” Leonardo, vol. 27, no. 1, pp. 65–73, 1994.

[36] M.-L. Ryan, Narrative as virtual reality 2: Revisiting immersion and interactivity in literature and electronic media. JHU press, 2015.

[37] A. Rollings and E. Adams, Andrew Rollings and Ernest Adams on game design. New Riders, 2003.

[38] A. McMahan, “Immersion, engagement and presence,” The video game theory reader, vol. 67, p. 86, 2003.

[39] F. Hammer, S. Egger-Lampl, and S. Möller, “Quality-of-user-experience: a position paper,” Quality and User Experience, vol. 3, pp. 1–15, 2018.

[40] R. Etemadpour, Y. Zhu, Q. Zhao, Y. Hu, B. Chen, M. A. Sharier, S. Zheng, and J. G. S Paiva, “Role of absence in academic success: an analysis using visualization tools,” Smart Learning Environments, vol. 7, no. 1, pp. 1–25, 2020.

[41] M. Hassenzahl and N. Tractinsky, “User experience-a research agenda,” Behaviour & information technology, vol. 25, no. 2, pp. 91–97, 2006.

[42] H. Väätäjä, P. Majaranta, P. Isokoski, Y. Gizatdinova, M. V. Kujala, S. Somppi, A. Vehkaoja, O. Vainio, O. Juhlin, M. Ruohonen, et al., “Happy dogs and happy owners: Using dog activity monitoring technology in everyday life,” in Proceedings of the Fifth International Conference on Animal-Computer Interaction, pp. 1–12, 2018.

[43] M. Hassenzahl, “User experience and experience design,” The encyclopedia of human-computer interaction, vol. 2, 2013.

[44] M. S. El-Nasr, A. Drachen, and A. Canossa, Game analytics. Springer, 2016.

[45] B. Zheng, X. Jiang, G. Tien, A. Meneghetti, O. N. M. Panton, and M. S. Atkins, “Workload assessment of surgeons: correlation between nasa tlx and blinks,” Surgical endoscopy, vol. 26, no. 10, pp. 2746–2750, 2012.

[46] C. S. Loh, Y. Sheng, and D. Ifenthaler, “Serious games analytics: Theoretical framework,” Serious games analytics: Methodologies for performance measurement, assessment, and improvement, pp. 3–29, 2015.

[47] T. W. Schubert, “The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness.,” Z. für Medienpsychologie, vol. 15, no. 2, pp. 69–71, 2003.

[48] B. G. Witmer and M. J. Singer, “Measuring presence in virtual environments: A presence questionnaire,” Presence, vol. 7, no. 3, pp. 225–240, 1998.

[49] J. Lessiter, J. Freeman, E. Keogh, and J. Davidoff, “A cross-media presence questionnaire: The itc-sense of presence inventory,” Presence: Teleoperators & Virtual Environments, vol. 10, no. 3, pp. 282–297, 2001.

[50] B. Yan, S. Ni, X. Wang, J. Liu, Q. Zhang, and K. Peng, “Using virtual reality to validate the chinese version of the independent television commission-sense of presence inventory,” SAGE Open, vol. 10, no. 2, p. 2158244020922878, 2020.

[51] S. G. Hart, “Nasa-task load index (nasa-tlx); 20 years later,” in Proceedings of the human factors and ergonomics society annual meeting, vol. 50, pp. 904–908, Sage publications Sage CA: Los Angeles, CA, 2006.

[52] L. Colligan, H. W. Potts, C. T. Finn, and R. A. Sinkin, “Cognitive workload changes for nurses transitioning from a legacy system with paper documentation to a commercial electronic health record,” International journal of medical informatics, vol. 84, no. 7, pp. 469–476, 2015.

[53] A. Ghanbary Sartang, M. Ashnagar, E. Habibi, and S. Sadeghi, “Evaluation of rating scale mental effort (rsme) effectiveness for mental workload assessment in nurses,” Journal of Occupational Health and Epidemiology, vol. 5, no. 4, pp. 211–217, 2016.

[54] M. Hassenzahl, M. Burmester, and F. Koller, “Attrakdiff: Ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität,” Mensch & Computer 2003: Interaktion in Bewegung, pp. 187–196, 2003.

[55] T. Walsh, J. Varsaluoma, S. Kujala, P. Nurkka, H. Petrie, and C. Power, “Axe ux: Exploring long-term user experience with iscale and attrakdiff,” in Proceedings of the 18th international academic mindtrek conference: Media business, management, content & services, pp. 32–39, 2014.

[56] L. Feng and W. Wei, "An empirical study on user experience evaluation and identification of critical ux issues," Sustainability, vol. 11, no. 8, p. 2432, 2019.

[57] C. Lallemand, V. Koenig, G. Gronier, and R. Martin, "Création et validation d'une version française du questionnaire attrakdiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs," European Review of Applied Psychology, vol. 65, no. 5, pp. 239–252, 2015.

[58] ITU-T, "Recommendation P.809," Subjective evaluation methods for gaming quality, 06 2018.

[59] ITU-T, "Recommendation P.911," Subjective audiovisual quality assessment methods for multimedia applications, 12 1998.

[60] ITU-T, "Recommendation P.919," Subjective test methodologies for 360º video on head-mounted displays, 10 2020.

[61] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented reality: A class of displays on the reality-virtuality continuum," in Telemanipulator and telepresence technologies, vol. 2351, pp. 282–292, Spie, 1995.

[62] J. M. Pawlowski, "The quality adaptation model: adaptation and adoption of the quality standard iso/iec 19796-1 for learning, education, and training," Journal of Educational Technology & Society, vol. 10, no. 2, pp. 3–16, 2007.

[63] A. Raake and S. Egger, "Quality and quality of experience," in Quality of Experience: Advanced concepts, applications and methods, pp. 11–33, Springer, 2014.

[64] S. Möller and A. Raake, Quality of experience: advanced concepts, applications and methods. Springer, 2014.

[65] C. Keighrey, R. Flynn, S. Murray, and N. Murray, "A qoe evaluation of immersive augmented and virtual reality speech & language assessment applications," in 2017 ninth international conference on quality of multimedia experience (QoMEX), pp. 1–6, IEEE, 2017.

[66] N. C. Nilsson, R. Nordahl, and S. Serafin, "Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence.," Human Technology, vol. 12, no. 2, 2016.

[67] H.-H. Liou, S. J. Yang, S. Y. Chen, and W. Tarng, "The influences of the 2d image-based augmented reality and virtual reality on student learning," Journal of Educational Technology & Society, vol. 20, no. 3, pp. 110–121, 2017.

[68] F. Biocca, "The evolution of interactive media," Narrative impact. Social and cognitive foundations, pp. 97–130, 2002.

[69] M. Lombard and T. Ditton, "At the heart of it all: The concept of presence," Journal of computer-mediated communication, vol. 3, no. 2, p. JCMC321, 1997.

[70] R. M. Baños, C. Botella, M. Alcañiz, V. Liaño, B. Guerrero, and B. Rey, "Immersion and emotion: their impact on the sense of presence," Cyberpsychology & behavior, vol. 7, no. 6, pp. 734–741, 2004.

[71] M. J. Schuemie, P. Van Der Straaten, M. Krijn, and C. A. Van Der Mast, "Research on presence in virtual reality: A survey," CyberPsychology & Behavior, vol. 4, no. 2, pp. 183–201, 2001.

[72] M.-L. Ryan, "From narrative games to playable stories: Toward a poetics of interactive narrative," StoryWorlds: a journal of narrative studies, vol. 1, pp. 43–59, 2009.

[73] J. D. Bolter, "Remediation and the desire for immediacy," Convergence, vol. 6, no. 1, pp. 62–71, 2000.

[74] S. Bronack, R. Sanders, A. Cheney, R. Riedl, J. Tashner, and N. Matzen, "Presence pedagogy: Teaching and learning in a 3d virtual immersive world.," International journal of teaching and learning in higher education, vol. 20, no. 1, pp. 59–69, 2008.

[75] G. G. Robertson, S. K. Card, and J. D. Mackinlay, "Three views of virtual reality: nonimmersive virtual reality," Computer, vol. 26, no. 2, p. 81, 1993.

[76] R. L. Gregory, "What do qualia do?," 1996.

[77] S. Shoemaker, "Qualities and qualia: What's in the mind?," Philosophy and Phenomenological Research, vol. 50, pp. 109–131, 1990.

[78] U. Jekosch, Voice and speech quality perception: assessment and evaluation. Springer Science & Business Media, 2005.

[79] J. Blauert and U. Jekosch, "A layer model of sound quality," Journal of the Audio Engineering Society, vol. 60, no. 1/2, pp. 4–12, 2012.

[80] S. Uhrig, Human information processing in speech quality assessment. Springer, 2021.

[81] S. Möller, Quality Engineering: Quality of Communication Technology Systems. Springer Nature, 2023.

[82] A. R. Damasio, "Review: Toward a neurobiology of emotion and feeling: Operational concepts and hypotheses," The Neuroscientist, vol. 1, no. 1, pp. 19–25, 1995.

[83] S. B. Gilbert, "Perceived realism of virtual environments depends on authenticity," Presence, vol. 25, no. 4, pp. 322–324, 2016.

[84] D. A. Bowman, R. P. McMahan, and E. D. Ragan, "Questioning naturalism in 3d user interfaces," Communications of the ACM, vol. 55, no. 9, pp. 78–88, 2012.

[85] A. Hameed, A. Perkis, and S. Möller, "Evaluating hand-tracking interaction for performing motor-tasks in vr learning environments," in 2021 13th International Conference on Quality of Multimedia Experience (QoMEX), pp. 219–224, IEEE, 2021.

[86] A. Rovira, D. Swapp, B. Spanlang, and M. Slater, "The use of virtual reality in the study of people's responses to violent incidents," Frontiers in behavioral neuroscience, vol. 3, p. 917, 2009.

[87] M. Slater, "Implicit learning through embodiment in immersive virtual reality," Virtual, augmented, and mixed realities in education, pp. 19–33, 2017.

[88] S. Bucolo, "Understanding cross cultural differences during interaction within immersive virtual environments," in Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry, pp. 221–224, 2004.

[89] D. Weibel, B. Wissmath, and F. W. Mast, "Immersion in mediated environments: The role of personality traits," Cyberpsychology, Behavior, and Social Networking, vol. 13, no. 3, pp. 251–256, 2010.

[90] E.-J. Lee, "Authenticity model of (mass-oriented) computer-mediated communication: Conceptual explorations and testable propositions," Journal of Computer-Mediated Communication, vol. 25, no. 1, pp. 60–73, 2020.

[91] D. Arsenault, "Dark waters: Spotlight on immersion," in GAMEON-NA International Conference, pp. 50–52, Eurosis, 2005.

[92] J. Mütterlein, “The three pillars of virtual reality? investigating the roles of immersion, presence, and interactivity,” 2018.

[93] D. A. Bowman and L. F. Hodges, “Formalizing the design, evaluation, and application of interaction techniques for immersive virtual environments,” Journal of Visual Languages & Computing, vol. 10, no. 1, pp. 37–53, 1999.

[94] J. Flotyński, A. Nowak, and K. Walczak, “Explorable representation of interaction in vr/ar environments,” in Augmented Reality, Virtual Reality, and Computer Graphics: 5th International Conference, AVR 2018, Otranto, Italy, June 24–27, 2018, Proceedings, Part II 5, pp. 589–609, Springer, 2018.

[95] J. Flotyński, M. Krzyszkowski, and K. Walczak, “Semantic composition of 3d content behavior for explorable virtual reality applications,” in Virtual Reality and Augmented Reality: 14th EuroVR International Conference, EuroVR 2017, Laval, France, December 12–14, 2017, Proceedings 14, pp. 3–23, Springer, 2017.

[96] N. Magnenat-Thalmann, H. Kim, A. Egges, and S. Garchery, “Believability and interaction in virtual worlds,” in 11th International Multimedia Modelling Conference, pp. 2–9, IEEE, 2005.

[97] A. Luciani, D. Urma, S. Marlière, and J. Chevrier, “Presence: the sense of believability of inaccessible worlds,” Computers & graphics, vol. 28, no. 4, pp. 509–517, 2004.

[98] T. Kim and F. Biocca, “Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion.,” Journal of computer-mediated communication, vol. 3, no. 2, p. JCMC325, 1997.

[99] S. Moller, K.-P. Engelbrecht, C. Kuhnel, I. Wechsung, and B. Weiss, “A taxonomy of quality of service and quality of experience of multimodal human-machine interaction,” in 2009 international workshop on quality of multimedia experience, pp. 7–12, IEEE, 2009.

[100] S. Möller, S. Schmidt, and J. Beyer, “Gaming taxonomy: An overview of concepts and evaluation methods for computer gaming qoe,” in 2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX), pp. 236–241, IEEE, 2013.

[101] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton, “Measuring and defining the experience of immersion in games,”

International journal of human-computer studies, vol. 66, no. 9, pp. 641–661, 2008.

[102] B. Laugwitz, T. Held, and M. Schrepp, "Construction and evaluation of a user experience questionnaire," in HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4, pp. 63–76, Springer, 2008.

[103] J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhart, and J. N. Pidruzny, "The development of the game engagement questionnaire: A measure of engagement in video game-playing," Journal of experimental social psychology, vol. 45, no. 4, pp. 624–634, 2009.

[104] W. A. IJsselsteijn, Y. A. de Kort, and K. Poels, "The game experience questionnaire," Eindhoven: Technische Universiteit Eindhoven, vol. 46, no. 1, 2013.

[105] J. Takatalo, G. Nyman, and L. Laaksonen, "Components of human experience in virtual environments," Computers in Human Behavior, vol. 24, no. 1, pp. 1–15, 2008.

[106] A. Hameed, S. Irshad, and A. Perkis, "Towards a quality framework for immersive media experiences: A holistic approach," in International Conference on Interactive Digital Storytelling, pp. 389–394, Springer, 2019.

[107] E. Gobbetti and R. Scateni, "Virtual reality: Past, present, and future," Virtual environments in clinical psychology and neuroscience: Methods and techniques in advanced patient-therapist interaction, 1998.

[108] K. Yin, Z. He, J. Xiong, J. Zou, K. Li, and S.-T. Wu, "Virtual reality and augmented reality displays: advances and future perspectives," Journal of Physics: Photonics, vol. 3, no. 2, p. 022010, 2021.

[109] W. R. Sherman and A. B. Craig, Understanding virtual reality: Interface, application, and design. Morgan Kaufmann, 2018.

[110] J. Mütterlein and T. Hess, "Immersion, presence, interactivity: Towards a joint understanding of factors influencing virtual reality acceptance and use," 2017.

[111] M. Falahee, K. Latham, and E. Geelhoed, "Safety and comfort of eyeglass displays," in International Symposium on Handheld and Ubiquitous Computing, pp. 236–247, Springer, 2000.

[112] E. Chang, H. T. Kim, and B. Yoo, “Virtual reality sickness: a review of causes and measurements,” International Journal of Human–Computer Interaction, vol. 36, no. 17, pp. 1658–1682, 2020.

[113] P. Vorderer, W. Wirth, F. R. Gouveia, F. Biocca, T. Saari, L. Jäncke, S. Böcking, H. Schramm, A. Gysbers, T. Hartmann, et al., “Mec spatial presence questionnaire,” Retrieved Sept, vol. 18, no. 2015, p. 6, 2004.

[114] M. M. Bradley and P. J. Lang, “Measuring emotion: the self-assessment manikin and the semantic differential,” Journal of behavior therapy and experimental psychiatry, vol. 25, no. 1, pp. 49–59, 1994.

[115] T. Baumgartner, L. Valko, M. Esslen, and L. Jäncke, “Neural correlate of spatial presence in an arousing and noninteractive virtual reality: an eeg and psychophysiology study,” CyberPsychology & Behavior, vol. 9, no. 1, pp. 30–45, 2006.

[116] D. P. Salgado, F. R. Martins, T. B. Rodrigues, C. Keighrey, R. Flynn, E. L. M. Naves, and N. Murray, “A qoe assessment method based on eda, heart rate and eeg of a virtual reality assistive technology system,” in Proceedings of the 9th ACM Multimedia Systems Conference, pp. 517–520, 2018.

[117] D. Egan, S. Brennan, J. Barrett, Y. Qiao, C. Timmerer, and N. Murray, “An evaluation of heart rate and electrodermal activity as an objective qoe evaluation method for immersive virtual reality environments,” in 2016 eighth international conference on quality of multimedia experience (QoMEX), pp. 1–6, IEEE, 2016.

[118] H. G. Hoffman, T. Richards, B. Coda, A. Richards, and S. R. Sharar, “The illusion of presence in immersive virtual reality during an fmri brain scan,” Cyberpsychology & behavior, vol. 6, no. 2, pp. 127–131, 2003.

[119] J. S. Taube, S. Valerio, and R. M. Yoder, “Is navigation in virtual reality with fmri really navigation?,” Journal of cognitive neuroscience, vol. 25, no. 7, pp. 1008–1019, 2013.

[120] V. Clay, P. König, and S. Koenig, “Eye tracking in virtual reality,” Journal of eye movement research, vol. 12, no. 1, 2019.

[121] N. A. Rappa, S. Ledger, T. Teo, K. Wai Wong, B. Power, and B. Hilliard, “The use of eye tracking technology to explore learning and performance within virtual reality and mixed reality settings: a scoping review,” Interactive Learning Environments, vol. 30, no. 7, pp. 1338–1350, 2022.

[122] A. Hameed, S. Möller, and A. Perkis, "How good are virtual hands? influences of input modality on motor tasks in virtual reality," Journal of Environmental Psychology, vol. 92, p. 102137, 2023.

[123] J.-N. Voigt-Antons, T. Kojic, D. Ali, and S. Möller, "Influence of hand tracking as a way of interaction in virtual reality on user experience," in 2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–4, IEEE, 2020.

[124] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale," Intl. Journal of Human–Computer Interaction, vol. 24, no. 6, pp. 574–594, 2008.

[125] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in Advances in psychology, vol. 52, pp. 139–183, Elsevier, 1988.

[126] J. R. Lewis, "Psychometric evaluation of an after-scenario questionnaire for computer usability studies: the asq," ACM Sigchi Bulletin, vol. 23, no. 1, pp. 78–81, 1991.

[127] K. Kilteni, R. Groten, and M. Slater, "The sense of embodiment in virtual reality," Presence: Teleoperators and Virtual Environments, vol. 21, no. 4, pp. 373–387, 2012.

[128] D. Saredakis, A. Szpak, B. Birckhead, H. A. Keage, A. Rizzo, and T. Loetscher, "Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis," Frontiers in human neuroscience, vol. 14, p. 96, 2020.

[129] M. Slater, "Presence and the sixth sense," Presence, vol. 11, no. 4, pp. 435–439, 2002.

[130] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter, "Spatial sound with loudspeakers and its perception: A review of the current state," Proceedings of the IEEE, vol. 101, no. 9, pp. 1920–1938, 2013.

[131] F. Rumsey, "Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm," Journal of the Audio Engineering Society, vol. 50, no. 9, pp. 651–666, 2002.

[132] R. Held and N. Durlach, "Telepresence. presence: Teleoperatorsand virtual environments, 1 (1), 109-112. hendrix, c, & barfield, w.(1996). presence within virtual environments as a function ofvisual display parameters,"

Presence: Teleoperators and Virtual Environments, vol. 5, no. 2, pp. 5–39, 1992.

[133] C. D. Murray, J. Fox, and S. Pettifer, “Absorption, dissociation, locus of control and presence in virtual reality,” Computers in Human Behavior, vol. 23, no. 3, pp. 1347–1354, 2007.

[134] M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, “Toward low-latency and ultra-reliable virtual reality,” IEEE Network, vol. 32, no. 2, pp. 78–84, 2018.

[135] K. Brunnström, E. Dima, T. Qureshi, M. Johanson, M. Andersson, and M. Sjöström, “Latency impact on quality of experience in a virtual reality simulator for remote control of machines,” Signal Processing: Image Communication, vol. 89, p. 116005, 2020.

[136] W. Wu, A. Arefin, Z. Huang, P. Agarwal, S. Shi, R. Rivas, and K. Nahrstedt, “" i’m the jedi!"-a case study of user experience in 3d tele-immersive gaming,” in 2010 IEEE International Symposium on Multimedia, pp. 220–227, IEEE, 2010.

[137] P. Harms, “Vr interaction modalities for the evaluation of technical device prototypes,” in Human-Computer Interaction–INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part IV 17, pp. 416–435, Springer, 2019.

[138] H. Holderied, “Evaluation of interaction concepts in virtual reality applications,” INFORMATIK 2017, 2017.

[139] G. Fontaine, “The experience of a sense of presence in intercultural and international encounters,” Presence: Teleoperators & Virtual Environments, vol. 1, no. 4, pp. 482–490, 1992.

[140] F. Rebelo, P. Noriega, E. Duarte, and M. Soares, “Using virtual reality to assess user experience,” Human factors, vol. 54, no. 6, pp. 964–982, 2012.

[141] A. S. Fernandes and S. K. Feiner, “Combating vr sickness through subtle dynamic field-of-view modification,” in 2016 IEEE symposium on 3D user interfaces (3DUI), pp. 201–210, IEEE, 2016.

[142] T. C. Peck, H. Fuchs, and M. C. Whitton, “The design and evaluation of a large-scale real-walking locomotion interface,” IEEE transactions on visualization and computer graphics, vol. 18, no. 7, pp. 1053–1067, 2011.

[143] E. B. Goldstein and J. Brockmole, Sensation and perception. Cengage Learning, 2016.

[144] D. A. Bowman, D. Koller, and L. F. Hodges, “Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques,” in Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality, pp. 45–52, IEEE, 1997.

[145] W. Chen, A. Plancoulaine, N. Férey, D. Touraine, J. Nelson, and P. Bourdot, “6dof navigation in virtual worlds: comparison of joystick-based and head-controlled paradigms,” in Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology, pp. 111–114, 2013.

[146] M. Slater, M. Usoh, and A. Steed, “Taking steps: the influence of a walking technique on presence in virtual reality,” ACM Transactions on Computer-Human Interaction (TOCHI), vol. 2, no. 3, pp. 201–219, 1995.

[147] B. E. Riecke and D. Feuereissen, “To move or not to move: can active control and user-driven motion cueing enhance self-motion perception (" vection") in virtual reality?,” in Proceedings of the ACM symposium on applied perception, pp. 17–24, 2012.

[148] K. Rahimi, C. Banigan, and E. D. Ragan, “Scene transitions and teleportation in virtual reality and the implications for spatial awareness and sickness,” IEEE transactions on visualization and computer graphics, vol. 26, no. 6, pp. 2273–2287, 2018.

[149] A. Hameed and A. Perkis, “Spatial storytelling: Finding interdisciplinary immersion,” in International Conference on Interactive Digital Storytelling, pp. 323–332, Springer, 2018.

[150] D. Quesnel and B. E. Riecke, “Are you awed yet? how virtual reality gives us awe and goose bumps,” Frontiers in psychology, vol. 9, p. 2158, 2018.

[151] S. Coren, L. M. Ward, and J. T. Enns, Sensation and perception. John Wiley & Sons Hoboken, NJ, 2004.

[152] C. Guastavino, V. Larcher, G. Catusseau, and P. Boussard, “Spatial audio quality evaluation: comparing transaural, ambisonics and stereo,” Georgia Institute of Technology, 2007.

[153] J. Hummel, R. Wolff, T. Stein, A. Gerndt, and T. Kuhlen, “An evaluation of open source physics engines for use in virtual reality assembly simulations,” in Advances in Visual Computing: 8th International Symposium,

ISVC 2012, Rethymnon, Crete, Greece, July 16-18, 2012, Revised Selected Papers, Part II 8, pp. 346–357, Springer, 2012.

[154] P. A. Rauschnabel, R. Felix, C. Hinsch, H. Shahab, and F. Alt, "What is xr? towards a framework for augmented and virtual reality," Computers in human behavior, vol. 133, p. 107289, 2022.

[155] A. G. Sutcliffe, C. Poullis, A. Gregoriades, I. Katsouri, A. Tzanavari, and K. Herakleous, "Reflecting on the design process for virtual reality applications," International Journal of Human–Computer Interaction, vol. 35, no. 2, pp. 168–179, 2019.

[156] M. Slater, D. Banakou, A. Beacco, J. Gallego, F. Macia-Varela, and R. Oliva, "A separate reality: An update on place illusion and plausibility in virtual reality," Frontiers in Virtual Reality, vol. 3, p. 914392, 2022.

[157] J.-C. Lepecq, L. Bringoux, J.-M. Pergandi, T. Coyle, and D. Mestre, "Afforded actions as a behavioral assessment of physical presence in virtual environments," Virtual reality, vol. 13, no. 3, pp. 141–151, 2009.

[158] J. M. Loomis, "Presence in virtual reality and everyday life: Immersion within a world of representation.," PRESENCE: Teleoperators & Virtual Environments, vol. 25, no. 2, 2016.

[159] V.-M. Karhulahti, "Suspending virtual disbelief: a perspective on narrative coherence," in Interactive Storytelling: 5th International Conference, ICIDS 2012, San Sebastián, Spain, November 12-15, 2012. Proceedings 5, pp. 1–17, Springer, 2012.

[160] R. Skarbez, F. P. Brooks, and M. C. Whitton, "Immersion and coherence: Research agenda and early results," IEEE transactions on visualization and computer graphics, vol. 27, no. 10, pp. 3839–3850, 2020.

[161] J. Llobera, K. J. Blom, and M. Slater, "Telling stories within immersive virtual environments," Leonardo, vol. 46, no. 5, pp. 471–476, 2013.

[162] S. Beckhaus and R. W. Lindeman, "Experiential fidelity: Leveraging the mind to improve the vr experience," in Virtual realities: Dagstuhl seminar 2008, pp. 39–49, Springer, 2011.

[163] M. Carrozzino and M. Bergamasco, "Beyond virtual museums: Experiencing immersive virtual reality in real museums," Journal of cultural heritage, vol. 11, no. 4, pp. 452–458, 2010.

[164] S. Styliani, L. Fotis, K. Kostas, and P. Petros, “Virtual museums, a survey and some issues for consideration,” Journal of cultural Heritage, vol. 10, no. 4, pp. 520–528, 2009.

[165] M. Schmid Mast, E. P. Kleinlogel, B. Tur, and M. Bachmann, “The future of interpersonal skills development: Immersive virtual reality training with virtual humans,” Human Resource Development Quarterly, vol. 29, no. 2, pp. 125–141, 2018.

[166] I. Bergström, S. Azevedo, P. Papiotis, N. Saldanha, and M. Slater, “The plausibility of a string quartet performance in virtual reality,” IEEE transactions on visualization and computer graphics, vol. 23, no. 4, pp. 1352–1359, 2017.

[167] A. Hameed and A. Perkis, “A subjective and behavioral assessment of affordances in virtual architectural walkthroughs,” Applied Sciences, vol. 11, no. 17, p. 7846, 2021.

[168] S. Irshad, A. Perkis, and W. Azam, “Wayfinding in virtual reality serious game: An exploratory study in the context of user perceived experiences,” Applied Sciences, vol. 11, no. 17, p. 7822, 2021.

[169] M. V. Sanchez-Vives and M. Slater, “From presence to consciousness through virtual reality,” Nature reviews neuroscience, vol. 6, no. 4, pp. 332–339, 2005.

[170] H. A. van Veen, H. K. Distler, S. J. Braun, and H. H. Bülthoff, “Navigating through a virtual city: Using virtual reality technology to study human action and perception,” Future Generation Computer Systems, vol. 14, no. 3, pp. 231–242, 1998. Virtual Reality in Industy and Research.

[171] R. L. Solso and J. E. McCarthy, “Prototype formation of faces: A case of pseudo-memory,” British Journal of Psychology, vol. 72, no. 4, pp. 499–503, 1981.

[172] T. Dorta, A. Lesage, E. Pérez, and J. C. Bastien, “Signs of collaborative ideation and the hybrid ideation space,” in Design Creativity 2010, pp. 199–206, Springer, 2011.

[173] N. Gu, M. J. Kim, and M. L. Maher, “Technological advancements in synchronous collaboration: The effect of 3d virtual worlds and tangible user interfaces on architectural design,” Automation in Construction, vol. 20, no. 3, pp. 270–278, 2011.

[174] A. Carassa, F. Morganti, and M. Tirassa, “A situated cognition perspective on presence,” in Proceedings of the Annual Meeting of the Cognitive Science Society, vol. 27, 2005.

[175] R. Skarbez, S. Neyret, F. P. Brooks, M. Slater, and M. C. Whitton, “A psychophysical experiment regarding components of the plausibility illusion,” IEEE transactions on visualization and computer graphics, vol. 23, no. 4, pp. 1369–1378, 2017.

[176] M. Hofer, T. Hartmann, A. Eden, R. Ratan, and L. Hahn, “The role of plausibility in the experience of spatial presence in virtual environments,” Frontiers in Virtual Reality, vol. 1, p. 2, 2020.

[177] H. Heft, “Affordances and the body: An intentional analysis of gibson’s ecological approach to visual perception,” Journal for the theory of social behaviour, vol. 19, no. 1, pp. 1–30, 1989.

[178] F. Pucillo and G. Cascini, “A framework for user experience, needs and affordances,” Design Studies, vol. 35, no. 2, pp. 160–179, 2014.

[179] K. M. Sheldon, A. J. Elliot, Y. Kim, and T. Kasser, “What is satisfying about satisfying events? testing 10 candidate psychological needs.,” Journal of personality and social psychology, vol. 80, no. 2, p. 325, 2001.

[180] J. H. Steffen, J. E. Gaskin, T. O. Meservy, J. L. Jenkins, and I. Wolman, “Framework of affordances for virtual reality and augmented reality,” Journal of Management Information Systems, vol. 36, no. 3, pp. 683–729, 2019.

[181] T. Regia-Corte, M. Marchal, G. Cirio, and A. Lécuyer, “Perceiving affordances in virtual reality: influence of person and environmental properties in perception of standing on virtual grounds,” Virtual Reality, vol. 17, no. 1, pp. 17–28, 2013.

[182] D.-H. Shin, “The role of affordance in the experience of virtual reality learning: Technological and affective affordances in virtual reality,” Telematics and Informatics, vol. 34, no. 8, pp. 1826–1836, 2017.

[183] D. Kahneman, “A perspective on judgment and choice: mapping bounded rationality.,” American psychologist, vol. 58, no. 9, p. 697, 2003.

[184] G. Makransky, L. Lilleholt, and A. Aaby, “Development and validation of the multimodal presence scale for virtual reality environments: A confirmatory factor analysis and item response theory approach,” Computers in Human Behavior, vol. 72, pp. 276–285, 2017.

[185] M.-A. Moinnereau, A. A. de Oliveira Jr, and T. H. Falk, "Immersive media experience: a survey of existing methods and tools for human influential factors assessment," Quality and User Experience, vol. 7, no. 1, p. 5, 2022.

[186] G. Thorsteinsson, "Developing an understanding of the pedagogy of using a virtual reality learning environment (vrle) to support innovation education," in The Routledge International handbook of innovation education, pp. 486–500, Routledge, 2013.

[187] V. Liagkou, D. Salmas, and C. Stylios, "Realizing virtual reality learning environment for industry 4.0," Procedia CIRP, vol. 79, pp. 712–717, 2019.

[188] S. G. Izard, J. A. Juanes, F. J. G. Peñalvo, J. M. G. Estella, M. J. S. Ledesma, and P. Ruisoto, "Virtual reality as an educational and training tool for medicine," Journal of medical systems, vol. 42, no. 3, pp. 1–5, 2018.

[189] D. B. Mekbib, J. Han, L. Zhang, S. Fang, H. Jiang, J. Zhu, A. W. Roe, and D. Xu, "Virtual reality therapy for upper limb rehabilitation in patients with stroke: a meta-analysis of randomized clinical trials," Brain injury, vol. 34, no. 4, pp. 456–465, 2020.

[190] D. Regazzoni, C. Rizzi, and A. Vitali, "Virtual reality applications: guidelines to design natural user interface," in International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 51739, p. V01BT02A029, American Society of Mechanical Engineers, 2018.

[191] A. Masurovsky, P. Chojecki, D. Runde, M. Lafci, D. Przewozny, and M. Gaebler, "Controller-free hand tracking for grab-and-place tasks in immersive virtual reality: Design elements and their empirical study," Multimodal Technologies and Interaction, vol. 4, no. 4, p. 91, 2020.

[192] M. Kim, C. Jeon, and J. Kim, "A study on immersion and presence of a portable hand haptic system for immersive virtual reality," Sensors, vol. 17, no. 5, p. 1141, 2017.

[193] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev, "An introduction to 3-d user interface design," Presence, vol. 10, no. 1, pp. 96–108, 2001.

[194] E. D. Ragan, D. A. Bowman, R. Kopper, C. Stinson, S. Scerbo, and R. P. McMahan, "Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task," IEEE transactions on visualization and computer graphics, vol. 21, no. 7, pp. 794–807, 2015.

[195] S. Weber, D. Weibel, and F. W. Mast, "How to get there when you are there already? defining presence in virtual reality and the importance of perceived realism," Frontiers in psychology, p. 1538, 2021.

[196] C. J. Hamblin, Transfer of* training from virtual reality environments. Wichita State University, 2005.

[197] A. Ricca, A. Chellali, and S. Otmane, "Study of interaction fidelity for two viewpoint changing techniques in a virtual biopsy trainer," in 2017 IEEE Virtual Reality (VR), pp. 227–228, IEEE, 2017.

[198] H. K. Kim, D. W. Rattner, and M. A. Srinivasan, "The role of simulation fidelity in laparoscopic surgical training," in International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 1–8, Springer, 2003.

[199] A. Chellali, H. Mentis, A. Miller, W. Ahn, V. S. Arikatla, G. Sankaranarayanan, S. De, S. D. Schwaitzberg, and C. G. Cao, "Achieving interface and environment fidelity in the virtual basic laparoscopic surgical trainer," International journal of human-computer studies, vol. 96, pp. 22–37, 2016.

[200] E. L. Hutchins, J. D. Hollan, and D. A. Norman, "Direct manipulation interfaces," Human–computer interaction, vol. 1, no. 4, pp. 311–338, 1985.

[201] B. E. Riecke, J. J. LaViola Jr, and E. Kruijff, "3d user interfaces for virtual reality and games: 3d selection, manipulation, and spatial navigation," in ACM SIGGRAPH 2018 Courses, pp. 1–94, 2018.

[202] C. M. Myers, A. Furqan, and J. Zhu, "The impact of user characteristics and preferences on performance with an unfamiliar voice user interface," in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–9, 2019.

[203] M. E. Nunnally and Y. Bitan, "Time to get off this pig's back?: the human factors aspects of the mismatch between device and real-world knowledge in the health care environment," Journal of Patient Safety, vol. 2, no. 3, pp. 124–131, 2006.

[204] C. Våpenstad, E. F. Hofstad, T. Langø, R. Mårvik, and M. K. Chmarra, "Perceiving haptic feedback in virtual reality simulators," Surgical endoscopy, vol. 27, no. 7, pp. 2391–2397, 2013.

[205] C. van de Kamp and F. T. Zaal, "Prehension is really reaching and grasping," Experimental Brain Research, vol. 182, no. 1, pp. 27–34, 2007.

[206] G. Schlesinger, "The mechanical structure of the artificial limbs," in Replacement links and work aids, pp. 321–661, knight, 1919.

[207] N. Kamakura, M. Matsuo, H. Ishii, F. Mitsuboshi, and Y. Miura, "Patterns of static prehension in normal hands," The American journal of occupational therapy, vol. 34, no. 7, pp. 437–445, 1980.

[208] C. Sollerman, "The use of eight main hand grips in activities of daily living," Handens greppfunktion, Thesis, 1980.

[209] T. Feix, J. Romero, H.-B. Schmiedmayer, A. M. Dollar, and D. Kragic, "The grasp taxonomy of human grasp types," IEEE Transactions on human-machine systems, vol. 46, no. 1, pp. 66–77, 2015.

[210] A. Viau, A. G. Feldman, B. J. McFadyen, and M. F. Levin, "Reaching in reality and virtual reality: a comparison of movement kinematics in healthy subjects and in adults with hemiparesis," Journal of neuroengineering and rehabilitation, vol. 1, no. 1, pp. 1–7, 2004.

[211] H. Sveistrup, "Motor rehabilitation using virtual reality," Journal of neuroengineering and rehabilitation, vol. 1, no. 1, pp. 1–8, 2004.

[212] R. R. Kaliki, R. Davoodi, and G. E. Loeb, "Evaluation of a noninvasive command scheme for upper-limb prostheses in a virtual reality reach and grasp task," IEEE Transactions on Biomedical Engineering, vol. 60, no. 3, pp. 792–802, 2012.

[213] F. Grimm, G. Naros, and A. Gharabaghi, "Closed-loop task difficulty adaptation during virtual reality reach-to-grasp training assisted with an exoskeleton for stroke rehabilitation," Frontiers in neuroscience, vol. 10, p. 518, 2016.

[214] A. S. Merians, H. Poizner, R. Boian, G. Burdea, and S. Adamovich, "Sensorimotor training in a virtual reality environment: does it improve functional recovery poststroke?," Neurorehabilitation and neural repair, vol. 20, no. 2, pp. 252–267, 2006.

[215] M. F. Levin, E. C. Magdalon, S. M. Michaelsen, and A. A. Quevedo, "Quality of grasping and the role of haptics in a 3-d immersive virtual reality environment in individuals with stroke," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 23, no. 6, pp. 1047–1055, 2015.

[216] M. P. Furmanek, L. F. Schettino, M. Yarossi, S. Kirkman, S. V. Adamovich, and E. Tunik, “Coordination of reach-to-grasp in physical and haptic-free virtual environments,” Journal of neuroengineering and rehabilitation, vol. 16, no. 1, pp. 1–14, 2019.

[217] A. M. Plooy and J. P. Wann, “Supporting natural prehension in virtual environments,” in Stereoscopic Displays and Virtual Reality Systems VII, vol. 3957, pp. 362–368, SPIE, 2000.

[218] F. Buttussi and L. Chittaro, “Locomotion in place in virtual reality: A comparative evaluation of joystick, teleport, and leaning,” IEEE transactions on visualization and computer graphics, vol. 27, no. 1, pp. 125–136, 2019.

[219] A. D. Blaga, M. Frutos-Pascual, C. Creed, and I. Williams, “Too hot to handle: An evaluation of the effect of thermal visual representation on user grasping interaction in virtual reality,” in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–16, 2020.

[220] B. Berki, “Experiencing the sense of presence within an educational desktop virtual reality,” Acta Polytechnica Hungarica, vol. 17, no. 2, pp. 255–265, 2020.

[221] M. Feick, N. Kleer, A. Tang, and A. Krüger, “The virtual reality questionnaire toolkit,” in Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology, pp. 68–69, 2020.

[222] S. Lackey, J. Salcedo, J. Szalma, and P. Hancock, “The stress and workload of virtual reality training: the effects of presence, immersion and flow,” Ergonomics, vol. 59, no. 8, pp. 1060–1072, 2016.

[223] F. Chen, Y.-C. Lin, J.-W. Chien, and C.-E. Tsai, “Virtual reality for digital user experience and interactive learning based on user satisfaction: A pilot study,” in 2016 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 374–377, IEEE, 2016.

[224] L. Sassatelli, M. Winckler, T. Fisichella, A. Dezarnaud, J. Lemaire, R. Aparicio-Pardo, and D. Trevisan, “New interactive strategies for virtual reality streaming in degraded context of use,” Computers & Graphics, vol. 86, pp. 27–41, 2020.

[225] D. Abdlkarim, M. Di Luca, P. Aves, S.-H. Yeo, R. C. Miall, P. Holland, and J. M. Galea, “A methodological framework to assess the accuracy of virtual reality hand-tracking systems: A case study with the oculus quest 2,” BioRxiv, pp. 2022–02, 2022.

[226] Oculus VR, Inc.", “Oculus Device Specifications | Oculus Developers,” 2020.

Appendix A

Publications in Chapter 2

I: Hameed, Asim, and Andrew Perkis. "A Holistic Quality Taxonomy for VR Experiences" *Submitted to Frontiers in Virtual Reality*, 2024: (in review).

II: Hameed, Asim, and Andrew Perkis. "Authenticity & Presence: Assessing the Quality of VR Experiences." *Frontiers in Psychology*, 2024, 15: 1291650.

III: Hameed, Asim, and Andrew Perkis. "Spatial storytelling: Finding interdisciplinary immersion." in *11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5–8, 2018, Proceedings 11,* pp. 323-332. Springer International Publishing, 2018.

IV: Hameed, Asim, Shafaq Irshad, and Andrew Perkis. "Towards a quality framework for immersive media experiences: a holistic approach." in *12th International Conference on Interactive Digital Storytelling, ICIDS 2019, Little Cottonwood Canyon, UT, USA, November 19–22, 2019, Proceedings 12,* pp. 389-394. Springer International Publishing, 2019.

I	is awaiting publication, included here but not available online
II	is available at https://doi.org/10.3389/fpsyg.2024.1291650
III	is available at https://doi.org/10.1007/978-3-030-04028-4_35
IV	is available at https://doi.org/10.1007/978-3-030-33894-7_41

frontiers

A Holistic Quality Taxonomy for Virtual Reality Experiences

Asim Hameed [1,*], **Sebastian Möller** [2], **and Andrew Perkis** [1]

[1] *Department of Electronic Systems, Norwegian University of Science & Technology (NTNU), Norway.*

[2] *Quality & Usability Lab, Technische Universität Berlin (TUB), Germany.*

Correspondence*:
Corresponding Author
asim.hameed@ntnu.no

frontiers | Frontiers in Psychology

TYPE Perspective
PUBLISHED 19 April 2024
DOI 10.3389/fpsyg.2024.1291650

OPEN ACCESS

EDITED BY
Pierfelice Rosato,
University of Bari Aldo Moro, Italy

REVIEWED BY
Wolfgang Leister,
Norwegian Computing Center, Norway

*CORRESPONDENCE
Asim Hameed
✉ asim.hameed@proton.me

RECEIVED 09 September 2023
ACCEPTED 26 March 2024
PUBLISHED 19 April 2024

CITATION
Hameed A and Perkis A (2024) Authenticity and presence: defining perceived quality in VR experiences. *Front. Psychol.* 15:1291650.
doi: 10.3389/fpsyg.2024.1291650

Authenticity and presence: defining perceived quality in VR experiences

Asim Hameed* and Andrew Perkis

Department of Electronic Systems, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

This work expands the existing understanding of quality assessments of VR experiences. Historically, VR quality has focused on presence and immersion, but current discourse emphasizes plausibility and believability as critical for lifelike, credible VR. However, the two concepts are often conflated, leading to confusion. This paper proposes viewing them as subsets of authenticity and presents a structured hierarchy delineating their differences and connections. Additionally, coherence and congruence are presented as complementary quality functions that integrate internal and external logic. The paper considers quality formation in the experience of authenticity inside VR emphasizing that distinguishing authenticity in terms of precise quality features are essential for accurate assessments. Evaluating quality requires a holistic approach across perceptual, cognitive, and emotional factors. This model provides theoretical grounding for assessing the quality of VR experiences.

KEYWORDS

virtual reality (VR), user experience, user-perceived quality, presence, plausibility, believability, authenticity

1 Introduction

Virtual reality (VR) is historically preoccupied with delivering realistic and immersive experiences, seamlessly transporting us into immersive worlds that blur the lines between the real and the virtual along the virtuality continuum (Milgram et al., 1995). It belongs to a range of emerging technologies that generate omnidirectional extended reality (XR) experiences for users. These are either mixed reality (MR) technologies that overlay digital images and information on the physical context or create a new reality by completely occluding the natural context, like VR (LaValle, 2016). In the past, the focus has remained on the perceived quality of VR, namely *presence* and *immersion* (Lombard and Ditton, 1997; Nilsson et al., 2016). Recently, we see the discourse expand that scope to include *plausibility* and *believability* as crucial judged quality aspects of VR experiences (Slater, 2018; Weber et al., 2021). These terms describe virtual environments in their lifelikeness, whose behavior makes sense and allows one to suspend disbelief. Both terms are interchangeably used to comment on the credibility of a VR experience in the degree to which a VR environment adheres to rules, constraints, and logic that harmonize with what users expect (Skarbez, 2016). Similarly, other terminologies, such as "coherent" and "congruent" also come up to describe the predictability and consistency of features and behaviors within virtual worlds. This work focuses on users' subjective judgments of a VR experience's credibility—referred to as its *authenticity*. We recognize that assessing the experiential quality of VR is not a one-dimensional task and concepts like plausibility and believability require invested research. It is also important to highlight that these concepts must not be reduced to singular notions of coherence or realism alone. At the same time, we believe that the use of both terms, though critical, often blurs at the edges, giving rise to conceptual confusion. Within this paradigm, we ask: how can the concepts of plausibility and believability be clearly defined and differentiated within the broader notion of authenticity in VR experiences? Further, what roles do coherence and congruence play in complementing plausibility and believability to evaluate the overall quality of VR experiences? To this end, the paper will first separate the two terms and accurately outline their differences and connections. Secondly, we propose a structured hierarchy that defines plausibility and believability as subsets of the overarching concept—authenticity. Through this, we hope to delineate the boundaries and intersections of these terms. Finally, we introduce coherence and congruence as quality functions complementing plausibility and believability. This dynamic interplay underscores the importance of considering both the internal and external logic of a VR experience and the alignment of its stimulus to the users' perceptions and experiences in evaluating its overall quality. The proposed authenticity paradigm integrates previous frameworks on presence, realism, and plausibility. We synthesize these perspectives into a cohesive structure that can guide the analysis and design of high-quality VR experiences. Adopting a nuanced perspective that approaches authenticity and presence as experienced quality can enhance theoretical clarity and provide stronger empirical grounding for studying user experiences in VR.

2 VR—Realistic, plausible, and believable

When assessing the overall quality of VR experiences, we have had a historical preoccupation with *realness* or *realism*. Realism in VR expands from the fidelity of available stimuli to the perception of how closely a virtual environment (VE) imitates the real world (Alexander et al., 2005; Lombard et al., 2009). VR experiences are commonly assessed in terms of two crucial dimensions: presence and immersion (Schuemie et al., 2001; Biocca, 2002; McMahan, 2003; Lombard et al., 2009; Slater, 2018). The richness of the VE profoundly influences both of these facets—its visual, aural, and fidelity—which play a pivotal role in captivating users and enhancing their sense of immersion (Steuer, 1992). Engaging a user with rich and exclusive sensory stimulation inside a head-mounted display (HMD) achieves a sense of *presence*—an objective property of the system (Bowman et al., 2012)—associated with a vivid sense of being "there" in the virtual world, interacting with virtual objects, engaging with virtual characters, and feeling emotions within the simulated world. The prevailing discourse in VR has often leaned heavily on the prominence of presence (Kim and Biocca, 1997; Lee, 2004) as the primary construct of a subjective experience of feeling transported into a virtual world. It is a psychological state influenced by the user's expectations, beliefs, and experiences. Immersion (Witmer and Singer, 1998; McMahan, 2003) meanwhile are the technological (or system) aspects that surround the user, as mentioned before. It is the extent to which any user would feel absorbed in the virtual world owing to its ability to produce and render scenarios and experiences with a high degree of realism (visual and audio fidelity), responsiveness (interactive fidelity), embodiment (sensorimotor stimulation and feedback) (Steuer, 1992; Baños et al., 2004; Kilteni et al., 2012). Presence has long been considered the defining quale of VR; however, an overemphasis risks overlooking other critical elements of the overall user experience. As VR technology advances and its applications expand, it becomes increasingly evident that presence alone is an insufficient framework to capture the richness and complexity of VR experiences fully.

Multiple other works exploring complementary phenomena influencing VR experiences share this point of view. Earlier on, Slater (2009) conceptualized a theoretical framework with two orthogonal components, namely *place illusions* and *plausibility illusion*. Place illusion denoted presence, while the additional plausibility illusion referred to the realism and likelihood of a VR scene. In their terms, "the overall credibility of the scenario being depicted" juxtaposed with user expectations, delivering an impression that the system-generated events were occurring. Later iterations of the concept have used both the term plausibility (Rovira et al., 2009; Hofer et al., 2020) as well as other classifications for the same theoretical principle; reality judgment (Baños et al., 2000), perceived realism (Lombard and Ditton, 1997; Schubert et al., 2001), coherence (Skarbez et al., 2017), and authenticity (Gilbert, 2016). These works view plausibility as a higher-order cognitive operation that involves a judgment on the credibility or authenticity of the VR scene, which is reflected by its consistency and the extent to which it meets a user's expectations.

Looking in detail at plausibility is essential to differentiate between various quality aspects of the phenomenon. For Skarbez (2016), this translates to when a VE projects situations that appear apparent to the users based on their existing knowledge of the world. Such knowledge can include their understanding of both the real world and their knowledge of the fictional world depicted inside VR. Internal plausibility is how well it follows its rules and makes sense within its framework. External plausibility is how consistent it is with real-world knowledge and whether it matches a user's understanding of the real world (Busselle and Bilandzic, 2008; Hofer et al., 2020). An updated review, published by Slater et al. (2022), added depth to their initial conceptualization by specifying different instances of plausibility inside VR: a reactive environment that responds to actions, contingent interactions that happen in relation to the user, and coherence with users' expectations based on their experiences and knowledge. A more recent contribution by Latoschik and Wienrich (2022) looks at plausibility alongside congruence—how we feel about the experience and how well it matches our expectations. Their model considers congruence as the objective match between the information processed by the user and their expectations at the sensory, perceptual, and cognitive levels. Plausibility results from the evaluation of congruence across the three levels. Sensory congruence is how well the experience matches our senses. Perceptual congruence is how well the experience fits our understanding of how the world functions. Finally, cognitive congruence is how well the experience matches our beliefs and expectations. Weber et al. (2021) have identified plausibility under the concept of *perceived realism*, which extends to (1) the realism of objects, sounds, and scenes in terms of their congruence to real-world textures, proportions, details, etc. (2) the plausibility of story and characters, evaluating their consistency rather than factual accuracy, and (3) judgment about the naturalness of interactions.

Another term often interchangeably used with plausibility is that of *believability*. Closely related to the historic literary notion of the "suspension of disbelief" on the part of the audience/reader to suspend their judgment concerning the implausibility of a given narrative (Chandler and Munday, 2011). The idea that audiences are willing to accept the premises of a fictional work, even if they are fantastical or unrealistic, as long as that world and its characters feel subjectively accurate and coherent enough. For VR, the suspension of disbelief is essential for creating believable experiences. If the virtual world is believable enough, users will accept its artificiality and immerse themselves in it, just as a reader would in a fictional story. Sheridan referred to it as "the active imagination in suppressing disbelief (and thus enhanced believability)" (Sheridan, 2000). Believability is also defined as elements operating at various levels of realism—sensory, perceptual, and emotional—manifested through realistic visual and aural effects, a consistent VE that allows natural interactions, as well as aesthetic, dramaturgical, and emotional aspects of the VR experience (Magnenat-Thalmann et al., 2005; Papagiannakis et al., 2005; Bogdanovych et al., 2015).

We recognize the significant contribution of the frameworks and models described in the previous section. Concurrently, we recognize the necessity of consistently refining concepts to enhance clarity, especially because using broad and repetitive terminologies adds uncertainty to quality assessments. We believe maintaining distinct terms for plausibility and believability is necessary to explain fully the characteristics and influences shaping VR experiences. This distinction also aligns with semiotic principles, given that the elements outlined in the frameworks and models discussed previously correspond to separate semantic and syntactic categories (Barricelli et al., 2016). Therefore, an explication using precise language for describing quality aspects and avoiding confusion between key constructs and factors is important. Further, we agree with the contention that fixating solely on presence does not encapsulate the multifaceted nature of VR experiences (Gilbert, 2016). In the following section, we propose a recalibration of focus toward plausible and believable VR experiences that we present as subsets of a quality model for *authenticity*.

2.1 The Presence–Authenticity Dyad

Gilbert (2016) described authenticity as how well the VR environment mirrors the expected regularities of the world it is trying to represent (Bowman et al., 2012). How faithfully does it replicate the behaviors, relationships, and rules consistent with its purported context? How closely an entity aligns with an individual's expectations, cognitive schemas, prior knowledge, personal experiences, preferences, and interaction reciprocity (Bucolo, 2004; Weibel et al., 2010). The longer a user stays in a VR environment, the more likely their initial sense of wonder will give way to a heightened awareness of the environment's authenticity. Once familiarized, users begin to notice incongruities in the VR setting. For instance, the inability to physically interact with virtual objects in an intuitive way (Hameed et al., 2021), the failure of non-player characters to respond to the user's existence (Rovira et al., 2009), or a disjunction between the realism of the user's avatar and the aesthetic of the world they inhabit (Slater, 2017). This shift from an initial enchantment to a heightened critical awareness of its features reflects the various quality aspects that influence the assessment of a VR experience. It highlights that while a robust place illusion is necessary, it may prove shallow and lose its spell if the virtual world gives the impression of being inauthentic.

Considering this, we introduce the "Presence-Authenticity Dyad," recognizing authenticity as a complementary dimension to presence and crucial for evaluating the quality of VR experiences. In similar vocabulary to that which characterizes presence as a feeling of "being there," this work defines authenticity as *"a sense of 'trueness and genuineness'"* felt in a virtual place.

In agreement with Lee (2020), we see it as users' individual judgment on the virtual world's trueness and genuineness regarding its stimuli, content, and behavior. We expand this to include two subtypes of believability and plausibility. Despite their interchangeable use, they refer to related but distinct characteristics in the virtual. Both contribute to overall authenticity and presence. We define them as follows:

- *Plausibility* is the extent to which a VR experience can be logically explained and remains consistent with real-world principles. What's happening is real. It refers to the degree to which the VR environment and its contents exhibit logical congruence and follow common sense. For example, perceptual constancy, the consistency of its physics, etc. Plausibility operates at the syntactic level, reflects in logical consistency, and has more objective thresholds. It reflects the trueness of the depicted world. An experience is plausible if the environment and its contents remain rational and conform to the principles of its rules-based reality.
- *Believability* is how much a VR can deliver an experience with the realism and internal coherence required to make it feel believable for the user. It goes beyond mere visual fidelity and taps into the user's emotions, senses, and overall engagement with the virtual world. If it is convincing, it's happening. Factors include narrative logic, engaging gameplay, etc. Believability carries semantic elements, includes emotional resonance, and aligns variably based on the subjective perceptions of the user. It reflects the genuineness of the depicted world in its subtle details and nuances that mimic reality and support a "suspension of disbelief," even if the experience itself is fantastical or fictional.

2.2 A quality interpretation of authenticity

In line with the notion that presence constitutes a subjective sensation contingent upon the immersive attributes of a system, it becomes evident that a system must first facilitate immersion to establish the semblance of "being there." In a parallel vein, one can argue that authenticity reinforces the illusion engendered by the VR system, thereby influencing the efficacy of a VR experience.

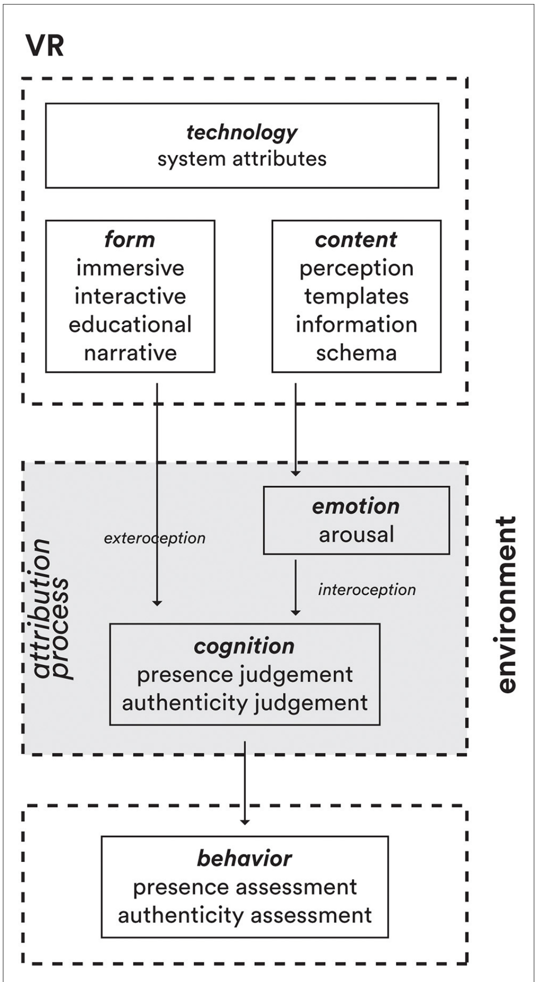


FIGURE 1
The attribution process leading to judgment inside VR, adapted from Diemer et al. (2015).

Presence and authenticity are two distinct facets of a VR experience, with multiple quality aspects underpinning the two phenomena (see Figure 1).

2.2.1 On quality formation

Quality involves perceptual, cognitive, emotional, and evaluative processes determining one's conscious perception of things. The word stems from qualia—qualitative, phenomenal aspects of consciousness and subjective experience (Gregory, 1996)—in the form of sensations, feelings, and mental imagery that reflect what it is like to experience something. They may include impressions of goodness, beauty, desirability, and virtue that arise in consciousness when encountering something (Shoemaker, 1990). Jekosch (2005) refers to *experienced quality* as the result of a mental evaluation where someone compares the actual composition of something to their expected or ideal composition. In other words, experienced quality is the subjective judgment of how well an entity's perceived composition aligns with desired expectations (Blauert and Jekosch, 2012). The term "entity" denotes any object or event, material or immaterial, that becomes an object of perception.

This may be straightforward in the physical realm where entities have objective physical attributes (or *quality elements*) that can be measured, e.g., display resolution. But in the subjective realm of perception, entities exhibit psychological features (or *quality features*) such as vividness and richness. A subject (user) perceives an entity's quality features and compares these to their internal ideals and expectations, which shapes a conscious impression of the entity's overall quality (Uhrig, 2021). Möller (2023) categorizes quality elements and features into complementary factors and aspects. For example, technical factors like throughput and jitter can affect perceptual aspects like immersion and embodiment.

Jekosch (2005) reformulations further identify a fundamental distinction between two facets of quality perception: *perceived quality* and *judged quality* (Uhrig, 2021). *Perceived quality*—akin to low-level thinking—is an immediate impression formed upon encountering a stimulus. Damasio (1995) argue that such a swift quality assessment upon encountering a natural environment or a technological stimulus need not require deep cognitive processing but results from integrating basic perceptual features into an abstract evaluation. For example, a user may perceive the quality of a depicted scene as poor upon first seeing it but without consciously analyzing why. Despite being an initial reflexive impression, perceived quality can still be intentionally contemplated and judged after the fact. This evaluative process produces a quality judgment that reflects cognitive analysis. The subjective experience of this quality judgment is termed the *judged quality*. Since cognitive evaluation activates complex associations and interpretations, judged quality encompasses richer perceptual, conceptual, and affective content than perceived quality. It is influenced by conscious analysis and reflection, not just direct perception.

The distinction between perceived quality and judged quality relates to the notion of experience, which can be defined as "the stream of perceptions (of feelings, sensory percepts, and concepts)" that occur in a given situation (Möller and Raake, 2014). In this respect, perceived quality aligns with the immediate experience of impressions and sensations within the VR world. It is an intuitive and phenomenological part of experiencing that world. However, judged quality goes beyond just experiencing. It requires additional cognitive processing to evaluate and consciously judge the quality of the depicted VR world (see Figure 2). So, while perceived quality is embedded in the direct experience, judged quality emerges from reflective analysis and interpretation of the VR experience. The former is instant, while the latter involves extra mental effort to reach an overall quality assessment.

2.2.2 Quality aspects of authenticity

Earlier in this paper, we defined authenticity as the trueness and genuineness of the displayed VR place. As evident, both words

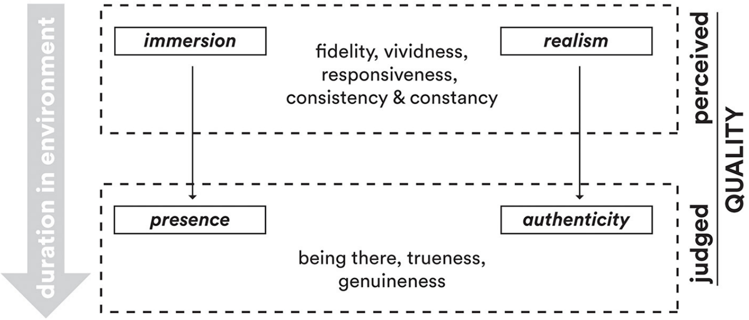


FIGURE 2
Judgment process from perceived quality to judged quality. Immersion-related attributes give rise to a sense of presence, whereas perceived attributes of consistency and constancy (or realism) foster an authenticity about the experience.

imply a deliberate judgment of the depicted place on the user's part. This is unlike a user's more immediate, intuitive impression of the VR setting's various sensory inputs and atmospherics. In fact, upon first encountering, users will be relying on sensory percepts to discern the visuals, sounds, and vectors available in the VR space. However, initial quality perceptions simultaneously evolve as they are compared to quality features internally desired by the user (expectations). We argue here that even at the most nascent stages of their embodied encounter, the *experienced quality* of the virtual world is enough to imbue a feeling of presence (ephemeral as it may be). The longer the visitors stay in the immersive VR world from here on, the more their awareness is heightened with respect to the lifelikeness, interactive intuitiveness, and audio-visual synchronicity, etc. of the VR setting. Quality judgments on trueness and genuineness entail conscious assessments of the virtual world's congruence and coherence and reflect this heightened state of intentional and reflective cognitive processing. Both quality descriptions go beyond initial impressions to include complex and nuanced evaluations of whether the VE maintains its integrity, i.e., credibility. These judgments pertain to perceptions and desires but also carry emotional and evaluative dimensions.

Trueness is defined as "conformity to reality and actuality" or "agreement to fact and reality" (Webster, 2014). It is focused on the accuracy of the information following reality and reflective of facts. Conversely, genuineness is defined as "the quality of being honest and sincere" and "the quality of being real and exactly what it appears to be" (Webster, 2014). It goes beyond mere accuracy and delves into sincerity. It encompasses the quality of being real and without pretense. In terms of determining experiential quality in VR, authenticity must then be understood as the sum of the factual accuracy of the world as well as the sincerity of its self-expression. Its trueness is evidence-based (objective), whereas its genuineness is internally driven (subjective).

In terms of a VR experience, both trueness and genuineness are distinct quality features of the authenticity of that experience. As such, we associate them with the quality aspects that determine authenticity, i.e., plausibility and believability, respectively. Trueness speaks to the plausibility of a VR experience and genuineness reflects its believability. Table 1 charts the differentiation of quality goals for the two aspects, the factors influencing them, and some evaluation methods to assess them. Moreover, we refer to the terms *congruence* and *coherence* as functions of the two quality aspects that specify either the fulfillment or nonfulfillment of authenticity. We ascribe the term coherence to believability and congruence to plausibility. The former describes an inner connectedness or integration of meaning within something, while the latter refers to an alignment or matching between two or more things (virtual-to-real). Our appropriation of both terms is consistent with how they regularly appear in VR research. Most definitions of coherence relate to Skarbez et al. (2017), who have referred to it as the internal consistency of a virtual experience and defined it "as the set of reasonable circumstances that the scenario can demonstrate without introducing unreasonable circumstances." How well the parts of something fit together logically. A coherent experience should reflect consistency through its story, visuals, sounds, characters, tasks, etc. Its parts must understandably indicate a unified whole, with ideas that make sense together. Correspondingly, the term congruence has been borrowed from environmental psychology to depict an agreement or consistency between things and is defined as "the degree to which different cues fit with each other" or "a similarity between perceptual variables" (Maffei et al., 2016; Flavián et al., 2021). Congruence may carry the processing of physical and relational information reflected in matching the logic, physical behaviors, and limitations within a virtual experience.

Our revised understanding of authenticity in VR suggests that while a user's initial engagement may stem from a feeling of presence within a computer-generated environment, a lasting impact of the VR experience hinges on its ability to instill a sense of genuineness and trueness within the virtual world (see Figure 3). It must carry qualities that inspire belief in its meaning and truth to sustain immersion. More than its illusions, a VR experience must resonate in its essence and significance to the user.

3 Discussion

In this paper, we explored the multifaceted nature of quality assessment for immersive VR experiences by drawing attention

TABLE 1 Key objective and subjective determinant for evaluating *plausibility* and *believability* in VR.

	Plausibility	*Believability*
Definition	Adheres to real-world principles, feels rational and explainable; trueness	Resonates with perceptions and emotions, feels subjectively "real"; genuineness
Function	Syntactic; Logical congruence	Semantic; Internal coherence
Descriptives	Perceptual constancy of objects Physics consistency Logical cause-and-effect chains Multi-sensory alignment Visual grounding Situated acoustics Fast interactive responsiveness	Scenario logic Atmospherics and randomness Narrative and Stylistic cohesion Environmental imperfections Nuanced reactions Resonance with memories and emotions Subjective presence
Factors	Stable geometry and optimized models Unrealistic forces and behavior Penetrations and incorrect scaling Audio synchronization and acoustics Lighting/shadow matches source Assets situated logically	Physically based rendering PBR materials and textures Detailed assets/expressive characters Subtle environmental cues Realistic audio sampling Natural conversational flow
Evaluations	Quality Metrics to evaluate 3D models Metrics for physics simulations Test logical contradictions Examine against physical rules Check sensory alignments Detect affordance mismatches	User testing and feedback Track user behavior Monitor user performance Assess emotional responses Review ecological realism Survey narrative realism and disbelief

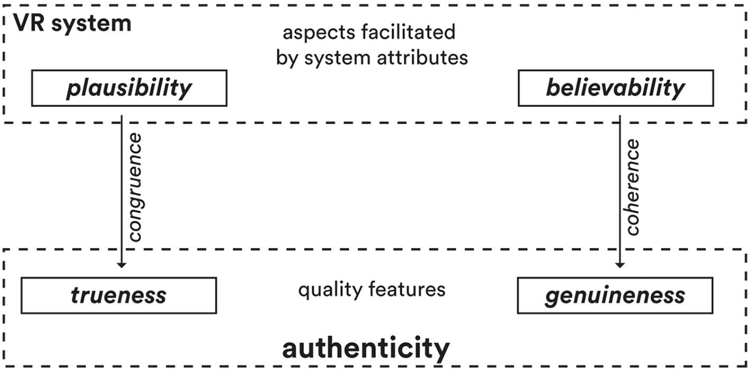


FIGURE 3
Quality formation of experienced *authenticity* inside VR.

to a conceptual distinction between perceived quality and judged quality. We proposed identifying authenticity as a key dimension of quality perception complementary to the feeling of presence. Existing literature on quality assessments of VR experiences emphasizing presence has often overlooked authenticity. This has led to multiple conceptualizations and questionnaires that remain preoccupied with system factors facilitating immersion and generating a one-time sense of "being there." There is a need to explicitly differentiate between realism as the fidelity and richness of the mediated environment vs. authenticity as the trueness and genuineness of virtual worlds. Clearly distinguishing these as two quality facets will allow for more precise definitions and measurement instruments.

To this end, we distinguished plausibility (trueness) and believability (genuineness) as distinct yet complementary aspects contributing to a VR experience's overall authenticity. Plausibility refers to the objective, logical congruence of the virtual world in adhering to real-world principles, natural laws, and common sense rationality. It operates at a syntactic level, reflected in consistencies like perceptual constancy of objects, accurate physics simulations, and logical cause-and-effect chains unfolding within the environment. In contrast, believability is more subjective, relating to how genuinely "real" the experience feels to an individual user based on their personal perceptions, prior experiences, and evoked emotions. While plausibility entails maintaining objective rules and realism, believability hinges on semantic details, stylistic nuances, and resonant engagement that suspends disbelief and facilitates immersive psychological involvement, even if the content is fantastical or imaginative. Thus, plausibility cues are more binary while believability varies across users.

It is important to highlight that VR experiences need not always mirror real-life scenarios. Experiences could involve unrealistic, fictional, or imaginative elements. Yet if these elements interact with the user congruently and coherently, they can feel authentic. For example, a virtual world that simulates real-life settings must meticulously adhere to real-world nuances and principles. In such a context, the VE should respect the laws of gravity, ensuring that objects behave as they would in the physical world. Conversely, objects may defy gravity in VR to provide a fantastical experience in a zero-gravity environment. Since such unnatural defiance aligns with the intended narrative, it will be acceptable in that depicted world. These flights of imaginative engagement encourage a willful “suspension of disbelief,” which may be construed as a momentary recalibration of one’s preconceived notions. Within this framework, individuals can momentarily adopt cognitive predispositions that harmonize with the fictitious realms they are immersing themselves in. This cognitive adaptability allows users to traverse and comprehend various VR experiences, from the meticulously realistic to the purely fantastical. This helps them appreciate the diversity of content and modalities within VR. Users bring their prior beliefs, but once inside the VE, new sensory input is integrated with these priors to update their beliefs, which influences their perception of the environment’s realism (Triantafyllou et al., 2014; Gilbert, 2016). For example, a user entering a virtual forest will compare the sensory input (like the appearance and sounds) with the priors of a real forest. If it aligns, the virtual forest will maintain its authenticity. The need for authenticity in VR extends to the consistency of interactions, relationships, and elements within the virtual space. If users perceive inconsistencies, mismatches in coherence, or behaviors that contradict their expectations, their sense of authenticity can be disrupted (Biocca and Delaney, 1995). This could, in effect, lead to a break-in-presence or a decrease in the overall quality of the VR experience.

Evaluating quality necessitates a holistic approach spanning perceptual, cognitive, and emotional factors. As users spend more time immersed in a virtual environment, perceived quality gives way to judged quality as inconsistencies become apparent. Achieving high-quality VR experiences involves optimizing both low-level processing and higher-level functions of congruency and coherence assessing events and interactions within the VE. Adopting a nuanced perspective that approaches authenticity and presence as experienced quality can enhance theoretical clarity and provide stronger empirical grounding for studying user experiences in VR. Below, we extend this discussion to briefly describe various technical and human-centric factors influencing plausibility and believability.

Evaluating plausibility

The technical factors contributing to a positive and immersive experience in VR are perceptual constancy, aliasing and sampling, audio synchronization, and physics consistency. Perceptual constancy ensures that objects maintain their appearance despite changes in environmental conditions (Coren et al., 2004; Jerald, 2015) while aliasing and sampling reduce visual artifacts like jagged edges and pixelated textures (Gibson and Mirtich, 1997; Lessiter et al., 2001). Audio synchronization improves the authenticity of the aural experience (Guastavino et al., 2007), while physics consistency requires emulating real-world scenarios with physics engines that behave realistically (Hummel et al., 2012). One suggestive evaluation approach uses quality metrics to assess 3D models and physics simulations based on their real-world physical properties and material types. Another recommendation is to evaluate how well the system adheres to established rules and cause-and-effect relationships within the defined world logic. This evaluation should be examined against physical rules and check sensory alignments (Chen et al., 2019). Additionally, it is suggested to use metrics such as collision detection, object interactions, and gravity behavior to assess the accuracy and realism of physics simulations in the virtual environment (Jiang et al., 2018). Lastly, tracking object interaction frequency and accuracy can help identify instances of affordance mistakes and analyze control mechanics to improve user experience (Hameed et al., 2021). Subjective measures involve questionnaires that assess how realistically users perceive the virtual world and how well it aligns with their prior expectations of similar environments (Regia-Corte et al., 2013). Self-reported measures can be employed to investigate emotional responses to implausible or nonsensical events. The overall pleasantness and engagement of the virtual experience can be assessed through questionnaires and surveys, gathering user feedback on their positive and negative affective responses to the features and elements within the VR environment (Möller et al., 2013; Hameed et al., 2023).

Evaluating believability

One of the crucial aspects of believability remains the use of high-fidelity stimuli, which includes various features such as render quality, physics engine, and spatial audio (Skarbez, 2016; Slater et al., 2022). In addition, internal coherence and consistency are essential, which means that all elements within the virtual world should make sense and be consistent with the established setting and rules (Lepecq et al., 2009). Details that reflect real-world experiences, such as environmental imperfections and character animations, can significantly enhance the feeling of naturalness within the environment (Loomis, 2016). Moreover, the complexity and realism of scripted events or narratives in the virtual world should remain logically consistent, and users should anticipate what comes next (Llobera et al., 2013; Skarbez et al., 2020). The sense that actions and experiences within the virtual world have value or significance also adds to their meaningfulness, which can heighten users’ cognitive absorption and emotional engagement (Murray et al., 2007; Beckhaus and Lindeman, 2011). To evaluate virtual assets, animations, and environments, use industry benchmarks and standards (Otto et al., 2019). Measure the world’s size, complexity, and dynamism with metrics like the number of environments, objects, and paths (Lugrin et al., 2013). Assess the level of detail and use of sound effects to enhance the virtual world’s believability (Tran et al., 2021). A human-centric factor that contributes to the believability of a virtual world is the suspension of disbelief, which refers to users’ willingness to temporarily accept the virtual world as real even though they know it’s not (Karhulahti, 2012). This can be achieved through immersive storytelling, visuals and audio, and minimal technical glitches.

Another important factor is narrative immersion and involvement, where users feel emotionally invested in the virtual world's characters, story, or situations. This emotional connection can be fostered through relatable characters, meaningful interactions, and engaging narratives (Rollings and Adams, 2003; Ryan, 2009). A user's prior VR experience can also affect their perception of realism and believability, with those with more experience having higher expectations for these qualities. Additionally, individuals with vivid imaginations and susceptibility to suggestion are more accepting of realistic and fictional VR experiences (Gilbert, 2016).

Further, several research lines can be pursued to examine the validity and refine the proposed model. Conducting user studies to empirically validate the assumptions about how the proposed quality aspects (plausibility, believability, coherence, congruence) contribute to perceived authenticity and overall quality judgments in VR experiences. Developing standardized scales and questionnaires to quantify and measure the different quality components outlined in the authenticity model (such as perceived realism, logical consistency, emotional resonance, and suspension of disbelief) would also be valuable. Designing controlled experiments systematically manipulating specific variables (e.g., physics accuracy, narrative logic, sensory alignments) to measure their impact on users' perceptions of plausibility, believability, and overall authenticity. Employing multimodal data collection combining subjective reports, behavioral tracking, physiological sensing, and qualitative interviews can capture the multidimensional nature of authenticity assessments. Finally, cross-domain evaluations can help assess the model's applicability and identify potential domain-specific nuances across various VR application areas like training, gaming, therapy, social VR, etc.

4 Conclusion

This work puts forth several key findings and contributions. It proposes authenticity as a complementary dimension to presence in evaluating the quality of VR experiences. It argues that while presence focuses on the sense of "being there," authenticity captures the sense of "trueness and genuineness" felt in the virtual place. It distinguishes between plausibility (adhering to real-world principles, reflecting trueness) and believability (resonating with user perceptions/emotions, capturing genuineness) as two key aspects of authenticity. The paper introduces a structured hierarchy that defines plausibility and believability as subsets under the broader umbrella of authenticity. It also positions coherence and congruence as complementary quality functions related to internal logic (believability) and external mapping (plausibility) respectively. Furthermore, it highlights the importance of considering perceived quality (immediate impressions) and judged quality (reflective evaluations) when assessing authenticity in VR experiences. The work provides a theoretical grounding for holistically evaluating authenticity by spanning perceptual (e.g., graphics, physics), cognitive (e.g., logical consistency, narrative), and emotional (e.g., engagement, resonance) factors.

The present contribution proposes a theoretical model rather than providing empirical validation. While empirical testing is outside the scope of this work, the quality framework for authenticity presented here can inform future research into assessments of VR experience. Comparing the authenticity model against existing frameworks for presence, immersion, realism, etc., to delineate conceptual boundaries and explore potential integrations would be insightful. Also, there is good potential for using empirical findings to iteratively refine and expand the proposed model's theoretical foundations. This paper puts forth a preliminary model to spur additional research that can advance knowledge on factors shaping authentic, high-quality VR user experiences.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

AH: Writing – review & editing, Writing – original draft, Conceptualization, Methodology. AP: Writing – review & editing, Project administration.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alexander, A. L., Brunyé, T., Sidman, J., and Weil, S. A. (2005). *From Gaming to Training: A Review of Studies on Fidelity, Immersion, Presence, and Buy-in and their Effects on Transfer in PC-Based Simulations and Games, Vol. 5*. Woburn, MA: DARWARS Training Impact Group, 1–14.

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., and Rey, B. (2004). Immersion and emotion: their impact on the sense of presence. *CyberPsychol. Behav.* 7, 734–741. doi: 10.1089/cpb.2004.7.734

Baños, R. M., Botella, C., Garcia-Palacios, A., Villa, H., Perpiñá, C., and Alcaniz, M. (2000). Presence and reality judgment in virtual environments: a unitary construct? *CyberPsychol. Behav.* 3, 327–335. doi: 10.1089/10949310050078760

Barricelli, B. R., Gadia, D., Rizzi, A., and Marini, D. L. (2016). Semiotics of virtual reality as a communication process. *Behav. Inf. Technol.* 35, 879–896. doi: 10.1080/0144929X.2016.1212092

Beckhaus, S., and Lindeman, R. W. (2011). "Experiential fidelity: leveraging the mind to improve the VR experience," in *Virtual Realities: Dagstuhl Seminar 2008* (Cham: Springer), 39–49. doi: 10.1007/978-3-211-99178-7_3

Biocca, F. (2002). "The evolution of interactive media," in *Narrative Impact. Social and Cognitive Foundations*, eds M. C. Green, J. J. Strange, and T. C. Brock (Boca Raton, FL: Taylor & Francis), 97–130.

Biocca, F., and Delaney, B. (1995). "Immersive virtual reality technology," in *Communication in the Age Virtual Reality, Vol. 15*, eds F. Biocca, and M. R. Levy (London: Routledge), 10–5555.

Blauert, J., and Jekosch, U. (2012). A layer model of sound quality. *J. Audio Eng. Soc.* 60, 4–12. doi: 10.21437/PQS.2010-4

Bogdanovych, A., Trescak, T., and Simoff, S. (2015). "Formalising believability and building believable virtual agents," in *Artificial Life and Computational Intelligence: First Australasian Conference, ACALCI 2015, Newcastle, NSW, Australia, February 5-7, 2015, Proceedings 1* (Cham: Springer), 142–156. doi: 10.1007/978-3-319-14803-8_11

Bowman, D. A., McMahan, R. P., and Ragan, E. D. (2012). Questioning naturalism in 3D user interfaces. *Commun. ACM* 55, 78–88. doi: 10.1145/2330667.2330687

Bucolo, S. (2004). "Understanding cross cultural differences during interaction within immersive virtual environments," in *Proceedings of the 2004 ACM Siggraph International Conference on Virtual Reality Continuum and its Applications in Industry* (New York, NY: ACM), 221–224. doi: 10.1145/1044588.1044634

Busselle, R., and Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: a model of narrative comprehension and engagement. *Commun. Theory* 18, 255–280. doi: 10.1111/j.1468-2885.2008.00322.x

Chandler, D., and Munday, R. (2011). *A Dictionary of Media and Communication*. Oxford: Oxford University Press. doi: 10.1093/acref/9780199568758.001.0001

Chen, M., Jin, Y., Goodall, T., Yu, X., and Bovik, A. C. (2019). Study of 3D virtual reality picture quality. *IEEE J. Sel. Top. Signal Process.* 14, 89–102. doi: 10.1109/JSTSP.2019.2956408

Coren, S., Ward, L. M., and Enns, J. T. (2004). *Sensation and Perception*. Hoboken, NJ: John Wiley & Sons.

Damasio, A. R. (1995). Review: toward a neurobiology of emotion and feeling: operational concepts and hypotheses. *Neuroscientist* 1, 19–25. doi: 10.1177/107385849500100104

Diemer, J., Alpers, G. W., Peperkorn, H. M., Shiban, Y., and Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Front. Psychol.* 6:26. doi: 10.3389/fpsyg.2015.00026

Flavián, C., Ibáñez-Sánchez, S., and Orús, C. (2021). The influence of scent on virtual reality experiences: the role of aroma-content congruence. *J. Bus. Res.* 123, 289–301. doi: 10.1016/j.jbusres.2020.09.036

Gibson, S. F., and Mirtich, B. (1997). *A Survey of Deformable Modeling in Computer Graphics*. Technical report. Cambridge, MA: Mitsubishi Electric Research Laboratories.

Gilbert, S. B. (2016). Perceived realism of virtual environments depends on authenticity. *Presence* 25, 322–324. doi: 10.1162/PRES_a_00276

Gregory, R. L. (1996). What do qualia do? *Perception* 25, 377–379. doi: 10.1068/p250377

Guastavino, C., Larcher, V., Catusseau, G., and Boussard, P. (2007). *Spatial Audio Quality Evaluation: Comparing Transaural, Ambisonics and Stereo*. Atlanta, GA: Georgia Institute of Technology.

Hameed, A., and Möller, S., and Perkis, A. (2023). How good are virtual hands? Influences of input modality on motor tasks in virtual reality. *J. Environ. Psychol.* 92:102137. doi: 10.1016/j.jenvp.2023.102137

Hameed, A., Perkis, A., and Möller, S. (2021). "Evaluating hand-tracking interaction for performing motor-tasks in VR learning environments," in *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)* (Montreal, QC: IEEE), 219–224. doi: 10.1109/QoMEX51781.2021.9465407

Hofer, M., Hartmann, T., Eden, A., Ratan, R., and Hahn, L. (2020). The role of plausibility in the experience of spatial presence in virtual environments. *Front. Virtual Real.* 1:2. doi: 10.3389/frvir.2020.00002

Hummel, J., Wolff, R., Stein, T., Gerndt, A., and Kuhlen, T. (2012). "An evaluation of open source physics engines for use in virtual reality assembly simulations," in *Advances in Visual Computing: 8th International Symposium, ISVC 2012, Rethymnon, Crete, Greece, July 16-18, 2012, Revised Selected Papers, Part II 8* (New York, NY: Springer), 346–357. doi: 10.1007/978-3-642-33191-6_34

Jekosch, U. (2005). *Voice and Speech Quality Perception: Assessment and Evaluation*. New York, NY: Springer Science & Business Media.

Jerald, J. (2015). *The VR Book: Human-Centered Design for Virtual Reality*. San Rafael, CA: Morgan & Claypool. doi: 10.1145/2792790

Jiang, B., Yang, J., Jiang, N., Lv, Z., and Meng, Q. (2018). Quality assessment for virtual reality technology based on real scene. *Neural Comput. Appl.* 29, 1199–1208. doi: 10.1007/s00521-016-2828-0

Karhulahti, V.-M. (2012). "Suspending virtual disbelief: a perspective on narrative coherence," in *Interactive Storytelling: 5th International Conference, ICIDS 2012, San Sebastián, Spain, November 12-15, 2012, Proceedings 5* (Cham: Springer), 1–17. doi: 10.1007/978-3-642-34851-8_1

Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence* 21, 373–387. doi: 10.1162/PRES_a_00124

Kim, T., and Biocca, F. (1997). Telepresence via television: two dimensions of telepresence may have different connections to memory and persuasion. *J. Comput.-Mediat. Commun.* 3:JCMC325. doi: 10.1111/j.1083-6101.1997.tb00073.x

Latoschik, M. E., and Wienrich, C. (2022). Congruence and plausibility, not presence: pivotal conditions for XR experiences and effects, a novel approach. *Front. Virtual Real.* 3:694433. doi: 10.3389/frvir.2022.694433

LaValle, S. (2016). *Virtual Reality*. Cambridge: Cambridge University Press

Lee, E.-J. (2020). Authenticity model of (mass-oriented) computer-mediated communication: conceptual explorations and testable propositions. *J. Comput.-Mediat. Commun.* 25, 60–73. doi: 10.1093/jcmc/zmz025

Lee, K. M. (2004). Presence, explicated. *Commun. Theory* 14, 27–50. doi: 10.1111/j.1468-2885.2004.tb00302.x

Lepecq, J.-C., Bringoux, L., Pergandi, J.-M., Coyle, T., and Mestre, D. (2009). Afforded actions as a behavioral assessment of physical presence in virtual environments. *Virtual Real.* 13, 141–151. doi: 10.1007/s10055-009-0118-1

Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. (2001). A cross-media presence questionnaire: the ITC-sense of presence inventory. *Presence* 10, 282–297. doi: 10.1162/105474601300343612

Llobera, J., Blom, K. J., and Slater, M. (2013). Telling stories within immersive virtual environments. *Leonardo* 46, 471–476. doi: 10.1162/LEON_a_00643

Lombard, M., and Ditton, T. (1997). At the heart of it all: the concept of presence. *J. Comput.-Mediat. Commun.* 3:JCMC321.

Lombard, M., Ditton, T. B., and Weinstein, L. (2009). "Measuring presence: the temple presence inventory," in *Proceedings of the 12th Annual International Workshop on Presence* (Los Angeles, CA: International Society for Presence Research), 1–15.

Loomis, J. M. (2016). Presence in virtual reality and everyday life: immersion within a world of representation. *Presence* 25, 169–174. doi: 10.1162/PRES_a_00255

Lugrin, J.-L., Wiebusch, D., Latoschik, M. E., and Strehler, A. (2013). "Usability benchmarks for motion tracking systems," in *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology* (New York, NY: ACM), 49–58. doi: 10.1145/2503713.2503730

Maffei, L., Masullo, M., Pascale, A., Ruggiero, G., and Romero, V. P. (2016). Immersive virtual reality in community planning: acoustic and visual congruence of simulated vs real world. *Sustain. Cities Soc.* 27, 338–345. doi: 10.1016/j.scs.2016.06.022

Magnenat-Thalmann, N., Kim, H., Egges, A., and Garchery, S. (2005). "Believability and interaction in virtual worlds," in *11th International Multimedia Modelling Conference* (Melbourne: IEEE), 2–9. doi: 10.1109/MMMC.2005.24

McMahan, A. (2003). "Immersion, engagement and presence," in *The Video Game Theory Reader*, eds M. J. P. Wolf, and B. Perron (New York, NY, US: Routledge), 67–86.

Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. (1995). "Augmented reality: a class of displays on the reality-virtuality continuum," in *Telemanipulator and Telepresence Technologies, Vol. 2351* (Bellingham, WA: SPIE), 282–292. doi: 10.1117/12.197321

Möller, S. (2023). *Quality Engineering: Quality of Communication Technology Systems*. New York, NY: Springer Nature. doi: 10.1007/978-3-662-65615-0

Möller, S., and Raake, A. (2014). *Quality of Experience: Advanced Concepts, Applications and Methods*. New York, NY: Springer. doi: 10.1007/978-3-319-02681-7

Möller, S., Schmidt, S., and Beyer, J. (2013). "Gaming taxonomy: an overview of concepts and evaluation methods for computer gaming qoe," in *2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)* (Klagenfurt am Wörthersee: IEEE), 236–241. doi: 10.1109/QoMEX.2013.6603243

Murray, C. D., Fox, J., and Pettifer, S. (2007). Absorption, dissociation, locus of control and presence in virtual reality. *Comput. Human Behav*. 23, 1347–1354. doi: 10.1016/j.chb.2004.12.010

Nilsson, N. C., Nordahl, R., and Serafin, S. (2016). Immersion revisited: a review of existing definitions of immersion and their relation to different theories of presence. *Hum. Technol*. 12, 108–134. doi: 10.17011/ht/urn.201611174652

Otto, M., Lampen, E., Agethen, P., Langohr, M., Zachmann, G., Rukzio, E., et al. (2019). A virtual reality assembly assessment benchmark for measuring VR performance & limitations. *Procedia CIRP* 81, 785–790. doi: 10.1016/j.procir.2019.03.195

Papagiannakis, G., Kim, H., and Magnenat-Thalmann, N. (2005). "Believability and presence in mobile mixed reality environments," in *IEEE VR2005 Workshop on Virtuality Structures* (Bonn).

Regia-Corte, T., Marchal, M., Cirio, G., and Lécuyer, A. (2013). Perceiving affordances in virtual reality: influence of person and environmental properties in perception of standing on virtual grounds. *Virtual Real*. 17, 17–28. doi: 10.1007/s10055-012-0216-3

Rollings, A., and Adams, E. (2003). *Andrew Rollings and Ernest Adams on Game Design*. Indianapolis, IN: New Riders.

Rovira, A., Swapp, D., Spanlang, B., and Slater, M. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Front. Behav. Neurosci*. 3:917. doi: 10.3389/neuro.08.059.2009

Ryan, M.-L. (2009). From narrative games to playable stories: toward a poetics of interactive narrative. *StoryWorlds* 1, 43–59. doi: 10.1353/stw.0.0003

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: factor analytic insights. *Presence* 10, 266–281. doi: 10.1162/105474601300343603

Schuemie, M. J., Van Der Straaten, P., Krijn, M., and Van Der Mast, C. A. (2001). Research on presence in virtual reality: a survey. *CyberPsychol. Behav*. 4, 183–201. doi: 10.1089/109493101300117884

Sheridan, T. B. (2000). "Interaction, imagination and immersion some research needs," in *Proceedings of the ACM symposium on Virtual Reality Software and Technology* (New York, NY: ACM), 1–7. doi: 10.1145/502390.502392

Shoemaker, S. (1990). Qualities and qualia: what's in the mind? *Philos. Phenomenol. Res*. 50, 109–131. doi: 10.2307/2108035

Skarbez, R. (2016). *Plausibility Illusion in Virtual Environments*. Chapel Hill, NC: The University of North Carolina at Chapel Hill.

Skarbez, R., Brooks, F. P., and Whitton, M. C. (2020). Immersion and coherence: research agenda and early results. *IEEE Trans. Vis. Comput. Graph*. 27, 3839–3850. doi: 10.1109/TVCG.2020.2983701

Skarbez, R. Brooks, Jr, F. P., Whitton, M. C. (2017). A survey of presence and related concepts. *ACM Comput. Surveys* 50, 1–39. doi: 10.1145/3134301

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

Slater, M. (2017). "Implicit learning through embodiment in immersive virtual reality," in *Virtual, Augmented, and Mixed Realities in Education*, eds D. Liu, C. Dede, R. Huang, and J. Richards (New York, NY: Springer), 19–33. doi: 10.1007/978-981-10-5490-7_2

Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *Br. J. Psychol*. 109, 431–433. doi: 10.1111/bjop.12305

Slater, M., Banakou, D., Beacco, A., Gallego, J., Macia-Varela, F., Oliva, R., et al. (2022). A separate reality: an update on place illusion and plausibility in virtual reality. *Front. Virtual Real*. 3:914392. doi: 10.3389/frvir.2022.914392

Steuer, J. (1992). Defining virtual reality: dimensions determining telepresence. *J. Commun*. 42, 73–93. doi: 10.1111/j.1460-2466.1992.tb00812.x

Tran, T., Foucault, G., and Pinquié, R. (2021). "Benchmarking of 3D modelling in virtual reality," in *CAD'21* (Riverview, FL: CAD Solutions LLC), 324–328. doi: 10.14733/cadconfP.2021.324-328

Triantafyllou, K., Lazaridis, L. D., and Dimitriadis, G. D. (2014). Virtual reality simulators for gastrointestinal endoscopy training. *World J. Gastrointest. Endosc*. 6:6. doi: 10.4253/wjge.v6.i1.6

Uhrig, S. (2021). *Human Information Processing in Speech Quality Assessment*. New York, NY: Springer. doi: 10.1007/978-3-030-71389-8

Weber, S., Weibel, D., and Mast, F. W. (2021). How to get there when you are there already? Defining presence in virtual reality and the importance of perceived realism. *Front. Psychol*. 12:628298. doi: 10.3389/fpsyg.2021.628298

Webster, M. (2014). *Merriam Webster Online*. Springfield, MA: Merriam-Webster Incorporated.

Weibel, D., Wissmath, B., and Mast, F. W. (2010). Immersion in mediated environments: the role of personality traits. *Cyberpsychol. Behav. Soc. Netw*. 13, 251–256. doi: 10.1089/cyber.2009.0171

Witmer, B. G., and Singer, M. J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence* 7, 225–240. doi: 10.1162/105474698565686

Towards a Quality Framework for Immersive Media Experiences: A Holistic Approach

Asim Hameed, Shafaq Irshad( ), and Andrew Perkis

Department of Electronic Systems,
Faculty of Information Technology and Electrical Engineering,
Norwegian University of Science and Technology, Trondheim, Norway
{asim.hameed,shafaq.irshad,andrew.perkis}@ntnu.no

Abstract. Immersive Media Technologies have emerged as popular media form. Their captivating nature makes them a powerful tool for participation and storytelling in a variety of domains attracting multidisciplinary interest. Existing frameworks for user-perceived quality in immersive media experiences are limited due to their exclusion of narrative dimensions. This research expands upon the current system-centered Quality of Experience framework by including Content Influence Factors based on learnings from IDN. Hence proposing a conceptual framework for measuring immersive media experiences, which comprise of four constructs: Form, Content, User, and Context. These components are interrelated through their overlapping dimensions, which is discussed through the course of this paper.

Keywords: Interactive Digital Narrative · Immersive Media Experiences · Quality of Experience · Virtual reality

1 Introduction

Over the years, immersive technologies have become inherently interactive and their dependence on narrative has gradually increased [7]. When the end user experiences these technologies it results in Immersive Media Experiences (IME). Underlying concepts and dimensions of IME have been developed from a technological perspective [10,12,21] however, quality measures are still rudimentary. Current Quality of Experience (QoE) frameworks limit their definition of content to its type (depth, texture, etc.) and reliability. Thereby, excluding the information and experiences it delivers. In turn, also excluding any narrative-based and/or task-based influences of the content on user-perceived quality. Hence, we believe that assessing quality in Immersive Media Experiences can benefit from the rich scholarship of Interactive Digital Narratives (IDN). It is not clear which factors of an IME are responsible for a user's emotion, involvement, and degree of interest for user-perceived quality. However, immersive media is widely

R. E. Cardona-Rivera et al. (Eds.): ICIDS 2019, LNCS 11869, pp. 389–394, 2019.
https://doi.org/10.1007/978-3-030-33894-7_41

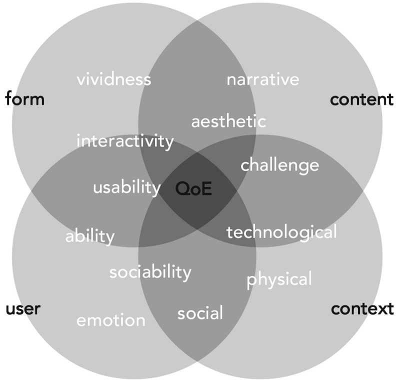


Fig. 1. Quality framework for Immersive Media Experiences (IME)

understood from an experiential perspective as a user's "sense of presence". This framework encapsulates physical, symbolic and psychological dimensions that must be considered for user perceived quality inside IMEs. Given the richness and complexity of emerging media environments, it is important to understand the dynamism of these contemporary media forms before developing quality frameworks. QoE measures are subject to a range of complex and strongly inter-related factors that fall into three categories of **Human**, **System** and **Context Influence Factors (IFs)** [13,18]. Despite their interest around user experience, existing frameworks are predominantly system-centric. With our work we want to focus on a human-centric paradigm by taking into account all those factors that reflect on the user's experience. For this, we accept the important of the above mentioned influence factors for our framework but also include *Content Influence Factors* for their role in overall user satisfaction, and QoE.

2 QoE Framework for Immersive Media Experiences

This research understands IME as a union of immersive, interactive and narrative. This section discusses our quality framework (Fig. 1) in terms of its four constructs: *Form*, *Content*, *User*, and *Context*, considering different dimensions and variables.

2.1 Form

We consider form to be the foundation upon which the entirety of IME is built. It comprises of a system-generated world that affords interaction to its users. Appropriating from Steuer, we denote form by its *vividness* and *interactivity.* One is the system's ability *"to produce a sensory rich mediated environment"*, and the latter is degree to which users can *"influence the form or content of the mediated environment"* [26]. To achieve flow inside any system the experience dimensions and quality dimensions needs to be measured.

1. Experience Dimension (Spatial Presence) is a sense of physical presence, specifically *Spatial Presence*, in the "immersive virtual environment" [26] referred as *Place Illusion* [24]. Ryan [20] refers to it as a new dimension of *Spatial Immersion* that comes from technology not narrative. System immersion is level of immersion (high or low) directly experienced by the user [16,25].

2. Quality Dimension (Vividness) is the sensorial encapsulation of the user is ensured by a distinct quality of technology, *vividness* [25,26]. It is the *"representational richness of a mediated environment ... that is, the way in which an environment presents information to the senses"* [26]. In this research, we consider vividness (extent and fidelity of sensory information) as a user-perceived quality of IVEs that depends on quantifiable system factors of tracking, latency, display persistence, resolution, optics (fov), and spatial audio.

Interaction is derived when a user responds to the *affordances* (action possibilities presented by digital elements, artifacts, and objects) inside a simulated environment [5]. It is a stimulus-driven variable that depends upon the technological formation of the IVE and is quantified under three factors: *speed of interaction* (system response time to user action), *range of interactivity* and *mapping* (system ability to map user input to changes in IVE). The degree to which the interactivity of an IVE, its controller, and feedback mechanisms match the real world has an affect on user's ability in applying natural navigation and manipulation techniques in IVEs.

2.2 Content

We introduce content as a new influence factor in our quality framework for IME. A user removed from their immediate context is immersed into *a reality represented by the medium*, i.e. the broad category objects, actors and events. We argue that an IVE with its inherent interactive qualities is a *live box of action possibilities* produced by the system. Content, on the other hand, is its *"meaning"*. It is the flow of events, inclusion of social elements, nature of task/activities performed. The overall meaningfulness of the content determines various kinds of presence [9,14,22]. Meaning, for the user, is derived from a combination of the content and the context within which the content exists [6]. We divide content into diegetic, non-diegetic, and aesthetic classes of information or experience. For our holistic framework, we have discussed the dimensions of two content factors in specific, i.e. narrative-based and task-based.

Narrative-Based: What storytellers achieved through expression, improvisation, theatrics, and exaggeration are now readily available to users as immersive environments produced by computers. Ryan [20] calls it Spatial Immersion (in her triad of spatial, temporal and emotional immersion). IVE is only a *presentation context* whereas its *narrative context* is the diegetic space of the story that takes place within it [2]. These dimensions are symmetrical to the four narrative-centric factors hypothesized by Rowe et al. [19]. These are *narrative consistency* (believability), *plot coherence* (logical order), *drama* (setup-conflict-resolution), and *predictability* (real-world authenticity). The result of which is a *Plausibility Illusion* - an acknowledgement of the truth of the environment [24].

Task-Based: Flow arises when perceived challenges correspond to perceived skills via *experience of flow* [4]. On the contrary, a mismatch between ability and challenge can lead to feelings of frustration and displeasure. A task inside a VE is determined by its nature and level of challenge (cognitive/motor). Additionally, tasks are also affected by context (e.g. temporal) and depend on the kind of interaction they require, i.e. navigation, selection or manipulation. Task performance improves when a user's ability is matched by the usability of a system. Another important factor is the introduction of aesthetic features (e.g. interface graphics, gamification, etc.) to enhance user performance. It can be hypothesized that tasks performed in IVEs influence the emotional state of the users and is directly influenced by the user's ability to use system [1,23,28].

2.3 User

User, or human, influence factors are deemed influential for the formation of quality [3]. User characteristics, their learning ability and assumed agency play a significant role in shaping the overall perceived quality of IME. *Characteristics* are demographic attributes as well as perceptual, cognitive and motor abilities of users [11]. Prior experiences of IVEs affect a willful suspension of disbelief as well as allocation of attentional resources [11] in turn, affecting presence. Other works [8,15,29] have identified the effects of age, gender, cultural background, and emotional state on user-perceived quality. Due to their characteristic similarity to the real-world, users have a higher chance of learning IVEs [17,27].

2.4 Context

Context factors are relevant situational properties that can be broken down into physical, temporal, social, economic, task and technical characteristics [18]. They have considerable effect on the quality levels of any media experience. But since fully immersive media (such as VR) occlude the real-world, we arrive at an inside and an outside. Simulated contextual changes inside virtual environments can affect user characteristics. IMEs are powerful because of the agency they give the end user. They are not mere simulations but entirely new spaces of signification as well. User do not just experience high-fidelity geometries with real-time responsiveness but the meanings those interactions deliver. This is why they require new inclusive measures for quality assessment. Hence, evaluating all the dimensions discussed above can depict the overall QoE of IMEs.

3 Conclusion

This research paper presents a modified quality framework of IMEs. In addition to immersivity and interactivity, the framework draws from theories and approaches in IDN to include narrativity as an important facet. The paper presents a four constructs i.e. Form, Content, Context and User, that determine quality in IMEs. For its practical use, the framework emphasizes on the importance of signification (the meaning delivered) aspects of these experiences for the user. We believe that any user-perceived experience evaluation is incomplete without considering narrative-related and task-related dimensions inside content.

References

1. Allcoat, D., von Mühlenen, A.: Learning in virtual reality: effects on performance, emotion and engagement. Res. Learn. Technol. **26**, 2140 (2018)
2. Brooks, K.: There is nothing virtual about immersion: narrative immersion for VR and other interfaces (2003). http://alumni.media.mit.edu/~brooks/storybiz/immersiveNotVirtual.pdf. Accessed May 2007
3. Brunnström, K., et al.: Qualinet white paper on definitions of quality of experience (2013)
4. Csikszentmihalyi, M.: Flow: The Psychology of Happiness. Random House, New York (2013)
5. Gibson, J.J.: The theory of affordances, Hilldale, USA, vol. 1, no. 2 (1977)
6. Gigliotti, C.: Aesthetics of a virtual world. Leonardo **28**(4), 289–295 (1995)
7. Gomes, R.: The design of narrative as an immersive simulation, Vancouver, CA (2005)
8. Guntuku, S.C., Lin, W., Scott, M.J., Ghinea, G.: Modelling the influence of personality and culture on affect and enjoyment in multimedia. In: 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 236–242. IEEE (2015)
9. Heeter, C.: Being there: the subjective experience of presence. Presence: Teleoperators Virtual Environ. **1**(2), 262–271 (1992)
10. Hupont, I., Gracia, J., Sanagustin, L., Gracia, M.A.: How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift. In: 2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX), pp. 1–6. IEEE (2015)
11. IJsselsteijn, W.A., De Ridder, H., Freeman, J., Avons, S.E.: Presence: concept, determinants, and measurement. In: Human vision and electronic imaging V, vol. 3959, pp. 520–529. International Society for Optics and Photonics (2000)
12. Keighrey, C., Flynn, R., Murray, S., Murray, N.: A QoE evaluation of immersive augmented and virtual reality speech & language assessment applications. In: 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6. IEEE (2017)
13. Le Callet, P., Möller, S., Perkis, A., et al.: Qualinet white paper on definitions of quality of experience. In: European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), vol. 3, no. 2012 (2012)
14. Lee, K.M.: Presence, explicated. Commun. Theory **14**(1), 27–50 (2004)
15. Murray, N., Qiao, Y., Lee, B., Muntean, G.M., Karunakar, A.: Age and gender influence on perceived olfactory & visual media synchronization. In: 2013 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6. IEEE (2013)
16. Nilsson, N.C., Nordahl, R., Serafin, S.: Immersion revisited: a review of existing definitions of immersion and their relation to different theories of presence. Hum. Technol. **12**, 108–134 (2016)
17. Pedroli, E., et al.: Characteristics, usability, and users experience of a system combining cognitive and physical therapy in a virtual environment: positive bike. Sensors **18**(7), 2343 (2018)

18. Reiter, U., et al.: Factors influencing quality of experience. In: Möller, S., Raake, A. (eds.) Quality of Experience. TSTS, pp. 55–72. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-02681-7_4
19. Rowe, J.P., McQuiggan, S.W., Lester, J.C.: Narrative presence in intelligent learning environments. In: AAAI Fall Symposium: Intelligent Narrative Technologies, pp. 127–134 (2007)
20. Ryan, M.L.: Narrative as Virtual Reality 2: Revisiting Immersion and Interactivity in Literature and Electronic Media, vol. 2. JHU Press, Baltimore (2015)
21. Schatz, R., Sackl, A., Timmerer, C., Gardlo, B.: Towards subjective quality of experience assessment for omnidirectional video streaming. In: 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6. IEEE (2017)
22. Sheridan, T.B.: Musings on telepresence and virtual presence. Presence: Teleoperators Virtual Environ. **1**(1), 120–126 (1992)
23. Shin, N.: Online learner's 'flow'experience: an empirical study. Br. J. Educ. Technol. **37**(5), 705–720 (2006)
24. Slater, M.: A note on presence terminology. Presence Connect **3**(3), 1–5 (2003)
25. Slater, M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos. Trans. Royal Soc. B: Biol. Sci. **364**(1535), 3549–3557 (2009)
26. Steuer, J.: Defining virtual reality: dimensions determining telepresence. J. Commun. **42**(4), 73–93 (1992)
27. Tyndiuk, F., Lespinet-Najib, V., Thomas, G., Schlick, C.: Impact of tasks and users' characteristics on virtual reality performance. Cyberpsychol. Behav. **10**(3), 444–452 (2007)
28. Tyng, C.M., Amin, H.U., Saad, M.N., Malik, A.S.: The influences of emotion on learning and memory. Front. Psychol. **8**, 1454 (2017)
29. Zhu, Y., Heynderickx, I., Redi, J.A.: Understanding the role of social context and user factors in video quality of experience. Comput. Hum. Behav. **49**, 412–426 (2015)

Spatial Storytelling: Finding Interdisciplinary Immersion

Asim Hameed(✉) and Andrew Perkis

National University of Science & Technology (NTNU),
7491 Trondheim, Norway
asim.hameed@ntnu.no

Abstract. This paper is part of an ongoing transdisciplinary research into immersion. In specific, it focuses on Spatial Storytelling to examine the narrative technique in conjunction with Spatial Presence, a commonly accepted subtype of Presence. How our real-life occupation is a constant narrative making exercise and how storytelling is ingrained in our movement in space. It is argued here that immersion and presence models stand to benefit from spatial theory, particularly, the body of work surrounding spatial practices and narratives. Further, that the incorporation of spatial theory adds to the necessary versatility required in approaching immersion, which has been thus far dominated by positivist empiricism. Contributions of a theorized space are also found missing from interactive storytelling and videogames where subject/object interactivity is seen as mere actions performed inside a given space whereas the paper argues that space is learnt through such involvement.

Keywords: Immersion · Presence · Spatial Storytelling · Spatial practice

1 Overview

When BBC unveiled its coverage of the FIFA World Cup 2018 in Russia, it did so by announcing a dedicated high-tech broadcast trial in VR. A first-time-ever VR experience that was designed to give audiences: "…taste of the future", said BBC [1]. The "fully immersive" experience of the matches transported viewers into a simulated hospitality box at the stadium. One did not only watch a live game but also had access to highlights packages and on-demand content. Additionally, it was possible to scan information on each game, lineups and overall stats of the tournament.

Such experiences are congruent with the state of contemporary society where pervasive media systems have rendered physical space into a data-space. Terms like "fully immersive" and "as-if-real" have become synonymous with the coming of age of audiovisual, multimodal and interactive media capable of occupying our perceptual system and simulating environments that evoke a feeling of 'being there' [2–4], or thereabouts [5]. Referring to the BBC Sports VR app, the user encounters a spatial experience inside an interior space, a hospitality box, which serves a virtual double of a generic hospitality box in some Russian stadium. The richness of experience here is extracted from providing a virtual experience where users could feel *as-if* they truly were in Russia. The experience does not limit itself to a mere delivery of a live

R. Rouse et al. (Eds.): ICIDS 2018, LNCS 11318, pp. 323–332, 2018.
https://doi.org/10.1007/978-3-030-04028-4_35

broadcast. In fact, to enrich this VR experience, the virtual hospitality box lets users interact with other media within, doubling on the illusion. The potential of content selection makes users feel more involved. This positively plays to secure user attention while providing interaction, both considered vital for rich experiences [6].

In this paper we discuss such efforts for richness, realness and/or believability. Is interactivity with virtual objects inside simulated environments enough to instill a sense of immersiveness or presence in the user? Or is it a multi-user shared experience of sociability in these virtual environments that makes them real? Perhaps it's the authorship and agency that comes with content generation and manipulation, which can summon that all evasive feeling of as-if-real? These questions are of interest when creating spectacular synthetic/narrative/virtual environments that would imbue a willful suspension of disbelief or presence. This paper is part of an ongoing body of work aimed at understanding and being able to use digital storytelling to create compelling new immersive media. In doing so, we must depart from the monodisciplinary, and/or multidisciplinary, approaches with an intent to support transdisciplinary endeavors in as far as concepts of immersion and presence are concerned.

The paper considers Spatial Storytelling, as a subgenre of Interactive Digital Storytelling, building on traditions of immersive theatre and invites input from media-psychology and spatial studies, particularly spatial thinking and spatial narratives. It adopts a media philosophical approach to examine, through case studies, the role of participatory spatial narratives to offer a reformulation of the theoretical modelling of topics related to electronic simulations and extended realities.

2 Immersive New Media

2.1 Immersion and the Spatial Presence Models

The consequential challenges posed by such immersive and interactive new media have resulted in an abundance of theory surrounding Immersion and Presence [7–10], while producing notable frameworks [11–13] over the years. However, these frameworks are many, and incoherent, which is effectively due to the interdisciplinarity and multidimensionality of Presence research. Apropos to media technologies, Spatial Presence has emerged as the most relevant subtype of Presence in line with the theories of machine-mediated telepresence and teleoperation [14]. This interest has yielded a more concentrated evaluation of Spatial Presence as a "psychological" [13] "state of consciousness" [12] defined as "the subjective experience of a user or onlooker to be physically located in a mediated space" [6] even though one is not. From a media-psychological standpoint, there are two aspects involved:

1. a simulated spatial environment where one feels located;
2. for that mediated environment to offer perceivable options for activity [15].

In effect, most Spatial Presence models view Immersion as a "sensation of being enveloped" [13] by such media-based environments. Wirth et al. [15], refine this to the "features that give rise to Presence" by stating that, "presence is conceptualized as the experiential counterpart of immersion". Mel Slater's framework for Immersive Virtual

Environments (FIVE) [12], divides the achievement of Spatial Presence into three phases of *place illusion* (I am here), *plausibility* (this is happening), and *body ownership* (it is my body). Each is a separate stage and arguably each requires a varied palette to be effective. However, this interdisciplinary potential is not fully utilized when immersion is limited to a system characteristic alone, i.e., the input properties of the mediated technology to provide stimuli (vividness) and afford action (interactivity) [11, 16, 17]. Immersion as technology or immersion as the experience of being enveloped by technology for place illusion empirically enables researchers to quantify otherwise subjective mediated experiences. As sensorimotor contingencies, i.e. to map and match the user's proprioception; and information it affords the senses (visual, haptic, aural, etc.), it's possible to study immersion as a technically measurable property of the system.

Such frameworks reinforce positivist models that favor data-oriented approaches to perception and representation in these media forms, i.e. to design a simulated spatial environment where one would feel present, and that any such design would be possible through thorough mapping. In this way, as a system property, immersion is thus reducible to a degree of correspondence—higher fidelity of display and tracking yields greater level of immersion—enabling a "productionist metaphysics" [18]. This has led to a vastly Euclidean interpretation of three-dimensional simulated space, which signifies a preoccupation with low-mimetic realism [19] or skeuomorphs; often confused with believability [20].

In contrast, we can also find works that bring interdisciplinarity vis-à-vis immersive and interactive media [21–24], and concentrate on the other two aspects of *plausibility* and *ownership* in the same way. These works are usually at the intersections of hard science and digital humanities, which discuss immersive and interactive new media drawing from fields as diverse as art, narratology, ludology, social anthropology, phenomenology, and psychology to name a few. That for believability, a place illusion is not enough and that the plausibility of reality is enriched through factors like sociability, delight, play, etc. Our interest remains in cultivating immersion on such interdisciplinary lines (See Fig. 1) finding encouragement in projects from within virtual reality and gaming sectors that are turning to low-tech features, such as involved narrative and social participation, to enhance the immersive qualities of their applications and products [25–28].

2.2 Immersion in Interactive Digital Narratives

We find Interactive Digital Storytelling, or Interactive Narrative Design, suitable because it propositions a position at the crossroads of narratology (the study of narratives and socio-cultural narrative structures), ludology (the study of gameplay and design) and HCI (human-computer interaction). Murray's [29] identification of the four essential properties found in computer-based narrative media can be viewed in parallel with Mel Slater's framework. Murray talks of *procedural* (computational), *participatory* (interactive), *spatial* (experiential) and *encyclopedic* (database) properties. Elements are utilized for believability, which must be achieved in congruity to real-life in order for immersive experiences to evoke presence. In other words, it is not only enough to immerse a user into a simulated space but to provide potential for

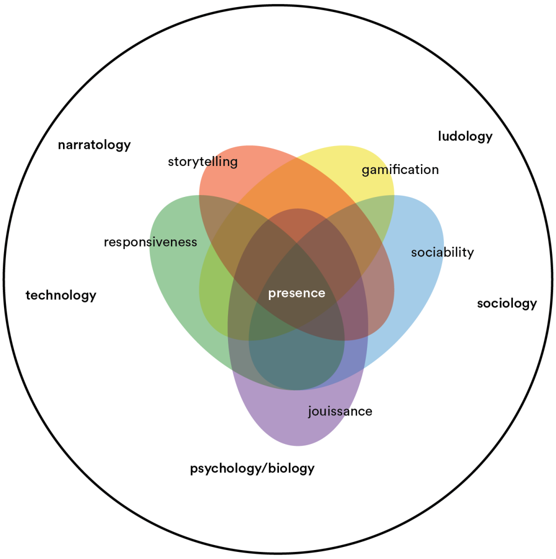


Fig. 1. Immersion radar illustrates the various overlapping influence factors on presence.

believability through additional properties, which too are immersions. Consequently, it can be reasoned that to attain better congruence immersion shall include system immersion [12], but also refer to *absorption & engagement* [30], *strategic and tactical immersion* [30], *imaginative immersion* [31], *challenge-based immersion* [31], *ludic immersion* [32], and *narrative immersion* [32]. Together, they make immersion an interdisciplinary concern.

With Interactive Digital Storytelling, we find a media experience, which utilizes a storytelling engine (system) offering action possibilities (interactivity) to intentionally influence the narrative (immersive) experience. Unlike most Spatial Presence models where an immersive experience is predominantly interpreted inside a simulated spatiality, narrative here, pursues that pivotal role. Narrative, however, is not to a binary categorization of stories non-stories. Instead, it is the potential for ‘storiness’ that is valuable [32]. Ryan’s theorization of narrative as a “semiotic object” is important in this regard since she conceptually develops narrative for use across varied media, i.e. as a cognitive template.

Spatial Storytelling. In continuation, we explicate the aforementioned cognitive template in its application inside Spatial Storytelling. Cognitive templates can be understood as mentally designed codes or stored templates used for the comprehension of our environment. This bottom-up information processing goes by piecing together disparate data to arrive at a bigger and bigger picture. Like this, they aid in the interpretation of experience and shaping an individual’s perception of reality [33]. The term ‘storiness’ can then be the potential for furnished possibilities of a given environment (natural or artificial) for a narrative unfolding.

A good example for ‘storiness’ is Spatial Storytelling, not just because it foregoes linguistic traditions and textual form, but for putting storiness to great effect. Spatial

Storytelling works by spatially engaging a user inside a mediated environment whose discovery through exploration advances a non-linear narrative, and where space is the essential communication medium. Inspired by ‘immersive theater’ [34], it closely follows environmental storytelling in creating preconditions for immersive narrative experiences in four possible ways [35]:

1. ability of spatial stories to evoke pre-existing narrative associations;
2. provide a backdrop where narrative events unfold;
3. embed narrative information within their mise-en-scene,
4. provide resources for emergent narratives.

Immersive theater has been learning from environmental storytelling. A good example is the theater company Punchdrunk [34] who allow their audience the liberty to watch and move as they choose. Further, they involve visual, aural, olfactory and tactile elements to evoke phenomenological multi-sensory experiences. More recently game designers have adopted similar approaches [36, 37]. Through Spatial Storytelling, they attempt to restructure narratives from temporal to spatial bodies of information—narratives distributed across the game space. This appropriation is easier for games since they do not rely on temporal markers common to narratives like “once upon a time…” or “the next day…”, etc. Games are usually characterized by spatiotemporal markers, that is, we point at a certain ‘thereness’ (dungeon, lake, downtown library, etc.) to communicate how far we are in a game; space relays information on time. Hence the readiness witnessed in game design towards Spatial Storytelling. This research considers it to be a compelling model useful for stimulating presence in immersive environments, largely due to its induction of a variety of immersions.

3 Immersive Spatial Narratives

3.1 We Are Immersed in Space

By shifting focus towards space, Spatial Storytelling turns to the narrative potential of locations and places in our everyday life. It is space, marked with disparate anchors of locations and places, each carrying meaning, temporal significance and past memories, which serves as the backdrop against which our individual life stories unfold. Space also works as a force field simultaneously accumulating formal, psychological and ideological histories, discourses, and economies over time—irreducible to and from any one aspect [38].

Everything ‘takes place’ in this space. Therefore, our actions are a “*spatial practice*” [39] “that shapes, and is shaped by, the social, economic, political and cultural” [40] forces within this space. In time, this enriches our *spatial literacy.* Our movement, participation, action, and recreation *inscribe* meaning within this space through the repetitive patterns of a daily routine. These spatial inscriptions emerge over denominated temporal cycles of days, weeks, months and years during our interaction with space, resulting in “*spatial narratives*”. “Through practice, we transform it into a place of meaning and value”: De Certeau [41].

3.2 Time and Space

In spatial theory, space is defined as the "physical setting in which everything occurs". Whereas, place is, "the outcome of the social process of valuing space; a product of the imaginary, of desire, and the primary means by which we articulate with space and transform it into a humanized landscape." [42]. While time and space have been long recognized as the criteria for studying everyday life. Western social theories have been favorably modeled around time, dispassionately assuming compliance from space. This position of dominance is most obvious when one considers the separation of history from geography. To this effect, spatial theory studies offer reflectivity and point at the "[…]implicit subordination of space to time […]" [43]. From a media-philosophical perspective previously absent cross-disciplinary discussions from politics, geography, archaeology, and narratology among other, are fundamental in framing discourses on the co-evolution of space and time with media systems, and for their re-conceptualization in the current age of extended realities.

Returning to our discussion on immersion and presence, one can observe similar binary tendencies in immersive media, particularly in the construction of immersive virtual environments (IVE) and a close "reenactment of Cartesian ontology" [22, 24]. This is evidenced in most Spatial Presence models that treat space as an a priori given; a Cartesian box; a Euclidean XYZ model. These are not self-acquired positions rather cultural values inculcated through traditions of Western technoscience. The conceptualization of space as a container is an attractive proposition for its ease of offering a completeness to its elusive nature [44]. Such an *ontic* position assumes the world (space) to be present-to-hand. In Heideggerian phenomenology we find a challenge to this model contending that space is learned—one learns it—through involvement [45]. Space, it postulates, is an "artifact" [46], which we constantly innovate and mold through our active participation. By being in space we create space, our agency is consumed by the continuous production of space [47]. Space is not a mere container nor an a priori. In fact, it is "an experiential environment whose qualia and character are produced through behavior, ritual, and human activity" [47].

3.3 Body and Space

Activity, our immediate involvement, finds a bodily interpretation in theories of embodied cognition. Space allows for action and movement, which is performed *through the body as a tool*, over a temporal cycle of time—making a narrative. Space is experienced through the body. We can observe this in terms of spatial literacy; if you compare spatial descriptions like north, south, vertical, horizontal, etc. to more experience-based descriptions such as lying down, in front of, straight up, etc. we'd find a better understanding of the latter set. This is because humans, from their childhood, develop through a bodily experience of space, which helps them in learning and understanding space (spatial literacy).

This spatial literacy is made possible by affordances, which are furnished action possibilities in an environment (space). A core concept for embodied psychology models and now widely adopted in interface and interaction (UX) design. It is vital not to confuse or restrict affordances to mere things one does inside an environment (natural or artificial). In fact, they are a relational complementarity between subject/environment, subject/object, object/environment, all at the same time [48]. Affordance are both projectable and non-projectable, for example, a door presents a projectable property of opening but can also have a non-projectable property of one being excited to open the door for your friends; the latter we learn from our experience in space. In his Spatial Presence model, Schubert et al., refer to these as “anticipated” actions that help in presence, they call them a “cognitive feeling”. Such research developments in theorizing Spatial Presence are refreshing for they bring a psychological model closer to its phenomenological counterpart.

3.4 A Way Forward with Spatial Storytelling

The insistence upon a Cartesian way of seeing-the-world (mind over body, the subject over the object) has more in common with renaissance perspectivism than with space. Our example of the BBC VR hospitality box is the most recent illustration of such representationalism. Such inclinations prevail over immersive media industries and, as previously discussed in relation to data-oriented system immersion, remain a popular conceptualization for research models and frameworks.

Alternatively, there are encouraging niche research projects like Holojam [27, 49], developed by the NYU Future Reality Lab, which creates interactive, participatory and shared immersive experiences. It is a nonpartisan approach with low-tech solutions integrating sociability through a collective activity. Holojam employs multiple immersions for effectivity, proving a useful precedent for study. Users are represented as stick figures as opposed to photorealist avatars, walking around a shared space where other users, local and remote, are there to interact with and contribute in making spatial art. Holojam makes use of a *participatory spatial narrative* experience inside an immersive environment where users can talk, observe, and physically interact with one another and the space.

Holojam can be viewed as a Spatial Storytelling model that favors believability to realism. It achieves this through: one, transporting physical objects into the ambient virtual space (place illusions) to create familiarity; two, requiring participation (plausibility); three, a participation not only with the virtual environment but social too, meaning, interaction with other users; four, these multi-user interactions are used for collaborative activity (increase sense of ownership); and five, the activity takes place in a shared virtual space allowing remote users to congregate. Through a shared (social) activity performed in space (spatial practice), participants create unique narratives that they can reflect on.

4 Conclusions

In conclusion, this paper notes that Spatial Storytelling presents promising theoretical interstices, which can help in the development of more cohesive models for immersion and presence. It creates an opportunity for technicians, designers, narrators, and theorists to contribute inside a diverse team. It identifies some immediate research directions for pushing forward interdisciplinary research on immersion, such as: evaluation of immersive experiences using system-based immersion against immersive media that involve multiple immersions; assessing narrative content-generation and manipulation as an influence factor on the quality of immersive experiences; using spatial literacy exercises to improve body ownership of subjects; and assessing ease-of-use in mixed-reality and virtual-reality applications through the benchmark of spatial practice.

There is a burgeoning growth of immersive media products dominated by gaming apps that provide entertainment material for a content-craving consumer market. To make the most of this anticipation, new media applications have to be seen as exciting new paradigms that require to be explored in their own right. Passive content inside simulated spaces shows little consideration for the potential of the medium. Spatial Storytelling offers a chance at agency to the user inside immersive media to focus on believability not realism that is congruent with the narrative of our daily lives.

References

1. BBC Sports page. https://www.bbc.com/sport/football/44305845. Accessed 1 July 2018
2. Pausch, R., Proffitt, D., Williams, G.: Quantifying immersion in virtual reality. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, pp. 13–18. ACM Press/Addison-Wesley Publishing Co., LA (1997)
3. Burdea Grigore, C., Coiffet, P.: Virtual Reality Technology. Wiley-Interscience, London (1994)
4. Spagnolli, A., Gamberini, L.: Immersion/emersion: presence in hybrid environments. In: Proceedings of Presence 2002: 5th Annual International Workshop, Porto, Portugal (2002)
5. Bowman, D.A., McMahan, R.P.: Virtual reality: how much immersion is enough? Computer **40**(7), 36–43 (2007)
6. Hartmann, T., Wirth, W., Vorderer, P., Klimmt, C., Schramm, H., Böcking, S.: Spatial presence theory: state of the art and challenges ahead. In: Lombard, M., Biocca, F., Freeman, J., IJsselsteijn, W., Schaevitz, R. (eds.) Immersed in Media, pp. 115–135. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-10190-3_7
7. Biocca, F.: Lighting a path while immersed in presence: a wayward introduction. In: Lombard, M., Biocca, F., Freeman, J., IJsselsteijn, W., Schaevitz, R. (eds.) Immersed in Media, pp. 1–9. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-10190-3_1
8. Lombard, M., Jones, M.T.: Defining Presence. In: Lombard, M., Biocca, F., Freeman, J., IJsselsteijn, W., Schaevitz, R. (eds.) Immersed in Media, pp. 13–34. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-10190-3_2
9. Slater, M.: A note on presence terminology. Presence Connect **3**(3), 1–5 (2003)
10. Mestre, D., Fuchs, P., Berthoz, A., Vercher, J.L.: Immersion et présence. In: Le traité de la réalité virtuelle, 1st volume, p. 309-38. Ecole des Mines de Paris, Paris (2006)
11. Morie, J.F.: Ontological implications of Being in immersive virtual environments. In: The Engineering Reality of Virtual Reality 2008. SPIE, USA, vol. 6804, article id. 680408 (2008)
12. Slater, M., Wilbur, S.: A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. Presence Teleoperators Virtual Environ. **6**(6), 603–616 (1997)

13. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence **7**(3), 225–240 (1998)
14. Steuer, J.: Defining virtual reality: dimensions determining telepresence. J. Commun. **42**(4), 73–93 (1992)
15. Wirth, W., et al.: A process model of the formation of spatial presence experiences. Media Psychol. **9**(3), 493–525 (2007)
16. Biocca, F.: The cyborg's dilemma: progressive embodiment in virtual environments. J. Comput.-Mediat. Commun. **3**(2) (1997)
17. Zeltzer, D.: Autonomy, interaction, and presence. Presence Teleoperators Virtual Environ. **1**(1), 127–132 (1992)
18. Coyne, R.: Heidegger and virtual reality: the implications of Heidegger's thinking for computer representations. Leonardo **27**(1), 65–73 (1994)
19. Davies, C., Harrison, J.: Osmose: towards broadening the aesthetics of virtual reality. ACM SIGGRAPH Comput. Graph. **30**(4), 25–28 (1996)
20. Vanderbilt, T.: These tricks make virtual reality feel real. Nautilus **32** (2016). http://nautil.us/issue/32/space/these-tricks-make-virtual-reality-feel-real. Accessed 28 July 2018
21. McMahan, A.: Immersion, engagement and presence. In: The Video Game Theory Reader, 1st edn., pp. 67–85. Routledge, NY (2003)
22. Davies, C.: Osmose: notes on being in immersive virtual space. Digit. Creat. **9**(2), 65–74 (1998)
23. McLellan, H.: Virtual realities. In: Handbook of Research for Educational Communications and Technology, 4th edn., pp. 457–487. Springer, London (1996)
24. Morie, J.F.: Performing in (virtual) spaces: embodiment and being in virtual environments. Int. J. Perform. Arts Digit. Media **3**(2–3), 123–138 (2007)
25. Qvortrup, L.: Introduction — welcome into the interface. In: Qvortrup, L. (ed.) Virtual Interaction: Interaction in Virtual Inhabited 3D Worlds, pp. 1–18. Springer, London (2001). https://doi.org/10.1007/978-1-4471-3698-9_1
26. Nilsson, N.C., Nordahl, R., Serafin, S.: Immersion revisited: a review of existing definitions of immersion and their relation to different theories of presence. Hum. Technol. **12**(2), 108–134 (2016)
27. Perlin, K.: Future Reality: how emerging technologies will change language itself. IEEE Comput. Graph. Appl. **36**(3), 84–89 (2016)
28. De la Peña, N., et al.: Immersive journalism: immersive virtual reality for the first-person experience of news. Presence Teleoperators Virtual Environ. **19**(4), 291–301 (2010)
29. Murray, J.H., Murray, J.H.: Hamlet on the Holodeck: The Future of Narrative in Cyberspace. The Free Press, New York (1997)
30. Adams, E.: Fundamentals of Game Design, 3rd edn. New Riders, Indianapolis (2014)
31. Ermi, L., Mäyrä, F.: Fundamental components of the gameplay experience: analysing immersion. In: Worlds in Play: International Perspectives on Digital Games Research, vol. 37(2), pp. 37–53 (2005)
32. Ryan, M.-L.: Narrative as virtual reality. In: Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media, 1st edn. John Hopkins University Press, Baltimore (2001)
33. Solso, R.L., McCarthy, J.E.: Prototype formation of faces: a case of pseudo-memory. Br. J. Psychol. **72**(4), 499–503 (1981)
34. Eglinton, A.: Reflections on a Decade of Punchdrunk Theatre. TheatreForum **37**, 46 (2010). Questia Online
35. Jenkins, H.: Game design as narrative architecture. In: First Person: New Media as Story, Performance and Game, 1st edn., pp. 118–130. MIT Press, Cambridge (2004)

36. Upload VR Gaming page. https://uploadvr.com/exclusive-preview-loading-human. Accessed 30 June 2018
37. IDNA page. https://www.metalocus.es/en/news/idna-%E2%80%93-spatial-storytelling. Accessed 3 July 2018
38. Rakatansky, M.: Spatial narratives. In: Strategies in Architectural Thinking, 1st edn., pp. 198–221. MIT Press, Cambridge (1992)
39. Gilded age plains city page. http://gildedage.unl.edu/narrative/topics.php?q=theory#refn01. Accessed 6 July 2018
40. Lefebvre, H., Nicholson-Smith, D.: The Production of Space. Basil Blackwell, Oxford (1991)
41. de Certeau, M.: The Practice of Everyday Life, 2nd edn. University of California Press, Berkeley (1984)
42. University of Texas Spatial Theory Homepage. https://sites.utexas.edu/religion-theory/bibliographical-resources/spatial-theory/overview. Accessed 7 July 2018
43. Soja, E.W.: Postmodern Geographies: The Reassertion of Space in Critical Social Theory. Verso, London (1989)
44. Neuhaus, F.: Body space and spatial narrative. In: Emergent Spatio-temporal Dimensions of the City: Habitus and Urban Rhythms, pp. 37–54. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-09849-4_3
45. Coyne, R.: Thinking through virtual reality: place, non-place and situated cognition. Techné Res. Philos. Technol. **10**(3), 26–38 (2007)
46. Dunn, S.: Space as an Artefact: a perspective on 'Neogeography' from the digital humanities. In: Digital Research in the Study of Classical Antiquity, 1st edn., pp. 53–72. Ashgate, Burlington VT (2010)
47. Drucker, J.: Humanities approaches to graphical display. Digit. Hum. Q. **5**(1), 1–21 (2011)
48. Gibson, J.: The theory of affordances. In: Perceiving, Acting, and Knowing: Toward an Ecological Psychology, 1st edn., pp. 67–82. University of Michigan, Michigan (1977)
49. Velho, L., Lucio, D., Carvalho, L.: Situated participatory virtual reality. In: Proceedings of XVI Simposio Brasileiro de Jogos e Entretenimento Digital, Curitiba, Brazil (2017)

Appendix B

Publications in Chapter 3

V: Hameed, Asim, and Andrew Perkis. "Affects of Perceived-Actions within Virtual Environments on User Behavior on the Outside." in *12th International Conference on Quality of Multimedia Experience (QoMEX),* pp. 1-6. IEEE, 2020.

VI: Hameed, Asim, and Andrew Perkis. "A Subjective and Behavioral Assessment of Affordances in Virtual Architectural Walkthroughs." in *Applied Sciences 11, no. 17 (2021): 7846.*

V is available at https://10.1109/QoMEX48832.2020.9123099
VI is available at https://doi.org/10.3390/app11177846

2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)

Affects of Perceived-actions within Virtual Environments on User Behavior on the Outside

Asim Hameed & Andrew Perkis
Department of Electronic Systems
Norwegian University of Science & Technology (NTNU)
Trondheim, Norway
asim.hameed@ntnu.no, andrew.perkis@ntnu@no

This paper is not included due to IEEE copyright restrictions available at
https://doi.org/10.1109/QoMEX48832.2020.9123099

applied sciences

MDPI

Article

A Subjective and Behavioral Assessment of Affordances in Virtual Architectural Walkthroughs

Asim Hameed * and Andrew Perkis

Department of Electronic Systems, The Norwegian University of Science & Technology (NTNU), NO-7491 Trondheim, Norway; andrew.perkis@ntnu.no
* Correspondence: asim.hameed@ntnu.no; Tel.: +47-412-856-87

Abstract: Immersive technologies, such as VR, offer first-person experiences using depth perception and spatial awareness that elucidate a sense of space impossible with traditional visualization techniques. This paper looks beyond the visual aspects and towards understanding the experiential aspects of two popular uses of VR in 3D architectural visualization: a "passive walkthrough" and an "interactive walkthrough". We designed a within-subject experiment to measure the user-perceived quality for both experiences. All participants (N = 34) were exposed to both scenarios and afterwards responded to a post-experience questionnaire; meanwhile, their physical activity and simple active behaviors were also recorded. Results indicate that while the fully immersive-interactive experience rendered a heightened sense of presence in users, overt behaviors (movement and gesture) did not change for users. We discuss the potential use of subjective assessments and user behavior analysis to understand user-perceived experiential quality inside virtual environments, which should be useful in building taxonomies and designing affordances that best fit these environments.

Keywords: virtual walkthrough; presence; user-perceived quality; subjective measurements; user behavior

check for updates

Citation: Hameed, A.; Perkis, A. A Subjective and Behavioral Assessment of Affordances in Virtual Architectural Walkthroughs. *Appl. Sci.* **2021**, *11*, 7846. https://doi.org/10.3390/app11177846

Academic Editor: Fotis Liarokapis

Received: 18 July 2021
Accepted: 17 August 2021
Published: 26 August 2021

1. Introduction

Architectural visualizations are uses of media (images, diagrams and more recently 3D modeling techniques) to express and externally reflect upon, design visions. Advances in computer-generated imagery have increased our appetite for life-like photorealistic visualizations of would-be environments and speculations on possible built futures. To this end, the emergence of immersive technologies, especially virtual reality (VR) applications, has presented a powerful first-person communication medium allowing users to step into, freely move about in and explore the environment. Instead of imagining a design, one can have a naturalistic experience akin to the real-world experience of a built environment. VR employs dimensions of immersion, interactivity and presence within computer-generated models to produce an explorable place illusion [1,2]. This feature makes it easier for users to understand spatial relationships, scale and depth. VR-driven architectural visualizations allow projects to be showcased in real-time, enabling immediate and critical feedback. With VR, ideas can be "considered, revised, developed, rejected and returned to" [3]. The synthetic environments of virtual architectural worlds invoke the sense of *being inside* them—a sense of presence [4]. This subjective feeling [5,6] is pivotal for virtual experiences and emerges out of the interplay of immersion and interaction [7–9]. Research indicates that merely a place illusion or spatial presence [8,10–12] alone is not sufficient to sustain prolonged interest in virtual environments (VE). In fact, users also require motivation through involvement and engagement within these worlds for a heightened sense of self-presence [13,14]. This could be a consequence of focusing one's energy and attention on the stimuli available in the virtual world, e.g., interaction possibilities, with which an involved a user experiences more presence [7,15]. Given this premise, we investigated two popular uses of VR in architectural visualization for their effects on users—a "passive walkthrough" and an "interactive walkthrough". We tested subjects within a virtual architectural interior (see Figure 1) with the objective of studying the effects of interactivity on the overall formation of a *sense of presence*, engagement, perceived naturalness and negative effects. This paper describes the experiment and discusses the results.

Figure 1. A user explores the immersive virtual environment using a tethered VR headset inside our laboratory. (Photo: Asim Hameed).

2. Background

2.1. Plausibility in Virtual Environments

Immersive virtual environments (IVE) are 360-degree spatial experiences that either superimpose or occlude the real-space altogether. With this, the ubiquity of real-time rendering has made it possible to experience virtual architectural environments with correct scale and depth precision. IVEs now offer visualization solutions for the design industry, environment models for immersive games, training environments for virtual learning [16], visualization solutions for collaborative design [17] and methods for the gamification of building information modeling (BIM) to test various physical dynamics and performances [18]. All current VR applications facilitate 360° viewing. Some are passive experiences along predefined paths or points with little exploration and interaction. Others allow freedom of movement (exploration) but no interaction, whereas in their most interactive form, they allow for both exploration and interactions with virtual objects.

IVEs are effective spaces because of their similarity to our real-world navigation, mapping and manipulation techniques. As humans we respond naturally and effortlessly to perceived actions. We take this behavior with us into virtual spaces when dealing with the affordances they offer [19]. Like presence, the phenomenon of plausibility illusion (Psi) [20] is also important for research within virtual reality applications. Psi refers to the illusion that a virtual scenario experienced is actually occurring [21]. This refers to the coherence and consistency of behaviors and events that transpire within the context of a given virtual scenario [22]. Psi fits well with the conceptualization of quality as a cognitive judgment. This was investigated by Skarbez et al. [23] in an empirical study where participants transitioned from lower-coherence to higher-coherence scenarios. Participants could change the characteristics and behaviors of their virtual avatars with the goal of matching them to themselves in the real world. The level of plausibility was higher in the highest-coherence scenario; i.e., users connected with the most well-behaved avatar.

For the user, quality is a judgment that distinguishes between perceived quality and expected quality. Refs. [21,24] previously examined the effects of plausibility mismatches on the formation of an overall sense of presence. Both studies underlined the need for developing protocols to assess coherence factors and their consistencies. This study observed selected affordances within an immersive virtual environment (IVE) and examined their effects in terms of user behavior and perceived experiential quality.

2.2. Affordances and Perceived Quality in Virtual Environments

We do not objectively perceive environmental properties or objects; rather, we perceive what we can do [25]. The perception-in-action [26] process is facilitated by the opportunities presented to an organism by its environment, or the *situated affordances* of the environment. The concept has been around for a long time, and a detailed explication is not within the mandate of this paper. However, it is important to emphasize that affordances are neither objective nor subjective; instead, they "cut across the dichotomy of subjective–objective... Both physical and psychical, yet neither" [27]. Affordances must therefore be understood in two ways: (a) affordances are properties of the environment; or (b) affordances are relations between an organism and its environment. Building on Hassenzahl's hedonic and pragmatic model [28,29], we define four distinct affordance types that expand from more immediate operational goals to deeper biological or psychological needs. These are:

1. Manipulation Affordances: The directly perceived affordances that speak to the physical/sensorial compatibility between the user and the object.
2. Effect Affordance: It describes the functioning of the object due to manipulation. It is also directly perceived based on cause-and-effect knowledge of the user.
3. Use Affordance: It relates to the physical and mental skills of the user utilizing the right cognitive or usage plans.
4. Experience Affordances: They are related to the psychological and biological needs of the user and are perceived only with correct knowledge and usage modes.

Manipulation affordances are at the lowest level and are signified by motor goals performed in order to accomplish do-goals, i.e., the effect and use affordances. At the highest level are be-goals (or experience affordances) that motivate actions towards purposes [29]. Between them, they highlight the how, what and why of interaction possibilities. For example, a VE can afford manipulation to a user in the form of pressing a button (motor-goal). The effect affordance of the environment can associate the pressing of a button with the activation of an illumination object, e.g., a light on a wall—cause and effect. This combination and sequence of actions could be intended towards a use, such as illuminating the scene; the effect-and-use affordances use do-goals. The failure or fulfillment of a do-goal results in emotional consequences, such as satisfaction or annoyance. Returning to Hassenzahl [30], the emotional (hedonic) aspects make up the experience affordance for a user—enabling the achievement of be-goals [31]. The pragmatic aspects of a user's experience come from the compatibility of the user's skills with the capabilities of manipulation—the effect and use affordances of the environment [32].

For a light to turn on when a button is pressed meets a users expectation of how life real-world action sequences work. A light illuminating a scene realizes the use for that action. When these actions are matched to the abilities of the user, their successful performance achieves be-goals inside VEs that can cause pleasure. On the contrary, failing to perform them can cause annoyance. The extent of the interactions available inside a VE and their resultant consequences afford experiences to users. Steffen et al. [33] examined how the availability of affordances in VR applications gives them an edge over physical reality in certain use-cases, e.g., simulation-based training. On the other hand, works [34] evaluating the perceptions of real-world affordances—such as texture, gradient, handle-size, hand-size, etc.—have found that they affect user's choices and emotional states within VEs. This paper builds upon the aforementioned works by focusing on the psychological aspects of affordances inside VEs. In particular, we focus on their influences on user-perceived quality and the *sense of presence and plausibility*. It is part of ongoing research into factors influencing user experience and performance inside IVEs [35].

2.3. Measuring User Behavior and Experience

User-perceived quality is the emotional response, involvement and degree of interest a user shows. Inside IVEs, a foremost user experience is that of a "sense of presence," characterized as the "human experience" of the environment [20,36]. Presence is classified into three categories [4,10]: *spatial presence, self-presence* and *co-presence*. Of these, *spatial presence*, is a subjective feeling of "being there" inside a mediated space. For a user, this is characterized by a temporary loss of attention to the physical environment, and a behavioral response to the mediated environment. A user is said to be in a state of immersion when he responds to the physical and symbolic affordances (action possibilities) of the mediated environment [37]; in how he interacts with its "continuous stream of stimuli" [7], appropriates the tools/interface at hand and moderates his actions. Understanding these points help optimize the overall user-perceived quality of immersive applications. Kahneman [38] proposed two systems of thought: System 1 (fast, instinctive, emotional); and System 2 (slow, deliberate, logical). Over the years, numerous surveys and questionnaires [7,39–41] have been used as subjective assessment measures of VEs to capture self-reported System 2 reflective processes—things that do not usually come naturally and require some sort of conscious mental exertion on the part of the user—skills, mental or emotional states, etc. Reflexive System 1 skills are more intuitive and automatic, such as the innate abilities to perceive the world and recognize objects. They are better captured using physiological measures to assess covert and fast behaviors. Considering that most IVEs make use of (or imitate) real-world navigation and manipulation techniques, we have employed observation methodologies in this work. We believe behavioral observation can be useful for assess ing data (overt-motor responses and movements patterns) of subjects collected while they explore IVEs.

3. Materials and Methods

We conducted a repeated measures user study in two visually identical virtual models, manipulating only the affordances of the environment. We had two independent scenarios:

- Passive-walkthrough (PW), an immersive environment with navigation affordances but no interactive features;
- Interactive-walkthrough (IW), an immersive-interactive environment with navigation affordances and a few manipulation and effect affordances.

The aim was to observe and investigate whether the addition of affordances—and interactivity features—within IVEs affected the overt behavior of users; and further, whether their behavioral performances correlated with their subjective evaluations of the IVEs. The use of a within-subject method and visually identical environments was to reduce errors associated with individual differences. For subjective assessments, the study collected profile surveys, and a presence questionnaire was used by users for post-experience self-reporting. The behavioral assessment was based on an active-time diary and ethograms (inventory of behaviours or actions) generated from the video-data for each participant in the study. The behavioral patterns were analyzed against the subjective experiential scores for each participant to verify whether:

> manipulation and effect affordances in the immersive-interactive (IW) scenario would result in higher perceived experiential quality and higher behavioral activity compared to the non-interactive scenario (PW).

3.1. Environment

The virtual environment was designed in-house and modeled in Sketch-Up Pro. Texturing, lighting and interactivity elements were applied in UnReal Engine (UE4). The VE in both scenarios represents an architectural interior of a one-bedroom (32 feet on either side) apartment. The open-plan layout has the kitchen extending into the living room and a balcony. The same balcony can be accessed from the bedroom. There is a separate bathroom and storage space to explore as well. Both scenarios, PW and IW, support natural/free walking. They also support *navigation affordances* using point-and-teleport technique for movement and navigation in the IVE. In addition, IW also used *manipulation affordances* and *effect affordances*. Figure 2 shows an image of the environment.

Figure 2. A view of the virtual living room inside the virtual architectural interior model (Unreal Engine © Epic Games. Photo: screenshot).

1. Passive Walkthrough (PW): This model uses high-poly assets from the UE Marketplace, and high-resolution images from an online repository. The model was prepared using datasmith in Unreal Engine 4. In order to simulate real-world materials, we used PBR-texturing (physically based rendering), realistic lighting and a spatial soundscape to enhance the immersive experience. The environment was optimized for used with HTC Vive Pro. Both handheld controllers can be used to exploit the *navigation affordance* of point-and-teleport. Additionally, hidden affordances in the form of collider components were also applied to surfaces in the model. They were activated to discourage teleportation or natural/free walking through surfaces (such as walls) to avoid unrealistic perforation effects of virtual surfaces.
   All doors in this environment were open by default to allow users free movement through the interior space.
2. Interactive Walkthrough (IW): This model has all the features from the PW. In addition, the IW scenario also uses additional *manipulation and effect affordances*. While one handheld controller was used for point-and-teleport, the second was used for the interactivity features that include:
   - *Two light toggles around the average eye-level*: A familiar design feature, a button (with a light-bulb icon) provided the required cognitive affordance and the opportunity for manipulation using the handheld controller. A laser pointer (similarly to real-life pointers) could be directed at the button to toggle on–off. The explicit manipulation affordance was immediately satisfied with the effect affordance, as the user should notice their actions resulting in additional scene lighting.
   - *Six operable doors around the average waist height*: Interaction with doors was communicated via *metaphorical affordance*, i.e., the imitation of real-life door

handles. The familiar and explicit affordance of hold-and-twist was, however, not present; instead, their manipulation was possible through a hidden affordance activated when a user clicked the handheld controller closer to the handle. The effect and pattern affordances were revealed to the user through successive movements resulting in learning how to open a virtual door.

- *Six cabinets and drawers at various heights*: Different height levels were used to assess the naturalness of the user's behavioral response. The cabinets used the same manipulation and effect affordances as the doors.

All doors in this scenario were closed by default so that users had to open them using the handheld controllers in order to access different spaces.

3.2. Participants

The study inducted 34 participants (18 male, 16 female, μ = 26.7 ± 6.7) over a period of two weeks via mailing lists, flyers and online forms. Participants tried both scenarios in a randomized order. Before this, none of the participants had tested VR in a laboratory scenario. Then participants reported no competence in VR, whereas 15 participants had basic competence and 9 reported intermediate competence in using VR applications. A total of 68 experiences (N = 34 × 2 scenarios/subject) were recorded. Out of 34, two entire sessions (for subject S4 and S10) were excluded on account of incomplete video data. Participants each signed a written consent form and were duly compensated for their participation. All participants were active users of multimedia technologies, and most had prior experience with head-mounted displays. The experiment was pre-approved, and data collection was in line with ethical principles for medical research involving human subjects. Figure 3 shows a participant engaged with the environment.

Figure 3. A participant in the IW scenario bent down to explore interactivity options for a kitchen appliance. (Photo: Asim Hameed).

3.3. Setup

3.3.1. Laboratory and Equipment:

The experiment was conducted in our VR laboratory, which is approximately 16 × 19 feet in size. The laboratory is equipped for subjective and physiological assessments. The VR simulation was run on a desktop PC operating with 64-bit Windows 10 Pro with an Intel Core i7 7700 3.6 GHz processor, 32 GB DDR4 SDRAM (2800 MHz) and a single 3 GB NVIDIA GeForce GTX 1060 graphics card. The participants explored the virtual environment using the HTC Vive Pro HMD supporting 6DOF and motion tracking. It has a total resolution of (1440 × 1600 per eye) at a 90 Hz refresh rate. The headset features a 110-degree FoV and supports 3D spatial audio. The play area was fixed at 10 × 14 feet inside the laboratory. The experience was externally displayed on a 65-inch Samsung Full-HD TV to examine the activity of the participants and look out for unwanted artifacts and/or graphic or interactivity malfunctions.

3.3.2. Procedure

Experimental sessions were pre-scheduled using Google Forms. They were limited to a single participant at a time. Each slot was allocated 60 min that included testing both scenarios and filling out the respective questionnaires. Participants were received by the moderator. They then filled out a 10-item background information survey. Next, participants tried on the headset (HMD) to familiarize themselves with the HTC Vive Pro controllers following a quick tutorial inside the SteamVR Home space. Participants were then provided a set of instructions explaining the experimental procedure. All participants confirmed their willingness by signing a consent form. The experiment was divided into two parts, i.e., PW and IW. The task order was deliberately randomized for each participant to prevent carryover effects. Subjects spent time in each scenario per their liking. Each experience was followed by the ITC-SOPI questionnaire for experiential evaluation. After testing both scenarios, participants were thanked and compensated for their time.

3.4. Instruments

3.4.1. Subjective Measure

The experiment used the Independent Television Company Sense of Presence Inventory (ITC-SOPI) as the prime instrument—a validated cross-media questionnaire for users to report their experiences of a “displayed environment” [40]. The protocol collected background information, such as demographics, digital proficiency and VR competency at the beginning. Afterwards, participants filled out a post-experience questionnaire following each scenario. The responses were recorded on a 1–5 Likert scale for the four aspects of the ITC-SOPI. Participants had the additional option to put down their comments at the end of all ratings. The ITC-SOPI included:

- *Spatial presence (SP)*—a sense of being there and/or encapsulated by a space.
- *Engagement (EN)*—feeling psychologically involved in, feeling moved by and/or enjoying the content.
- *Ecological validity, or naturalness (NV)*—perceiving the mediated environment as lifelike and/or natural.
- *Negative effects (NE)*—an adverse psychological reaction towards the mediated environment.

3.4.2. Behavioral Observations

Active run-time logs were created by the application for each use. Click activities were also logged within the game. Additionally, over 10 h of video data of participant activity was recorded. Video-based observations made it possible for subjects to express themselves unobtrusively, feel at ease and facilitate more natural. Video-based behavior observation enables frame-accurate annotation of behavior. The data were post-processed for analysis and observation coding inside open-source event logging software, BORIS (Behavioral Observation Research Interactive Software) [42]. All behaviors were coded based on manual video analysis by a single person to ensure reliability. Codings were done in an ethogram (details follow in the next subsection). Observations were coded for each subject in each scenario. Two main types were determined:

1. State Events: durational events that have beginnings and ends.

  - *Still*: The subject remains stationary in one position in the physical space. They might sway, bend or rotate while on the same point without an intentional step.
  - *Stride*: The subject moves intentionally in a forward or backward direction from their stationary position. This can be one complete stride or more.
  - *Sit*: The subject assumes a sitting position.

2. Point Events: non-durational events that are generally momentary only.
   - *Click or Point*: The subject clicks the controller in space either close to the body or away from it.
   - *Turn*: The subject rotates 90-degrees to 180-degrees about an axis either in a stationary position or during movement.
   - *Bend*: The subject bends forwards or backwards.
   - *Extend*: The subject extends their limbs or part of their body outward to touch, kick or peek at objects inside the virtual environment.
   - *Shrink*: The subject draws their limbs or part of their body inwards as a gesture of cautiousness or alertness inside the virtual environment.

Since active run-times for participants varied considerably, a uniform 3-min observation time was used. A 3-min interval/slice was randomly selected from the active run-time sequence of each user.

4. Results

An experiment was designed with a single categorical group at two levels: PW and IW. Four dimensions from the ITC-SOPI, namely, SP, EN NV and NE, make up the measured quantitative variables for our study. The manually coded participant behavior types from video analysis were also variables. Thirty-four participants evaluated two virtual scenarios, out of which, two sessions were excluded due to incomplete data. Sixty-four data entries (2 per subject X 32) were received and analyzed for the four dimensions of the ITC-SOPI and overt user behavior.

4.1. From ITC-SOPI

The results of the questionnaire data were compiled in a mean opinion score (MOS)—average judgment for one scenario over all subjects. A multivariate analysis of covariance (MANCOVA) was run after controlling for the covariate of active run-time. This was done considering the possible effects of active run-time (duration of time spent within the VE) on the four dimensions of the ITC-SOPI. Statistical significance was assumed at $p = 0.05$. There was a statistically significant difference between the two categorical scenarios (PW and IW) on the combined dependent variables of spatial presence (SP), engagement (EN), naturalness (NV) and negative effects (NE) after controlling for active run-time: $F(4, 59) = 4.662$, $p = 0.02$, Wilk's $\lambda = 0.76$, $\eta^2 p = 0.24$. We ran separate ANOVAs for each dependent variable SP, EN, NV and NE. We found significant differences between PW & IW for SP & EN, but no statistically significant difference was found for NV or NE:

- Spatial presence (SP) : $F(1, 62) = 43.50$, $p = 0.001$, $\eta^2 p = 0.166$.
- Engagement (EN) : $F(1, 62) = 26.00$, $p = 0.017$, $\eta^2 p = 0.089$.
- Naturalness (NV) : $F(1, 62) = 133.0$, $p = 0.279$, $\eta^2 p = 0.019$.
- Negative Effects (NE) : $F(1, 62) = 75.40$, $p = 0.383$, $\eta^2 p = 0.012$.

Table 1 shows the comparison of the means of the three items under both scenarios. The mean levels in the table indicate higher values for the IW scenario in at least two categories of SP and EN. These differences can be visualized in the min–max plots available for the four variables in Figure 4.

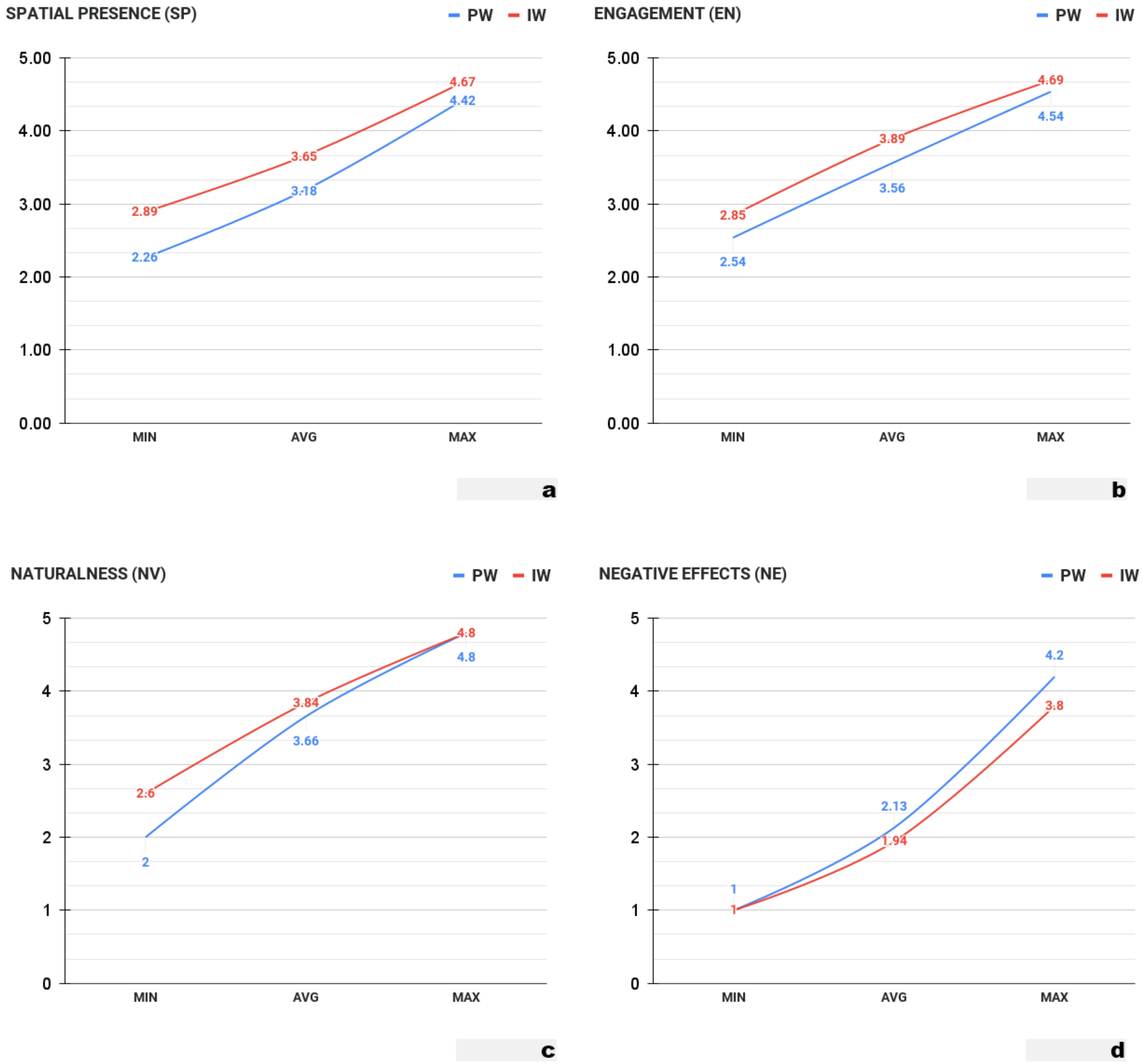


Figure 4. Minimum and maximum for the four ITC-SOPI dimensions under PW and IW: (**a**) Spatial Presence (SP); (**b**) Engagement (EN); (**c**) Naturalness (NV); (**d**) Negative Effects (NE).

Table 1. Collective means for ITC-SOPI items under both scenarios. Means were calculated for every question under each scenario, for all participants.

Item	Condition	Mean (µ)	Std.Dev
Spatial Presence (SP)	PW	3.18	0.61
	IW	3.65	0.50
Engagement (EN)	PW	3.56	0.57
	IW	3.89	0.50
Naturalness (NV)	PW	3.66	0.73
	IW	3.84	0.64
Negative Effects (NE)	PW	2.13	0.91
	IW	1.94	0.79

4.2. From the Time-Log and Observations

Our understanding of behavior begins with the collective means for run-time and activities of all participants (shown in Table 2). The collective run-time for PW was 284 min 25 s (17,065 s). It was 337 min 38 s (20,258 s) for IW. Differences were observed in the run-time, durational and non-durational activities between the two scenarios, PW and IW.

The ANOVA results below indicate no statistically significant results for all behaviors, barring click events, which demonstrated notable differences. This was expected, as subjects had more manipulation opportunities in IW compared to PW. Figure 5 illustrates the min–max plots for participant behavior.

- Run-Time : $F(1, 62) = 3.71$, $p = 0.06$, $\eta^2 p = 0.056$
- Stride State : $F(1, 62) = 1.70$, $p = 0.19$, $\eta^2 p = 0.027$
- Still State : $F(1, 62) = 0.61$, $p = 0.44$, $\eta^2 p = 0.01$
- Turn Event : $F(1, 62) = 2.21$, $p = 0.14$, $\eta^2 p = 0.034$
- Bend Event : $F(1, 62) = 2.0$, $p = 0.16$, $\eta^2 p = 0.031$
- Click Event : $F(1, 62) = 4.77$, $p = 0.033$, $\eta^2 p = 0.071$

Table 2. Collective means for observed behavior types of each participant in both scenarios.

Item	Condition	Mean (µ)	Std.Dev
Run-Time (s)	PW	480.5 s	221.2
	IW	593.47	247.2
Stride (s)	PW	36.21 s	23.1
	IW	44.92 s	29.8
Still (s)	PW	139.48 s	26.9
	IW	33.58 s	32.8
Turn (no.)	PW	17.72	5.2
	IW	15.75	5.4
Bend (no.)	PW	0.66	1.04
	IW	1.16	1.7
Click (no.)	PW	23.28	12.0
	IW	30.47	14.3

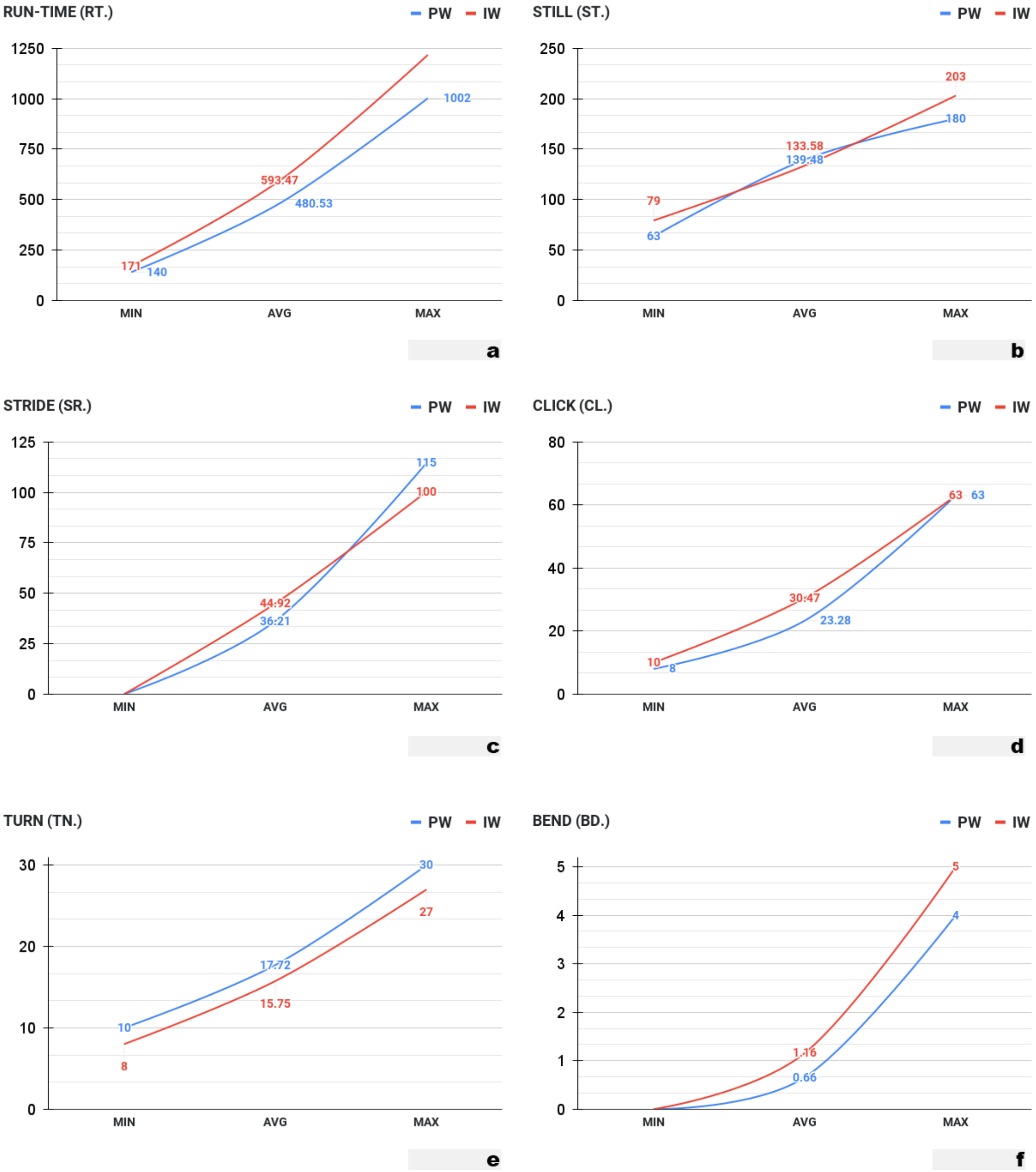


Figure 5. The collective minima and maxima for user behaviors in PW and IW: (**a**) Run-time. (**b**) Still state. (**c**) Stride state. (**d**) Click event. (**e**) Turn event. (**f**) Bend event.

5. Discussion

The above results indicate that while user-perceived experiential quality improved from IW to PW, user behavior showed no significant difference across the sample. Click (or point) events were an exception, showing an increase in IW owing to more action possibilities.

We found that the addition of even a few manipulation and effect affordances markedly increased the place illusion inside the IVE. User-perceived spatial presence increased from PW to IW with a *p-value = 0.001*. Insofar as PW provided multi-directional viewing, it nonetheless remained passive, whereas the interactions in IW made the environment seem more active. Subjects did not feel surrounded by a lifeless world, but one which responded to their actions. This was also expressed in writing by some subjects. The same possibility for action positively affected the level of engagement or involvement (*p-value = 0.017*) that directly correlates with the significance a user attaches to the stimuli or activity of the virtual environment [13,18]. This further alludes to the importance of affordances in creating opportunities for action inside IVEs [22]. Both scenarios had quite high mean scores. However, manipulation and effect affordances did not convince the subjects more of the naturalness, life-likeness or persuasiveness of the virtual environment. Subjects found no difference; *p-value = 0.279*. The same was true for negative effects, as the presence of an adverse psychological reaction did not vary between scenarios: *p-value = 0.383*.

In this section, we expand the collective results with a by-subject comparison for further understanding. We conducted observational assessments of three individual subjects and used event plots for their activity. The three subjects were selected based on the similarity of their logged run-times in both scenarios (in Table 3). The run-time variance between scenarios for other subjects was far greater.

Table 3. Individual activity figures for the three selected subjects (S.) under the two scenarios (Cdn.). Still (St.) and stride (Sr.) events are shown as percentages of total run-time (RT.). Point events of click (Cl.), turn (Tn.) and bend (Bd.) are indicated as frequencies.

S.	Cdn.	RT.	St.	Sr.	Cl.	Tn.	Bd.
S15	PW	4 m 8 s	52	48	20	26	1
	IW	3 m 56 s	76	22	43	26	3
S27	PW	9 m 13 s	68	29	11	14	0
	IW	9 m 52 s	74	26	10	18	0
S05	PW	12 m 45 s	99	0	25	20	2
	IW	16 m 31 s	99	0	51	10	0

5.1. Subject S15

The subject logged the lowest run-time in both scenarios. This indicates that the change in scenario did not effect the time-use tendency of the user. Despite this, the subject reported a higher SP score in IW compared to PW ($\mu = 3.5 > \mu = 2.72$). The same is also true for a higher EN score in IW ($\mu = 3.31 > \mu = 2.62$), and a marginally higher score for NV in IW too ($\mu = 4.2 > \mu = 3.8$). Considering the increases in SP, EN and NV, there is a visible difference in still-to-stride ratio from scenario to PW (52:48) to IW (76:22). In IW, the subject remained stationary for longer to interact with objects. This is visible from the click events that almost doubled from 20 to 43. Figure 6 compares the events plots.

Figure 6. An ethogram showing event log for S15 under both scenarios, PW and IW. The x-axis shows time, whereas the y-axis shows durational and non-durational behavior.

5.2. Subject S27

The subject produced a similar median run-time log for both scenarios. Once again, the change in scenario did not effect the time-use tendency of the user. Compared to PW, the subject reported a minimal score increase for all three dimensions in IW: SP ($\mu = 3.7 > \mu = 3.1$), EN ($\mu = 4.15 > \mu = 3.8$) and NV ($\mu = 4.4 > \mu = 4.0$). The overt behavior for the subject also barely shifted from one scenario to the other. From the plot in Figure 7 we can see the similarities of the events. It can be confirmed with Table 3 that there were next to no behavioral changes by this subject.

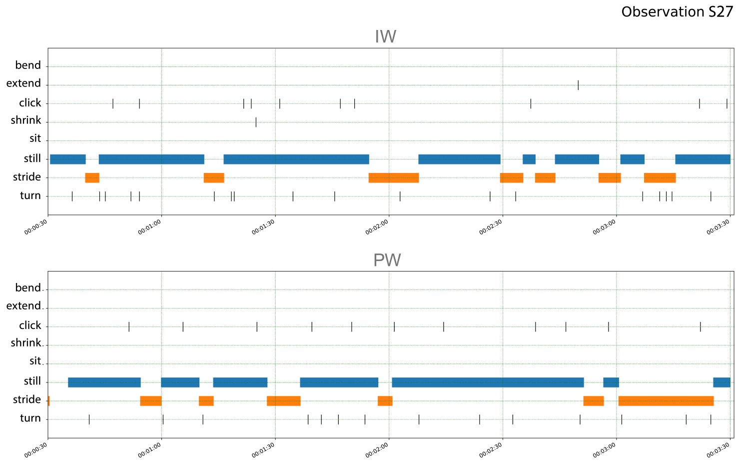


Figure 7. An ethogram showing event log for S27 under both scenarios, PW and IW. The x-axis shows time, whereas the y-axis shows durational and non-durational behavior.

5.3. Subject S05

The subject recorded long run-time logs in regard to the whole sample in both scenarios, consecutively. As evident from the event plot in Figure 8, the subject barely moved from one position in both scenarios. However, it doubled its click events from 25 to 51 in the IW scenario. This, however, did not effect the SP score at all. We see a hair-line increase in IW ($\mu = 3.4 > \mu = 3.5$). Interestingly enough, the score for EN was higher in scenario PW for this subject ($\mu = 4.03 > \mu = 3.85$). The same was true for NV with a higher score in PW ($\mu = 3.6 > \mu = 3.4$). The overt behavior remained similar in both scenarios for this subject as well.

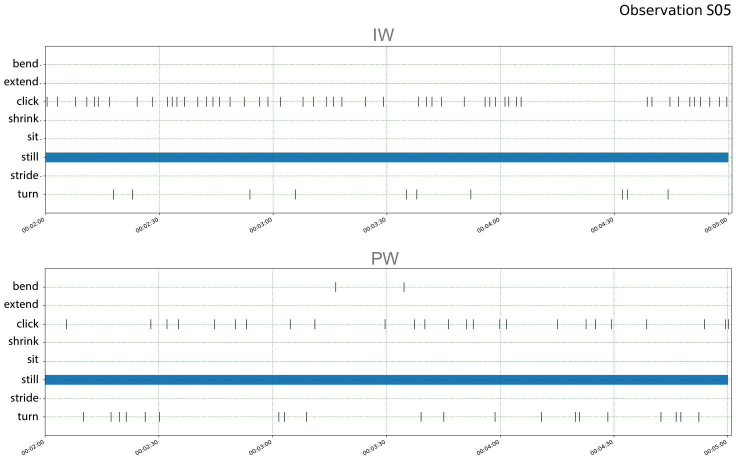


Figure 8. An ethogram showing event log for S05 under both scenarios, PW and IW. The x-axis shows time, whereas the y-axis shows durational and non-durational behavior.

6. Conclusions

The aim of this study was to conduct a comparative assessment of two VR experiences and cross-examine their user-perceived experiential quality against how users behaved in them. We analyzed the effects of manipulation and effect affordances inside a virtual architectural interior on the overall *sense of presence* in users and how they modified their behavior with respect to these affordances.

On the one hand, our study confirms that IVEs are more than just passive geometries and that users feel cognitively and emotionally more involved in virtual environments with action possibilities. The results validated that affordances do positively affect the presence and user-perceived quality. However, results from observation analysis nullified our hypothesis that subjects in the IW scenario would demonstrate higher overt-motor responses to manipulation and effect affordances. Subjects' overt behavior remained predominantly unmoved between the scenarios. The longevity of durational events and frequency of momentary events did not show any significant changes. It is perhaps this lack of overt activity that caused users not to notice any difference in negative effects (arising from exaggerated head-movements, etc.) between the scenarios either. We could observe that:

1. The representationalism (metaphorical affordance) of virtual environments in its imitation of real-life objects creates expectations that can not be physically met, e.g., the door handle.
2. Affordance mismatches resulted in the users appropriating the ready-at-hand tool (i.e., the handheld controller) in a manner most familiar to them.
3. While VR creates an illusion of real-life behavior with objects, users did not use spatial literacy; instead, they felt more comfortable relying on familiar digital literacies (like the pointing and clicking of a mouse).

4. Metaphorical affordances can be useful when the emphasis is on physical exploration, and one-on-one imitation of a function may not be preferred—e.g., when designing immersive-interactive architectural or exhibition tours.
5. Explicit affordances will help when a realistic one-on-one imitation of a function is required in VEs—design prototyping support, test fixture solutions, etc.

It was also observed that most subjects preferred the point-and-teleport technique to natural walking. They avoided extending out in space, even when the situation required so within the IVE. This establishes a premise: investigating whether users' background knowledge of multimedia technologies influences their locomotion preferences. Future studies could include users with lower technological proficiencies to test this. There is definitely a need to further understand the taxonomy of affordances with respect to virtual environments. If most VR experiences are to remain similar to real-life, then the designs of objects and their affordances have to be adjusted to the "human experience" of immersive media. Our future work will focus more on affordance mismatches and their effects on coherence and overall plausibility within VEs. We will further work on subjective and computer-based observation methods to understand mental, behavioral and emotional affordances and their effects on experiential quality inside IVEs.

Author Contributions: The individual contributions of the authors are (according to CRediT taxonomy): Conceptualization, A.H.; methodology, A.H. and A.P.; software, A.H.; validation, A.H.; formal analysis, A.H.; investigation, A.H.; resources, A.H. and A.P.; data curation, A.H.; writing—original draft preparation, A.H.; writing—review and editing, A.H. and A.P.; visualization, A.H.; supervision, A.P.; project administration, A.H.; funding acquisition, A.H. and A.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Norwegian University of Science and Technology (723046).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy protection.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Burdea, G.C.; Coiffet, P. *Virtual Reality Technology*; MIT Press: Cambridge, MA, USA, 2003; pp. 663–664. [CrossRef]
2. Slater, M. Immersion and the illusion of presence in virtual reality. *Br. J. Psychol.* **2018**, *109*, 431–433. [CrossRef] [PubMed]
3. Cross, N. Designerly ways of knowing: Design discipline versus design science. *Des. Issues* **2001**, *17*, 49–55. [CrossRef]
4. Skarbez, R.; Brooks, F.P., Jr.; Whitton, M.C. A survey of presence and related concepts. *ACM Comput. Surv. (CSUR)* **2017**, *50*, 1–39. [CrossRef]
5. Insko, B.E. Measuring Presence: Subjective, Behavioral and Physiological Methods. In *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments*; IOS Press: Amsterdam, The Netherlands, 2003; pp. 109–119.
6. Schuemie, M.J.; Van Der Straaten, P.; Krijn, M.; Van Der Mast, C.A. Research on presence in virtual reality: A survey. *CyberPsychol. Behav.* **2001**, *4*, 183–201. [CrossRef] [PubMed]
7. Witmer, B.G.; Singer, M.J. Measuring presence in virtual environments: A presence questionnaire. *Presence* **1998**, *7*, 225–240. [CrossRef]
8. Nilsson, N.C.; Nordahl, R.; Serafin, S. Immersion Revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Hum. Technol.* **2016**, *12*, 108–134. [CrossRef]
9. McRoberts, J. Are we there yet? Media content and sense of presence in non-fiction virtual reality. *Stud. Doc. Film.* **2018**, *12*, 101–118. [CrossRef]
10. Biocca, F. The evolution of interactive media. *Narrat. Impact. Soc. Cogn. Found.* **2003**, 97–130. [CrossRef]
11. Lee, K.M. Presence, explicated. *Commun. Theory* **2004**, *14*, 27–50. [CrossRef]
12. Hameed, A.; Perkis, A. Spatial storytelling: Finding interdisciplinary immersion. In *Proceedings of the International Conference on Interactive Digital Storytelling*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 323–332.

13. Hameed, A.; Irshad, S.; Perkis, A. Towards a Quality Framework for Immersive Media Experiences: A Holistic Approach. In *Proceedings of the International Conference on Interactive Digital Storytelling*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 389–394.
14. Baños, R.M.; Botella, C.; Alcañiz, M.; Liaño, V.; Guerrero, B.; Rey, B. Immersion and emotion: Their impact on the sense of presence. *Cyberpsychol. Behav.* **2004**, *7*, 734–741. [CrossRef]
15. Schubert, T.W. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Z. Medien.* **2003**, *15*, 69–71. [CrossRef]
16. Dorta, T.; Lesage, A.; Pérez, E.; Bastien, J.C. Signs of collaborative ideation and the hybrid ideation space. In *Design Creativity 2010*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 199–206.
17. Gu, N.; Kim, M.J.; Maher, M.L. Technological advancements in synchronous collaboration: The effect of 3D virtual worlds and tangible user interfaces on architectural design. *Autom. Constr.* **2011**, *20*, 270–278. [CrossRef]
18. Yan, W.; Culp, C.; Graf, R. Integrating BIM and gaming for real-time interactive architectural visualization. *Autom. Constr.* **2011**, *20*, 446–458. [CrossRef]
19. Carassa, A.; Morganti, F.; Tirassa, M. A situated cognition perspective on presence. In Proceedings of the Annual Meeting of the Cognitive Science Society, Stresa, Italy, 21–23 July 2005; Volume 27. Available online: https://escholarship.org/uc/item/7sm3n3bj (accessed on 17 June 2021).
20. Steuer, J. Defining virtual reality: Dimensions determining telepresence. *J. Commun.* **1992**, *42*, 73–93. [CrossRef]
21. Hofer, M.; Hartmann, T.; Eden, A.; Ratan, R.; Hahn, L. The role of plausibility in the experience of spatial presence in virtual environments. *Front. Virtual Real.* **2020**, *1*, 2. [CrossRef]
22. Skarbez, R. Plausibility Illusion in Virtual Environments. Front. UNC-Chapel Hill Libraries Catalog. 2016. Available online: https://catalog.lib.unc.edu/catalog/UNCETDv118rf14z (accessed on 17 June 2021).
23. Skarbez, R.; Neyret, S.; Brooks, F.P.; Slater, M.; Whitton, M.C. A psychophysical experiment regarding components of the plausibility illusion. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 1369–1378. [CrossRef]
24. Bergström, I.; Azevedo, S.; Papiotis, P.; Saldanha, N.; Slater, M. The plausibility of a string quartet performance in virtual reality. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 1352–1359. [CrossRef]
25. Heft, H. Affordances and the body: An intentional analysis of Gibson's ecological approach to visual perception. *J. Theory Soc. Behav.* **1989**, *19*, 1–30. [CrossRef]
26. Goldstein, E.B.; Brockmole, J. Sensation and Perception. Cengage Learning. 2016. Available online: http://www.cengage.com (accessed on 25 August 2021)
27. Gibson, J. The Theory of Affordances. *Perceiving Act. Knowing* **1977**, 67–82. [CrossRef]
28. Hassenzahl, M. User Experience and Experience Design. *Encycl. Hum.-Comput. Interact.* **2013**, *2*. Available online: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/user-experience-and-experience-design (accessed on 18 June 2021).
29. Pucillo, F.; Cascini, G. A framework for user experience, needs and affordances. *Des. Stud.* **2014**, *35*, 160–179. [CrossRef]
30. Hassenzahl, M.; Diefenbach, S.; Göritz, A. Needs, affect, and interactive products–Facets of user experience. *Interact. Comput.* **2010**, *22*, 353–362. [CrossRef]
31. Sheldon, K.M.; Elliot, A.J.; Kim, Y.; Kasser, T. What is satisfying about satisfying events? Testing 10 candidate psychological needs. *J. Personal. Soc. Psychol.* **2001**, *80*, 325. [CrossRef] [PubMed]
32. Hassenzahl, M. Experience design: Technology for all the right reasons. *Synth. Lect. Hum.-Cent. Inform.* **2010**, *3*, 1–95. [CrossRef]
33. Steffen, J.H.; Gaskin, J.E.; Meservy, T.O.; Jenkins, J.L.; Wolman, I. Framework of affordances for virtual reality and augmented reality. *J. Manag. Inf. Syst.* **2019**, *36*, 683–729. [CrossRef]
34. Regia-Corte, T.; Marchal, M.; Cirio, G.; Lécuyer, A. Perceiving affordances in virtual reality: Influence of person and environmental properties in perception of standing on virtual grounds. *Virtual Real.* **2013**, *17*, 17–28. [CrossRef]
35. Hameed, A.; Perkis, A.; Möller, S. Evaluating Hand-tracking Interaction for Performing Motor-tasks in VR Learning Environments. In Proceedings of the 2021 13th International Conference on Quality of Multimedia Experience (QoMEX), Montreal, QC, Canada, 14–17 June 2021; pp. 219–224.
36. Perkis, A.; Timmerer, C.; Baraković, S.; Husić, J.B.; Bech, S.; Bosse, S.; Botev, J.; Brunnström, K.; Cruz, L.; De Moor, K.; et al. QUALINET White Paper on Definitions of Immersive Media Experience (IMEx). *arXiv* **2020**, arXiv:2007.07032.
37. Shin, D.H. The role of affordance in the experience of virtual reality learning: Technological and affective affordances in virtual reality. *Telemat. Inform.* **2017**, *34*, 1826–1836. [CrossRef]
38. Kahneman, D. A perspective on judgment and choice: Mapping bounded rationality. *Am. Psychol.* **2003**, *58*, 697. [CrossRef]
39. Lombard, M.; Ditton, T.B.; Weinstein, L. Measuring presence: The temple presence inventory. In Proceedings of the 12th Annual International Workshop on Presence, Los Angeles, CA, USA, 11–13 November 2009; pp. 1–15
40. Lessiter, J.; Freeman, J.; Keogh, E.; Davidoff, J. A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence Teleoper. Virtual Environ.* **2001**, *10*, 282–297. [CrossRef]

41. Makransky, G.; Lilleholt, L.; Aaby, A. Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Comput. Hum. Behav.* **2017**, *72*, 276–285. [CrossRef]
42. Friard, O.; Gamba, M. BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods Ecol. Evol.* **2016**, *7*, 1325–1330. [CrossRef]

Appendix C

Publications in Chapter 4

VII: Hameed, Asim, Andrew Perkis, and Sebastian Möller. "Evaluating hand-tracking interaction for performing motor-tasks in VR learning environments." in *13th International Conference on Quality of Multimedia Experience (QoMEX),* pp. 219-224. IEEE, 2021.

VIII: Hameed, Asim, Andrew Perkis, and Sebastian Möller. "How Good Are Virtual Hands? Influences of Input Modality on Motor Tasks in Virtual Reality." in *Journal of Environmental Psychology*, 2023: (paper accepted).

VII is available at https://10.1109/QoMEX51781.2021.9465407
VIII is is available at https://doi.org/10.1016/j.jenvp.2023.102137

2021 Thirteenth International Conference on Quality of Multimedia Experience (QoMEX)

Evaluating Hand-tracking Interaction for Performing Motor-tasks in VR Learning Environments

Asim Hameed
Dept. of Electronic Systems
NTNU
Trondheim, Norway
asim.hameed@ntnu.no

Prof. Andrew Perkis
Dept. of Electronic Systems
NTNU
Trondheim, Norway

Prof. Sebastian Möller
Quality and Usability Lab
Technical University Berlin
Berlin, Germany

Journal of Environmental Psychology 92 (2023) 102137

Contents lists available at ScienceDirect

Journal of Environmental Psychology

journal homepage: www.elsevier.com/locate/jep

How good are virtual hands? Influences of input modality on motor tasks in virtual reality

Asim Hameed[a,*], Sebastian Möller[b], Andrew Perkis[a]

[a] *Department of Electronic Systems, NTNU, NO-7491, Trondheim, Norway*
[b] *Quality & Usability Lab, TU Berlin, 10587, Berlin, DE, USA*

ARTICLE INFO

Keywords:
Virtual reality
Input modalities
Gesture interaction
User experience
Performance metrics
Usability
Presence
Perceived realism

ABSTRACT

Hand-tracking enables controller-free interaction with virtual environments, which can make virtual reality (VR) experiences more natural and immersive. As naturalness hinges on both technological and human influence factors, fine-tuning the former while assessing the latter can be used to increase overall experience. This paper investigates a reach-grab-place task inside VR using two input modalities (hand-tracking vs. handheld-controller). Subjects (N = 33) compared the two input methods available on a consumer grade VR headset for their effects on objective user performance and subjective experience of the perceived sense of presence, cognitive workload, and ease-of-use. We found that virtual hands (with hand-tracking) did not influence the subjective feelings of perceived presence, naturalness, & engagement; neither did it inspire the overall ease-of-use while performing the task. In fact, subjects completed the task faster and felt a lower mental workload and higher overall usability with handheld-controllers. The result found that in this particular case, hand-tracking did not improve the psychological and emotional determinants of immersive VR experiences. The study helps expand on our understanding of the two input modalities in terms of their viability for naturalistic experiences in VR akin to real-world scenarios.

1. Introduction

Advances in consumer-grade VR devices are paving the way for natural interactions and direct manipulation of objects inside immersive virtual environments (VE). From a behavioral standpoint, VR is an advanced human-computer interface that allows users to ”immerse” into computer-generated environments and interact in a naturalistic manner within them (Slater, 2009). In general, the user-interaction paradigm for VR has relied on the use of head-mounted displays (HMD) and handheld controllers to freely move, look around, and respond to the various directional, visual, auditory, and haptic stimuli within the 360-degree omnidirectional VEs (Sheridan, 2016; Slater, 2018). More recently, improvements in input modalities on devices like HoloLens, Magic Leap, Oculus, etc. have now made non-mediated realistic interactions a possibility (LaViola Jr et al., 2017) – making VR further attractive for an array of training and learning applications (Liagkou et al., 2019; Thorsteinsson, 2013). VR applications are now useful in safety training in mining (Zhang, 2017), virtual assembly and manufacturing (Abidi et al., 2019; Palmas et al., 2019), medical training (Izard et al., 2018; Pottle, 2019), motor learning and rehabilitation (Crocetta et al., 2018; Mekbib et al., 2020), etc. Advances in the field have made it possible for users to apply their spatial awareness, literacies, and skills while performing in naturalistic real-world paradigms made possible by realistic, dynamic, and multi-sensory VEs (Pfeuffer et al., 2017). In addition to their real-world similitude, another benefit of VEs is that they are fully controllable. They allow perceptual modifications, task scaling, performance measurements, and behavioral observation of participants undertaking activities – well-suited for user experience studies and research.

Interactions within VR are predominantly mediated but natural and intuitive interaction has always been the goal (Regazzoni et al., 2018). Until recently, hand-controllers have been the primary means to interact within VEs but recent advances have made natural interactions a possibility (Kim et al., 2017; Masurovsky et al., 2020). The availability of hand-tracking (use of virtual hands) has been pivotal towards this end. In theory, at least, hand-tracking promises more realistic experiences by facilitating direct control of objects. Though still nascent, the possibility of closer-to-real interactions with hand-tracking will greatly influence the effectiveness and reach of VR. Hand-tracking technologies support gesture-based interactions with virtual objects/artifacts without the

---

* Corresponding author.
E-mail address: asim.hameed@ntnu.no (A. Hameed).

https://doi.org/10.1016/j.jenvp.2023.102137
Received 30 December 2022; Received in revised form 17 June 2023; Accepted 4 September 2023

need for handheld-controllers. With time, interactions inside VEs are expected to emulate the nuances of the real world with increased authenticity.

Given their potential, this work looks at hand-tracking technology. In particular, our study examines how hand-tracking modality compares to handheld-controllers in managing basic reach-and-grasp (or, grab-and-place) actions. For this purpose, we chose the commercially available Oculus Quest VR headset that supports inside-out hand-tracking. We intentionally used this standalone device since it offered affordable VR solutions for everyday entertainment and use. Our user study looks at how well the two input modalities available on the device, in comparison, support basic selection and manipulation operations.

Further, from a usability and user experience point of view, it is not always the case that higher naturalness may also lead to higher performance since the association between realism (i.e., the degree to which device interaction resembles features of real life) and the experience of naturalness is not a linear one (McMahan et al., 2016). Various other works have explored the influences of hand visualization on learning and task execution within VR (Pastel et al., 2022) (Punako Jr and Thropp) (Ricca et al., 2020; Shin et al., 2022). This makes it a point of interest to research the influences, and comparison, of input modalities on task performance and overall user experience inside VEs. Studies have previously used performance and quality metrics for quantitative assessments of VEs (Chiu et al., 2019; Hameed et al., 2021; Suznjevic et al., 2017; Wang et al., 2021). Subjective measures such as psychological aspects (Argelaguet et al., 2016; Lougiakis et al., 2020), cognitive loads (Luro & Sundstedt, 2019; Steed et al., 2016), and usability (Jankowski & Grabowski, 2015; Masurovsky et al., 2020; Voigt-Antons et al., 2020) are also commonly used evaluation methods. These studies indicate that user experience and performance inside VR are eventually influenced by a combination of technological determinants as well as human factors. That the two often overlap and one is not without the other. This paper presents, and discusses, the finding from our study that compared differences in user experience and user performance while performing a reach-grab-place task in VR using two different input modality types.

2. Related Works

2.1. Interactivity inside virtual environments

Interactivity is when users can respond, and make changes, to the contents of a mediated space. This is built upon three basic components: input devices that capture user actions (input modality); display devices that present the effect of these actions back to the user (HMDs); and, transfer functions that map device movements into movements of controlled display or interface elements (mapping method) (Bowman et al., 2001). VR systems translate our head and body positions, map them to the geometries/elements of the VE, and present it as immersive VR inside an HMD.

A similarity between our interactions within VEs to how we interact in the real-world can be marked on a continuum; on the higher end are realistic interactions that are loyal imitations of real-world interactions whereas non-realistic (symbolic/referential) interactions fall on the lower end of the spectrum (Bowman et al., 2012). Both have their uses depending on the nature of the task/challenge at hand. The objective degree of exactness with which real-world interactions can be reproduced inside VEs is defined as the *interaction fidelity* of the system (Ragan et al., 2015). In terms of experience, it is also understood as the degree to which actions involved in the performance of a virtual task correspond to the actions required for an equivalent real-world task (Jerald, 2015). The honesty, or congruence, of interactivity, has noted effects on the degree of perceived realism as well (Bowman et al., 2001). *Perceived realism* refers to how closely a virtual world resembles and feels like the real world (Weber et al., 2021). It is determined by evaluating both the subjective sense of reality (verisimilitude) of the environment and its overall perceived credibility (veridicality). Weber et al. (Weber et al., 2021) note that when using VR, a user will invariably judge the degree of realism of the virtual world in terms of its congruence: (1) its sights and sounds, and the virtual objects/artifacts contained therein, (2) the credibility and plausibility of the plot/scene/situation, and (3) the naturalness and ease of interaction within the VE.

It is important to note that these expectations will inevitably be different depending on the content presented to the user. For example, a fictitious story or experience in VR may warrant a very different set of expectations compared to a VRLE, such as a VR training simulator. The latter requires realistic interactions matched to the real-world due to the dexterity and precision involved. A user may judge if the size and proportion of the virtual body match their real body, and whether it corresponds well to the task at hand. It is intuitive to assume that higher interaction fidelity will naturally improve the training effectiveness of a VRLE (Hamblin, 2005). But interaction fidelity may also vary in requirement with the nature of the task at hand. For example, a higher interaction fidelity technique was found suitable for a Virtual Biopsy Trainer (Ricca et al., 2017), whereas a low-fidelity system was found sufficient for a Laparoscopic Surgical Trainer (Chellali et al., 2016; Kim et al., 2003).

2.2. Natural interaction paradigms for VR

Interactions inside VR comprise *selection* and *manipulation*. An entity can be selected via controller input, gestures, or gaze. It can then be manipulated via resizing, re-orienting, scaling, rotating, or translating the selected object.

VR systems employ direct and indirect methods for interaction. Indirect manipulation involves interacting with virtual objects through a proxy object like a controller and relying on symbolic referents to build an association with the virtual entities (e.g. pressing a button to move a box) (Holderied, 2017). Direct manipulation methods involve using our bodies to directly interact with entities. An entity can be grabbed naturally, using virtual hands, when it is close enough. If it is not in reach for direct grasping with the hand, ray-casting (laser-pointer) or gaze-based approaches can be used to create that association.

Current VR systems are increasingly employing interface paradigms suited for natural user interactions (NUI), such as, to support direct manipulation of objects (via gestures or body movements) without the need for communicating with intermedial devices. Hypothetically, direct interactions are considered to have a better sense of proximity or psychological closeness because of their use of action cues (Hutchins et al., 1985). This can bear a positive effect on a user's perceived realism of the VE. It is not to say however that direct manipulation methods do not have shortcomings. In fact, they can be exceedingly impractical due to either input device constraints like limited tracking range or any limitations of the human operators, for example, anatomical challenges. Generally speaking, either method appropriately matched to the interaction modality and corresponding to the demands of the task will yield better performance. Rieke et al. (Riecke et al., 2018) note that interaction effectiveness improves when the selection and manipulation options of a virtual interface correspond to the speed, range, and tracking capabilities of the system, and are complemented by the comprehension, skill, and learning capacity of the user (Myers et al., 2019). Unrealistically complicated interactions and/or expectation mismatches lead to adaptation problems and negative human factor implications (Nunnally & Bitan, 2006; Våpenstad et al., 2013). If, or not, this also increases user experience is another question and one which is not widely explored. In their experiments, Voigt-Antons et al. (Voigt-Antons et al., 2020) observed that even though participants felt lower control with hand-tracking, they still reported a positive user experience. In the case of Masurovsky et al. (Masurovsky et al., 2020), lower performance metrics were contrasted by high subjective scores reported by users for controller-free interaction. Given the novelty and interest in naturalistic interactions for VR, it is important to study the various influences input modalities may have on user performance and overall experience.

2.3. Performing reach-grab-place tasks in VR

Prehension is the ability to reach for and grasp onto an object. It is a fundamental motor skill that allows us to manipulate objects in our environment. Prehension involves the coordination of various muscles and joints in the hand, arm, and shoulder to position the hand and fingers around an object in a way that allows us to hold onto it firmly (van de Kamp & Zaal, 2007). It is an important skill for carrying out basic actions in daily life – for tasks such as holding utensils, picking up and carrying objects, or manipulating tools and other objects. Several classifications of human grasp types have emerged over the years (Feix et al., 2015; Kamakura et al., 1980; Schlesinger, 1919; Sollerman, 1980). Various taxonomies classify grasp types into broad categories of power, precision, and intermediate type. Further subcategories are based on thumb positions and finger/palm contacts. In short, the various prehension types depend on the specific movement and positioning of the hand and fingers.

In VR, prehension would relate to the ability of users to select and manipulate virtual objects using either:

- hand-held controllers, which mimic the movement and position of the user's hands in the VE. These controllers can include buttons, triggers, and other inputs that allow the user to perform a range of actions, such as grasping, releasing, and manipulating virtual objects; or
- hand gestures, also hand-tracking, which allows the VR system to detect and interpret the movement and position of the user's hands in the real world and map them onto corresponding hand gestures in the VE. This can be done using sensors on the user's hands or using cameras and other sensors to track the movement of the user's hands and fingers.

In theory, both these methods can enhance the sense of immersion in a VE and a user's perceived realism of the virtual world. But obvious caveats remain. For instance, handheld controllers are technically brilliant since they allow seamless real-time tracking of the hand location within virtual space, but there is still an intermediate device in between. Hand-tracking, on the other hand, promises a direct method but one can still not fully act upon virtual objects. One can pick them up by enveloping them, it is not possible to squeeze or lift them because virtual objects don't have weight, volume, or texture.

As such, there is good reason for understanding prehension for VR given its usefulness in a wide range of tasks, activities, and scenarios suited for simulation-based virtual training, assembly, prototyping, etc. The abundance of research around prehension in VR is concentrated in areas of motor therapy (Kaliki et al., 2012; Sveistrup, 2004; Viau et al., 2004) and rehabilitation (Grimm et al., 2016; Levin et al., 2015; Merians et al., 2006). Separately, in a study with 13 participants, Furmanek et al. (Furmanek et al., 2019) compared reach-to-grasp movement patterns inside VE to those performed in the physical environment. The comparison was based on established kinematic variables and carried out in three phases of initiation, shaping, and closure. They found that user performances remain similar in both environments with the exception of differences found in the closure phase, which was prolonged in VE. In another experiment, participants performed a reach-and-grasp task under monocular, motion parallax, and binocular viewing conditions using a telepresence system (Plooy & Wann, 2000). While a prehension-based activity was used, the study however focused on depth and distance judgements only.

2.4. User experience and performance evaluations

The relationship between user performance and user experience in VEs is complementary and multifaceted. Users who are able to effectively interact, perform actions, and achieve their goals are likely to have a more positive experience. Similarly, users who are more immersed, engaged, and satisfied are likely to perform better. The interplay of these aspects bears an influence on the overall quality of the immersive experience (Brunnström et al., 2013; Perkis et al., 2020). From a user perspective, a VR experience amounts to where they are, what they are doing, and how they are doing it. Also, otherwise theorized into concepts of immersion, presence, and immediacy. To have *presence* or a "sense of being there" in a mediated space is the success of a *place illusion* (Biocca, 2002; Slater, 2018). Such an *Immersion* into the medium is a user's response to either system characteristics or the content presented or both simultaneously (Nilsson et al., 2016). Finally, *immediacy*, speaks to the interaction fidelity of an immersive media experience (McMahan et al., 2016). Realistic interactions (closer to natural) have higher immediacy as opposed to unrealistic interactions (Liou et al., 2017).

Broadly speaking, an immersive and authentic VR experience (realism) may draw users in but their interest and motivation depend on the engagement/challenge offered by the content/task (involvement) and how readily and intuitively the system/interface allows them to perform it (usability). Both objective and subjective measures are common practices for assessing the aforementioned aspects. In most cases, ratings are conducted post-experience after participants have removed their head-mounted displays. More recently, researchers (Alexandrovsky et al., 2020; Feick et al., 2020; Graf & Schwind, 2020) have tried to optimize subjective tools for within-experience use as well.

For the purposes of this paper, we specifically look at the following:

- Performance data: provides game-based quantitative data to assess the quality of a VR product from the perspective of the user, for example, analyzing log data to identify patterns and trends in user behavior, and conducting user testing to measure task completion time, error rates, and other performance indicators. This data is useful for measuring a user's performance but it can also be crucial for understanding and optimizing learning and training environments (Loh et al., 2015).
- Psychological and emotional determinants: of the VE are measured with the widely used Igroup Presence Questionnaire (IPQ) (Schubert et al., 2001). It also broadly considers the verisimilitude and veridicality of the virtual world known in terms of its degree of perceived realism (discussed in an earlier sub-section). Buttussi et al. (Buttussi & Chittaro, 2019) used it in a user study to assess the effects of the three locomotion techniques (joystick, teleportation, and leaning) on participants. The measurement tool is useful for determining perceived interaction fidelity and visual render quality experienced by users (Berki, 2020; Blaga et al., 2020; Fromberger et al., 2015). The possibility of action is generally seen to influence VR experiences (Schubert et al., 2001; Slater, 2018).
- Perceived mental workloads: provide a good overview of a user's state of mind during the performance of tasks inside VEs (Feick et al., 2020; Lackey et al., 2016; Zheng et al., 2012). The multi-dimensional NASA Task Load Index (NASA-TLX) (Hart, 2006) has been extensively used for this purpose. High levels of cognitive workload can lead to frustration, reduced performance, and an overall negative experience. Whereas lower levels of workload allow users to effectively engage with the training material and achieve their learning objectives.
- Usability: is primarily the "ease of use" of the VR interface. Usability issues can include difficulty navigating the VR environment, difficulty understanding or interacting with the training material, or confusion with the interface. In other words, how easy or difficult is it to act in the VE. This is especially crucial for successfully using VRLEs. Toolkits like the AttrakDiff (Hassenzahl et al., 2008) have been used to good effect for measuring and identifying usability and desirability issues influencing the effectiveness of VR learning applications (Chen et al., 2016; Sassatelli et al., 2020).

3. Method

We conducted an empirical study to systematically compare two available input modalities on a consumer-grade VR system. It is widely hypothesized that the possibility of closer-to-real interactions using hand-tracking modality potentially improves our overall experience of VR (Haar et al., 2021; Kilteni et al., 2012; Perez-Marcos et al., 2017). So we investigate modality influences on user experience and performance in simulation-based reach-grab-place tasks in VR. The study additionally considers if a visual factor (e.g. changes in the environmental color) will also bear influence on users' perceptions while undertaking the task as previously observed (Billger et al., 2004; Felton, 2021). The collected results were assessed for the users' subjective experience (e.g. sense of presence, perceived cognitive workload, and ease of use) in correlation to their game-based performance metrics.

A repeated-measures 2 × 2 design was used to compare two input modalities across participants performing a reach-grab-place task inside a VE rendered at two levels of representational realism. The modality type (M) and representation level (R) form our two independent variables (IV). Each had two levels: M, hand-tracking (M1) and handheld controller (M2); and R, saturated (R1) and grayscaled (R2). Giving us four variations, see Fig. 1:

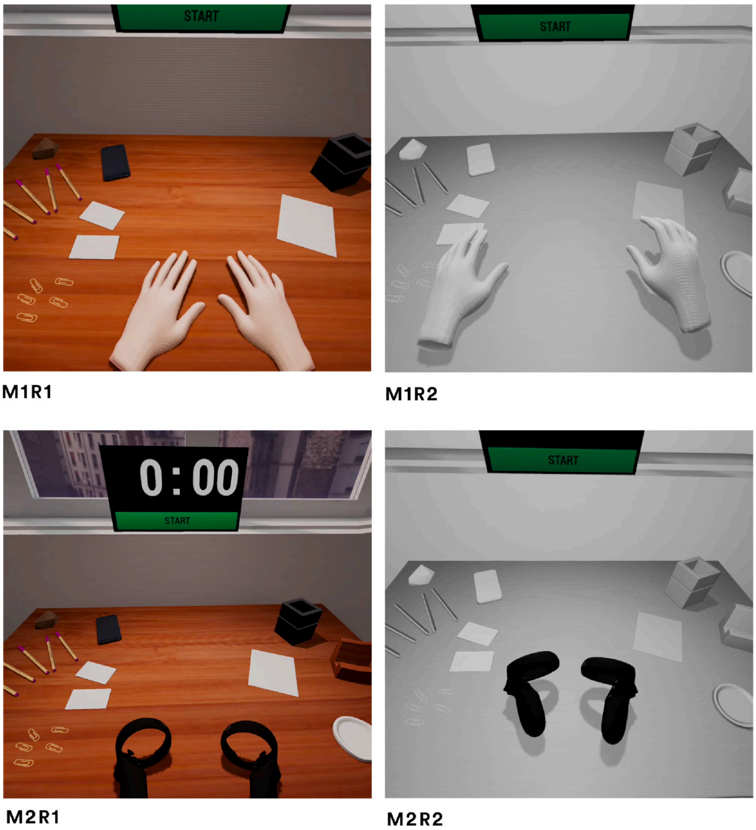
- M1R1 - hand-tracking x colors VE
- M1R2 - hand-tracking x grayscaled VE
- M2R1 - handheld-controller x saturated VE
- M2R2 - handheld-controller x grayscaled VE

Dependent variables (DV) included user performance metrics and scores for user-perceived overall presence, mental workload, and usability from self-reported questionnaires.

3.1. Participants

The sample consisted of a mixed demographic of N = 33 participants (15 male, 18 female, μ = 24.7 ± 2.3). Participants were recruited via an online volunteer portal of the university open to the general public. People from diverse backgrounds signed up for the study. They received gift cards for their participation. Participants reported their demographic data and knowledge/proclivity for immersive technology at the start of the test. Google Forms were used for the pre-study survey. In terms of user experience of VR: the majority of participants (n = 22) reported "no" experience; a few (n = 6) reported "some" prior experience; others (n = 4) had "intermediate/moderate" experience; and a single (n = 1) participant had "good" prior experience of VR. Similarly, only 1 participant had used VR in a lab study before. All participants (n = 32) reported at least "some" prior experience with video game controllers and a single (n = 1) participant had previous experience with hand-tracking. All participants had normal visual acuity or normal corrected visual acuity.

The study was conducted at the VR/XR labs of the university, see Fig. 2. Informed consent was obtained from all subjects involved in the study, and institutional ethics were sought prior to the commencement of this study. All procedures performed in this study were in accordance with the ethical standards of the Institutional Review Board (or Ethics



M1R1 M1R2

M2R1 M2R2

Fig. 1. Four variations based on Modality Types (M1, M2) and Representation Levels (R1, R2). (Unreal Engine © Epic Games. Photo: Screenshot Image).

Fig. 2. A participant using hand-controller modality interacts with the environment displayed inside the HMD.

Committee), and with the 1964 Helsinki Declaration and its later amendments. Experiment protocols and data collection followed the guidelines described in ITU-T Recommendations P.809 (ITU-T, 2018), P.911 (ITU-T, 1998), and P.919 (ITU-T, 2020). Data presented in this study is not publicly available due to privacy protection and is available on request only. This research was funded by the NTNU IE Faculty, Project No. 63350581770958. The study was conducted during the second Covid-19 wave. Strict hygiene protocols were followed during test sessions based on the specified SARS-CoV-2 guidelines for educational institutions.

3.2. Stimuli

The VE comprised a virtual room with a table overlooking a window. Participants sat at the virtual table to carry out a simple reach-grab-place task. The VE was modeled on a 1:1 exactness to the physical laboratory where the experiment was conducted; size, 5.4 m by 4.4 m. The model was prepared in Sketchup Pro (Version 20.2.172). The door and windows inside the VE matched the physical lab. Similarly, the furniture location and orientation were also matched to the physical room. The virtual table overlapped the physical table, and the heights were matched so that participants would feel a surface under their arms while at the virtual table. Textures, lighting, interactivity features, and gamification elements were applied inside Unreal Engine (Version 4.26).

3.3. Task

The VR task required the participants to re-organize a number of assorted items (N = 15) on a virtual table. At the start of each variation, all 15 items would appear on the left-hand side of the virtual table. Participants were then required to reach-grab-place each of the 15 items one by one to their designated positions on the right-hand side. It was important that participants accurately place the objects in their desired spots. Participants could orient the items to their liking and ease. Figs. 3 and 4 show a few select screenshots of the two modalities in action (see captions for details). The task required grabbing and moving the

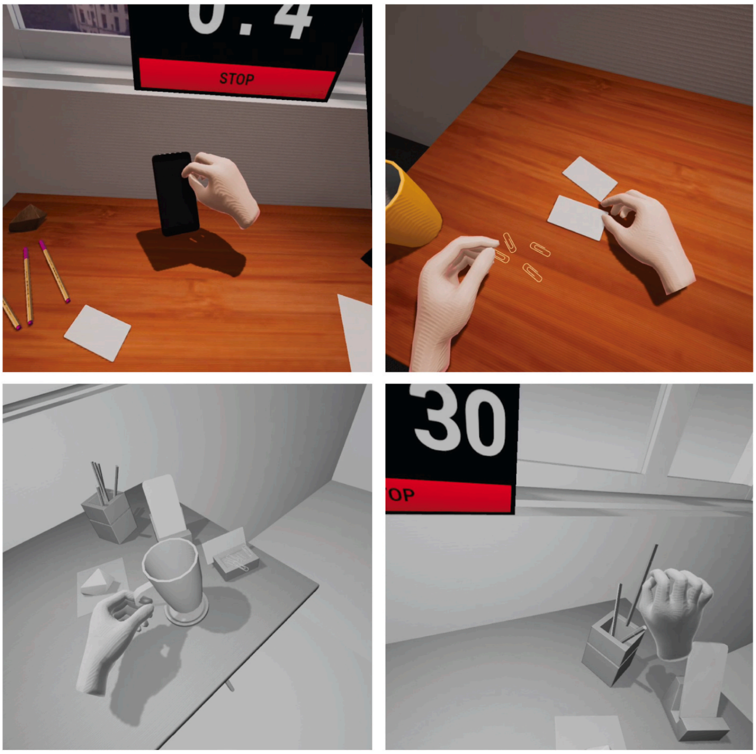


Fig. 3. Top Left: mobile phone grabbed with virtual hands in M1R1. Top Right: user picks a business card with virtual hands in M1R1. Bottom Left: user places the coffee mug with virtual hands in M1R2. Bottom Right: A marker is placed using virtual hands in M1R2. (Unreal Engine © Epic Games. Photo: Screenshot Image).

Fig. 4. Top Left: coffee mug placed on the coaster using controllers in M2R1. Top Right: marker placed in the holder using controllers in M2R1. Bottom Left: user picks a paper clip using controllers in M2R2. Bottom Right: paper weight placed on the paper in M2R2. (Unreal Engine © Epic Games. Photo: Screenshot Image).

following items:

- 5 pens into a pen holder
- 5 paperclips into the clip saucer
- 2 business cards into the cardholder
- 1 mobile phone to the mobile phone holder
- 1 paperweight on top of the paper
- 1 coffee mug onto a coaster

All six objects selected for the task required distinct modes of prehension (see Fig. 5). Each required a separate static grip, described below (Feix et al., 2015; Kapandji, 1987):

1. *Terminal opposition* (Fig. 5A) is a precision grip, which allows one to hold a thin object or to pick up a very fine object like a needle or a paperclip. The thumb and the tip of the index (or the middle finger) come into contact during opposition when fine objects are being grasped.
2. *Tetradigital grip* (Fig. 3B) is for holding larger objects firmly. With a tetradigital grip involving pulp contact, one can hold a pencil, a brush, or a pen. The pulp of the thumb presses the object firmly against the pulps of the index, middle, and ring fingers.
3. *Subterminal opposition* (Fig. 3C) is a grip that involves the sides of the fingers and thumb coming into contact with the object, rather than the fingertips or palms. It is often used for holding onto thin, elongated objects like paper.
4. *Tridigital grip* (Fig. 3D) involves the thumb, index, and middle finger. It is a popular grip used for bringing food to the mouth. Subterminal tridigital prehension in the case of small round or irregular objects.
5. *Panoramic pentadigital grip* (Fig. 3E) involves holding large, flat objects with the fingers widely separated and the thumb positioned in maximal counter opposition. This grip allows one to securely grasp such objects.
6. *Three-finger pinch-dorsal contact grip* (Fig. 3F) is used for holding objects like cups. It involves placing the pad of the index finger at the level of the middle phalanx and the radial aspect of the middle finger on the cup to provide balance and support.

A virtual timer kept time for each task. The timer could be stopped if the user got bored, frustrated, or annoyed with the input modality or the task itself. In such cases, the result would be reported as ”incomplete”.

3.4. Apparatus

The VR game was optimized for use in the Oculus Quest VR headset (Oculus VR, Inc.”, 2020). The system comprises of a standalone device capable of running games and software wirelessly under an Android-based operating system. The headset provided a stereoscopic viewing using OLED display for each eye, with an individual resolution of 1440 × 1600 and a refresh rate of 72 Hz. This headset comes with internal cameras for inside-out, positional tracking of movements that afford six-degrees-of-freedom, 6DOF. Oculus Quest uses both the Oculus Touch (Hand-controller) and also supports controller-free gestures. Fig. 6 shows the use of the two modalities.

Oculus hand-tracking analyzes discrete hand poses and tracks the

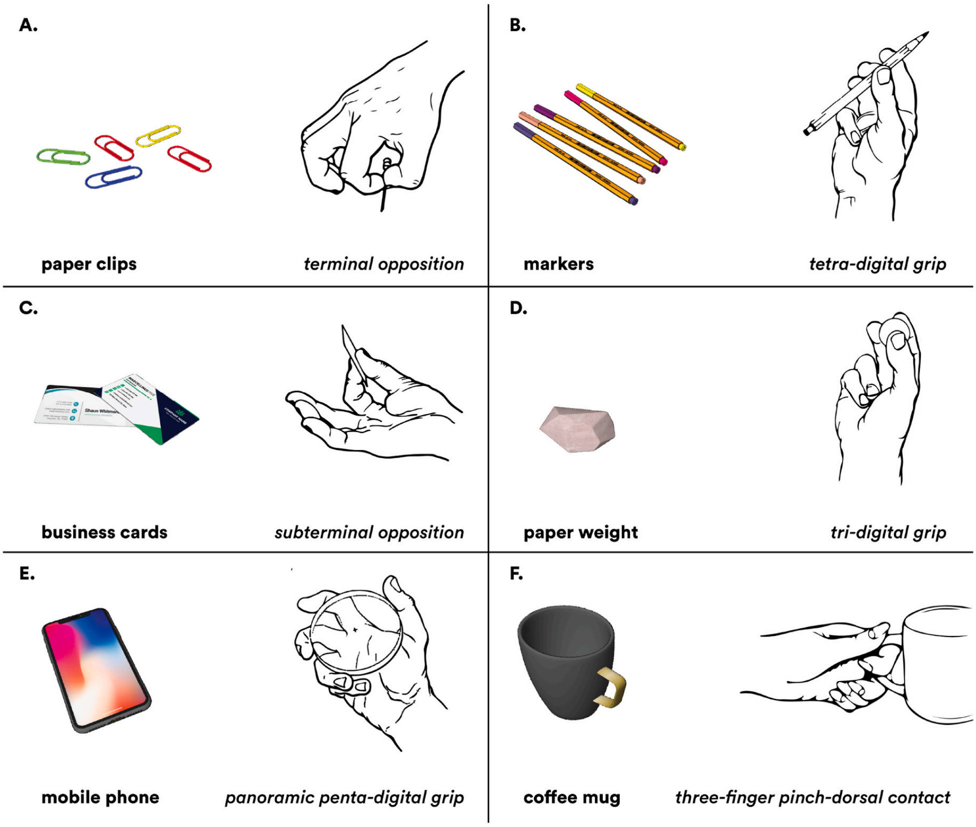


Fig. 5. Six objects used in the reach-grab-place task. Each object required a specific static grip to be handled. (prehension illustrations by I.A. Kapandji (Kapandji, 1987)).

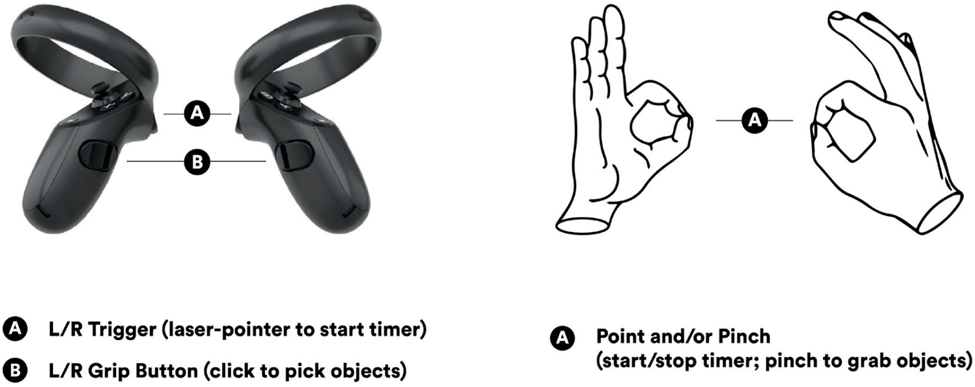


Fig. 6. Two input modalities available on Oculus Quest (Oculus Quest © Oculus VR, Inc.).

position of certain key points on hands in real-time, such as knuckles or fingertips. No official accuracy rate was specified for the Oculus Quest. Separately, AbdlKarim et al. (Abdlkarim et al., 2023) measured the overall performance of the hand-tracking system. At an average, their results showed a fingertip positional error of 1.1 cm, a finger joint angle error of 9.60, and a temporal delay of 38.0 ms.

On the handheld-controller, the left or right grip button could be pressed to lift any of the six objects. One object could be picked at a time by bringing the controller next to it. For hand-tracking, object interactions on the device are limited to basic gestures that include (Oculus, 2020):

- pointing, users extend their index finger and point it forward. The system recognizes it as a gesture for selecting items.
- pinching, when the thumb and index finger are together, the system recognizes it as a gesture used for grabbing the items.

3.5. Procedure

The study had four experimental conditions – M1R1, M1R2, M2R1, and M2R2. Each participant was to perform under all four conditions. The order of the experimental conditions was randomized across the participants to avoid sequence effects. The task requirement remained the same across the four conditions. However, the input modality and degree of representational realism varied in each case (refer to Fig. 2).

Upon arrival at the lab, participants were briefed on the purpose of the study. After signing the consent form, participants filled out a demographic survey. Afterward, participants were seated at a table facing a window. The virtual table within the game was matched at 1:1 to the physical table. All participants first attempted a physical version of the task to familiarize themselves with it so that while performing the task within the virtual game they could focus on *how to do* it rather than *what to do*. Each condition was preceded by a brief instructional phase explaining how to use the input modality. Participants were familiarized with the manipulation techniques through written instructions, which were complemented by a live presentation by the experimenter. The procedure followed throughout the test.

A typical testing session lasted 60 min – divided into four separate condition runs. All four experimental conditions were conducted back-to-back. A 5–7 min resting period between each test condition allowed the participants to recover and avoid eye strain. Each condition proceeded as follows:

- A participant would put on the HMD.
- A task would begin as soon as a participant ”started” the timer.
- The Participant would perform the reach-grab-place task.
- Once finished, the user would ”stop” the timer.
- The Participant would take off the HMD.
- They would rest for a few mins.
- The participant then fills out the post-game questionnaires.

3.6. Measures

Performance Metrics: refer to the patterns and trends identified in user behavior while performing tasks. User performance was judged on the basis of a game log generated at the end of each completed task, which marked:

1. Active-Time (AT), the time taken by a user to complete the task.
2. Object-Pick (OP), number of objects grabbed in the task
3. Click-Frequency (CF), number of attempted clicks to grab the objects.

Igroup Presence Questionnaire (IPQ): The IPQ is a self-report questionnaire to measure the sense of presence in virtual reality environments (Schubert et al., 2001). It contains 14 items rated on a seven-point Likert scale ranging from 0 to 6. The IPQ contains three sub-scales that measure:

1. Spatial Presence (SP), related to the sense of physically being in the VE.
2. Involvement (INV), is meant to evaluate the attention devoted to the VE.
3. Realness (REAL), evaluates the sense of reality attributed to the VE.

These concepts have been covered in the ’Related Works’ section of this paper. Additionally, the IPQ contains one general item which assesses the general “sense of being there” (GP), and has high loadings on all three factors, with an especially strong loading on Spatial Presence. The IPQ has a high reliability (Cronbach’s $\alpha = 0.87$).

Cognitive Load (NASA-TLX): The NASA task load index (NASA TLX) is a tool for measuring and conducting a subjective mental workload (MWL) or cognitive load assessment (Hart, 2006). It allows to determine the perceived cognitive load of users while they are performing a task. The index rates performance across six dimensions to determine an overall workload rating. These are the requirements of:

1. Mental demand (MD), thinking, decisions, or calculations.
2. Physical demand (PD), the amount and intensity of physical activity.
3. Temporal demand (TD), time pressure involved for completion.
4. Effort (ED), how hard is it to maintain performance?
5. Performance (RD), the level of success in completing the task.
6. Frustration level (FD), do you feel secure/insecure or discouraged/content.

Each question has a rating from 1 to 10, where 1 represents the lowest task demand, and 10 represents the highest, with the exception of the performance question, where 1 indicates the highest, and 10 indicates the lowest.

UX Evaluation (AttrakDiff): The Attrakdiff assesses UX-related quality perceptions of the game application (Hassenzahl et al., 2008). The tool applies a hedonic/pragmatic model of UX. This questionnaire evaluates the perceived pragmatic (4 items) and hedonic (4 items) attributes of the interface, as well as its overall attractiveness (2 items), using a 7-point semantic differential scale. The tool is effective for surveys due to its easy and intuitive handling. Furthermore, AttrakDiff offers fast and well-visualized illustrations of the results of comparisons of different products in terms of the user experience. The use of semantic differential makes it possible to narrow down how strongly a user might connect a survey term with a certain property of the application. With the help of opposite adjectives, users can indicate their perception of the application. The word-pairs make a collation of the evaluation dimensions possible; these being:

1. Pragmatic Quality (PQ), describes the usability of a given product and how successfully users achieve their goals using it.
2. Hedonic Quality (HQ), measures emotional reactions of how stimulating/inspiring a product is and whether users identify with it.
3. Attractiveness (ATT), describes the global value of the product based on quality perception, either positive or negative.

3.7. Analysis

The IBM SPSS Statistics (Version 28.0.1.0) software was used for analysis. Two-way MANOVA (Multivariate Analysis of Variance) was used to examine the effects of two independent variables (IV), modality type and representation level, on user performance metrics and the user-perceived sense of presence, mental workload, and usability. Independent observations were collected from a sufficient sample size (N = 33) as per ITU-T recommendations (ITU-T, 1998; ITU-T, 2020). Participants were equally distributed across all conditions (X4). Scores for N = 1 participants were excluded on account of being ”incomplete”. Multivariate normality, outliers, linearity, multicollinearity, and homogeneity of variance were tested before analysis to ensure the assumptions of MANOVA (Pallant, 2020). In the event of significant results, further univariate analyses were conducted. All analyses considered two independent variables (IV x 2) with two categorical groups: M-type (M1, M2), and R-Level (R1, R2). Subjects evaluated all four variations of the application in a randomized order. 128 data entries (4 per subject X 32) were received and analyzed. Below we look at the various dependent variables (DV) that included quantitative game metrics (x 3 DVs) and the different subscales for IPQ (x 4 DVs), NASA-TLX (x 6 DVs), and AttrakDiff (x 3 DVs). The AttrakDiff scores were also additionally analyzed using the official online eSurvey tool for AttrakDiff with outputs of portfolio-presentation, diagram of average values, and description of word-pairs. (User Interface Design - UID GmbH").

4. Results

In this section, we look at the results of our reach-grab-place task in VR. A total of 33 adults were randomly assigned to the four conditions of the 2×2 design. Scores for only 32 subjects were considered for results since $N = 1$ subjects failed to complete the task. We evaluated the main effects and interaction of M (M1, M2) and R (R1, R2) based on the performance of the users and their self-reported responses to the subjective measures. Below, we present our results.

4.1. Performance data

A two-way MANOVA examined the effects of M and R on the measures of user performance: AT, OP, and CF. The dataset for AT and OP had a positively skewed distribution so logarithmic transformation was applied uniformly. Results of Box's M Test (Equality of Covariance Matrices) showed no violation of the assumption ($p = 0.26$). Significance values for Levene's Test (Equality of Error Variances) ($P_{AT} = 0.73$, $OP_{OP} = 0.85$, and $P_{CF} = 0.37$) exceeded 0.05.

The MANOVA revealed a significant main effect of modality type, $F(3, 122) = 195.8$, $p = < 0.001$, Wilk's $\lambda = 0.172$, $\eta^2 p = 0.83$. There was no significant main effect of representation level, $F(3, 122) = 0.85$, $p = 0.47$, Wilk's $\lambda = 0.98$, $\eta^2 p = 0.02$. Nor was there a significant interaction effect between modality and representation, $F(3, 122) = 0.20$, $p = 0.90$, Wilk's $\lambda = 0.99$, 0.005. Follow-up univariate analyses indicated that modality type produced significant differences across all three performance scores,

AT: $F(1, 124) = 62.74$, $p = < 0.001$, $\eta^2 p = 0.34$

OP: $F(1, 124) = 20.1$, $p = < 0.001$, $\eta^2 p = 0.14$

CF: $F(1, 124) = 374.8$, $p = < 0.001$, $\eta^2 p = 0.75$

Descriptive statistics for the dependent variables, presented in Table 1, shows the number of observations(N), the mean(μ), and the standard deviation (SD.) in each group. The table shows both the log-transformed and the original data. As shown, users using hand-tracking took longer to complete the task (μ = 314s, SD. = 169.42; μ = 272s, SD. = 116.58) compared to those using handheld-controllers (μ = 152s, SD. = 67.52; μ = 154s, SD. = 104.12). There was no significant separation for the mean scores on the basis of levels of representation. The mean score for a number of objects grabbed OP, was more in hand-tracking (μ = 32, SD. = 17.04; μ = 29, SD. = 11.82) compared to handheld-controller (μ = 21, SD. = 9.20; μ = 18, SD. = 8.3). Similarly, the number of attempted grabs, CF, was exponentially higher for hand-tracking (μ = 399, SD. = 243.65; μ = 371, SD. = 92.1) as well. It indicated that subjects found it harder to grab objects using hand-tracking. This difficulty, therefore, can be one reasonable explanation for the longer completion times.

Table 1
The mean and standard deviation for the four different dependent variables, which have been split by the independent variables for all $N = 33$ participants.

Item	Condition	Mean(μ)		Std.Dev (SD.)		N
		org	*log*	*org*	*log*	
AT	M1R1	314s	2.45	169.43	0.21	32
	M2R1	152s	2.14	67.52	0.19	32
	M1R2	272s	2.40	116.58	0.19	32
	M2R2	154s	2.12	104.12	0.24	32
OP	M1R1	32	1.44	17.04	0.25	32
	M2R1	21	1.26	9.20	0.31	32
	M1R2	29	1.43	11.82	0.17	32
	M2R2	18	1.20	8.30	0.28	32
CF	M1R1	399	2.54	243.65	0.23	32
	M2R1	58	1.70	37.15	0.23	32
	M1R2	371	2.52	192.1	0.20	32
	M2R2	66	1.70	65.10	0.29	32

4.2. Sense of presence

A two-way MANOVA with covariate was conducted to control for AT. This was done in consideration of the possible effects the duration of time spent within the VE may have on the four IPQ items. The maximum Mahalanobis distance value was checked for assumption testing. It was 15.55, which is less than the critical value of 18.47 ($df = 4$) required for multivariate normality (refer to Table 2). Results of Box's M Test of Equality of Covariance Matrices showed no violation of the assumption ($p = 0.59$). All significance values ($P_{GP} = 0.64$, $P_{SP} = 0.59$, $P_{INV} = 0.47$, and $P_{REAL} = 0.55$) in Levene's Test of Equality of Error Variances were more than 0.05. The scores for the IPQ-items were compiled on a likert scale. No statistically significant effects were obtained from the MANOVA results. There was no difference in the means of the four dependent variables of IPQ (GP, SP, INV, and REAL) (see Table 3).

Firstly, there is a non-significant effect of modality type on IPQ scores, M: $F(4, 120) = 0.734$, $p = 0.57$, Wilk's $\lambda = 0.98$, $\eta^2 p = 0.024$. Secondly, there is a non-significant effect of representation level on IPQ scores, R: $F(4, 120) = 0.670$, $p = 0.61$, Wilk's $\lambda = 0.98$, $\eta^2 p = 0.022$. Finally, there is a non-significant interaction effect between modality and representation on IPQ scores, M x R: $F(4, 120) = 0.158$, $p = 0.96$, Wilk's $\lambda = 0.995$, $\eta^2 p = 0.005$.

We, therefore, fail to reject the null hypothesis, and that neither input modality nor visual representation bore significant influence when considered jointly for the user-reported IPQ scores. Table 3 shows Means (μ) and SDs of the IPQ sub-scales as a function of M-type and R-Level; charts are shown in Fig. 5.

4.3. Mental WorkLoad – MWL

The perceived cognitive workload, or MWL, for the virtual reach-grab-place task was evaluated using the NASA-TLX. A two-way MANOVA assessed the effects of modality and representation on users' NASA-TLX scores. As a preliminary step, we checked for MANOVA assumptions. Logarithmic transformation was applied so that all data presented normal distribution. The result of Box's M Test of Equality of Covariance Matrices was $p = 0.50$. Values for the Levene's Test of Equality of Error Variances were: $P_{MD} = 0.31$, $P_{PD} = 0.48$, $P_{TD} = 0.53$, $P_{RD} = 0.16$), $P_{ED} = 0.42$), and $P_{FD} = 0.05$.

All four conditions were analyzed across the six sub-scales. No significant effects for representation levels were found on the NASA-TLX scores, R: textitF $(6, 119) = 0.136$, $p = 0.99$, Wilk's $\lambda = 0.99$, $\eta^2 p = 0.007$. There were also no significant interaction effects between modality and representation on the indeces, M x R: textitF $(6, 119) = 0.684$, $p = 0.66$, Wilk's $\lambda = 0.97$, $\eta^2 p = 0.033$.

The MANOVA revealed a significant main effect of modality type on the NASA-TLX indeces, M: textitF $(6, 119) = 8.374$, $p = < 0.001$, Wilk's $\lambda = 0.703$, $\eta^2 p = 0.30$. Univariate ANOVAs were conducted to examine the specific effects of modality type on each subscale item. Below, we look at the test for between-subject effects:

Table 2
The critical chi-square values for evaluating Mahalanobis Distance at a critical alpha of 0.001 are shown below (Pallant, 2020). Values are shown from 2 to 10 degrees of freedom.

df	critical value	df	critical value	df	critical value
2	13.82	5	20.52	8	26.13
3	16.27	6	22.46	9	27.88
4	18.47	7	24.32	10	29.59

Table 3
The mean(μ) and standard deviation (SD.) for the four different dependent variables of IPQ, which have been split by the independent variables for all $N = 32$ participants.

IPQ	M1R1		M1R2		M2R1		M2R2	
	μ	SD.	μ	SD.	μ	SD.	μ	SD.
GP	4.63	1.62	4.31	1.67	4.93	1.54	4.77	1.62
SP	4.21	0.71	4.18	0.77	4.27	0.78	4.20	0.89
INV	4.50	1.37	4.37	1.21	4.81	1.20	4.52	1.23
REAL	3.61	0.65	3.42	0.70	3.49	0.79	3.35	0.63

Table 4
The mean and standard deviation for the six different dependent variables of NASA-TLX, which have been split by the independent variables for all $N = 32$ participants.

Item	Condition	Mean(μ)		Std. Dev (SD.)		N
		org	*log*	*org*	*log*	
MD	M1R1	4.06	1.82	1.86	0.87	32
	M2R1	2.53	1.05	1.65	0.92	32
	M1R2	4.06	1.89	1.52	0.70	32
	M2R2	2.38	1.06	1.24	0.75	32
PD	M1R1	4.06	1.88	1.66	0.69	32
	M2R1	2.59	1.17	1.39	0.79	32
	M1R2	4.19	1.95	1.49	0.60	32
	M2R2	2.53	1.20	1.08	0.66	32
TD	M1R1	4.25	1.92	1.90	0.73	32
	M2R1	2.81	1.257	1.73	0.85	32
	M1R2	4.13	1.91	1.56	0.70	32
	M2R2	2.78	1.27	1.48	0.80	32
RD	M1R1	3.03	1.39	1.51	0.84	32
	M2R1	3.16	1.45	1.51	0.84	32
	M1R2	3.22	1.55	1.31	0.70	32
	M2R2	2.84	1.26	1.65	0.89	32
ED	M1R1	3.60	1.62	1.90	0.85	32
	M2R1	2.75	1.21	1.72	0.85	32
	M1R2	3.97	1.86	1.62	0.62	32
	M2R2	2.47	1.04	1.63	0.88	32
FD	M1R1	4.84	2.15	1.72	0.68	32
	M2R1	3.34	1.53	1.62	0.85	32
	M1R2	4.81	2.18	1.47	0.55	32
	M2R2	3.06	1.47	1.30	0.69	32

- MD: F (1, 124) = 30.5, $p = < 0.001$, $\eta^2 p = 0.20$
- PD: F (1, 124) = 36.41, $p = < 0.001$, $\eta^2 p = 0.23$
- TD: F (1, 124) = 23.2, $p = < 0.001$, $\eta^2 p = 0.16$
- RD: F (1, 124) = 0.60, $p = 0.44$, $\eta^2 p = 0.005$
- ED: F (1, 124) = 18.8, $p = < 0.001$, $\eta^2 p = 0.132$
- FD: F (1, 124) = 28.9, $p = < 0.001$, $\eta^2 p = 0.19$

Significant difference was observed across all indices but PD ($p = 0.44$). The results indicate that the overall perceived workload was higher for hand-tracking (M1) compared to hand-controllers (M2) but barely diverged between saturated (R1) and grayscaled (R2) representations. Table 4 shows both the log-transformed and original means(μ) and standard deviations (SD.) in each group. Less cognitive workload was required when subjects used the handheld-controller to complete the task. Handheld-controllers were least imposing as evidenced by the means for M2R2 across all indices: MD (μ = 2.38, SD = 1.24), PD (μ = 2.53, SD = 1.08), TD (μ = 2.78, SD = 1.48), RD (μ = 2.84, SD = 1.65), ED (μ = 2.47, SD = 1.63), and FD (μ = 3.06, SD = 1.30).

4.4. User experience – UX

The results from the AttrakDiff were compiled into three dimensions of pragmatic (PQ) and hedonic quality (HQ), and attractiveness (ATT). Prior to a two-way MANOVA, the Mahalanobis distance was checked for assumption testing. It maximum value was 14.95, which is less than the critical value of 16.27 (df = 3) required for multivariate normality (refer to Table 2). Results of Box's M Test of Equality of Covariance Matrices showed no violation of the assumption ($p = 0.57$). All significance values $P_{PQ} = 0.16$, $P_{HQ} = 0.41$, and $P_{ATT} = 0.15$) in Levene's Test of Equality of Error Variances were more than 0.05.

No significant main effects were noticed for R-Level: textitF (3, 122) = 1.953, $p = 0.125$, Wilk's $\lambda = 0.95$, $\eta^2 p = 0.05$. Nor were any interaction effect revealed for, (M x R): textitF (3, 122) = 0.335, $p = 0.80$, Wilk's $\lambda = 0.99$, $\eta^2 p = 0.08$. The two-way MANOVA did however reveal statistically significant the means of PQ, HQ and ATT when based on M-type: F (3, 122) = 7.953, $p = < 0.001$, Wilk's $\lambda = 0.84$, $\eta^2 p = 0.16$.

Follow-up univariate ANOVAs examined the specific differences for M and R on the three dependent variable individually to identify specific differences. Modality type had a significant influence on the usability (PQ: $p = < 0.001$) and desirability (ATT: $p = 0.033$) of the VR application. Whereas representation levels significantly affected intrigue (HQ: $p = 0.025$) and desirability (ATT: $p = 0.038$) only.

We further examined the results with the online AttrakDiff tool that outputs the following "Result Diagrams": portfolio-presentation, diagram of average values, and description of word-pairs. Fig. 7 depicts the overlapped portfolio-presentation results for the four conditions. M1R1 is located in "neutral" edging towards "self-oriented"; M1R2 is within "neutral"; M2R2 falls within "task-oriented"; and M2R1 can be seen in the "desired" position with a slight tendency towards "task-oriented". All confidence intervals are pretty similar in size, indicating that there was a general agreement amongst participants in terms of the hedonic and pragmatic qualities of the conditions. However, M2R1 visibly has the smallest confidence rectangle in the group implying higher reliability and a less coincidental result. The overlapped portfolio-presentation clearly demonstrates that confidence intervals are overlapping based on M-type, for example, blue over red, and green over yellow. This corroborates with the results from our analysis above. Subjects found the handheld-controller modality (M2R1 and M2R2) better oriented for performing tasks than the hand-tracking modality

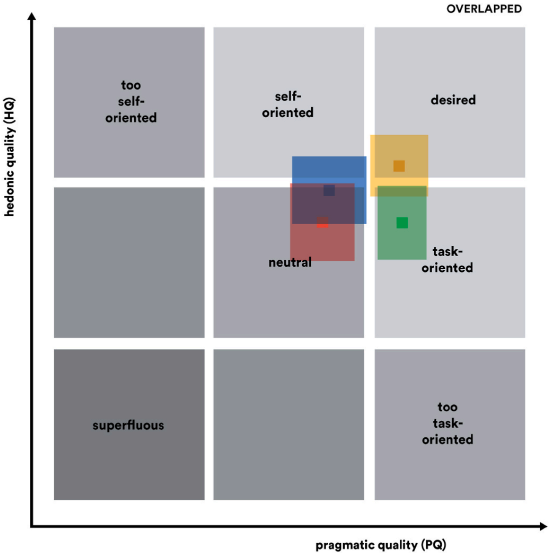


Fig. 7. An overlapped Portfolio-Presentation for the four experimental conditions. Color legend: M1R1 (blue), M1R2 (red), M2R2 (green), M2R1 (yellow). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(M1R1 and M1R2). Overall, M2R1 exhibited higher perceived quality for participants.

Findings from the portfolio-presentation can be compared against the diagram of average values (Fig. 8) for further insight into perceived user experience. The results confirm that M2R1 (PQ = 1.40, HQ = 1.24, ATT = 1.85) generally does better (evident by the yellow line). However, M2R2 scores higher at PQ. M1R2 has the lowest performance (PQ = 0.43, HQ = 0.53, ATT = 0.90). The scores for M2R2 oscillate from a high for perceived pragmatic quality to a low score for hedonic quality, (PQ = 1.44, HQ = 0.86, ATT = 1.28). Overall, perceived usability remains higher than emotional response across all conditions.

Finally, (Fig. 9) shows the results of the description of word-pairs with all scores of the 10 different word-pairs. From the figure, the higher scores for M2R1 in terms of user experience are obvious. This condition shows superior performance in almost all word pairs (except confusing—clearly structured). An encouraging takeaway is that all conditions are within the positive user experience range. We can deduce that hand-controllers engender higher perceived usability for participants while at the same time visual realism had a positive emotional impact. This trend is most evident in M2R1, which also had the highest global value for quality perception in terms of appeal and pleasantness.

5. Discussion

For a reach-grab-place task in immersive VR, we systematically compared two input modalities (hand-tracking vs.handheld-controller) in two visual representation levels (saturated vs. grayscaled). We measured objective user performance metrics and subjective user experiences of perceived sense of presence, mental workload, and ease-of-use.

Generally speaking, the two input modalities can be used in VR to different effects. Hand-tracking allows users to interact with the virtual environment using natural hand gestures whereas a handheld-controller uses buttons and triggers to perform actions. Since virtual hands enable a more naturalistic interaction compared to handheld-controllers, we hypothesized higher ratings on naturalness and intuitiveness.

5.1. On performance and effectiveness

We found that the handheld-controller input modality was more effective for completing the reach-grab-place task, as it resulted in faster task completion times and better object manipulation compared to the hand-tracking. This was reflected in the mental workload scores where subjects recorded lower loads for handheld-controllers, indicating that it may be less cognitively demanding to use. Similarly, subjects also ranked the handheld-controller modality higher on task-orientedness and the overall appeal of the modality for the said task. However, subjects did not find any significant differences vis-a-vis immersion, involvement, and realness between the two input modalities inside VR. This suggests that while the handheld-controller may be more effective and preferred in terms of practical performance, it does not necessarily have a greater impact on the subjective experience of VR for the user. Surprisingly, the level of visual representation (saturated to grayscaled) also did not seem to have influenced the subjective experience of VR for the subjects.

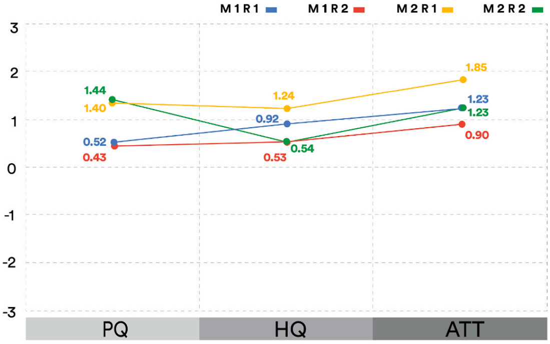


Fig. 8. Diagram of Average Values for the four experimental conditions. Color legend: M1R1 (blue), M1R2 (red), M2R2 (green), M2R1 (yellow). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

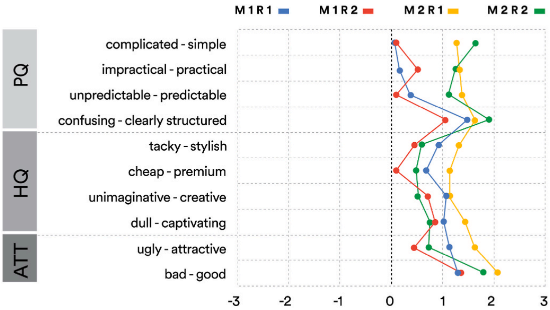


Fig. 9. Description of Word-Pairs for the four experimental conditions.

It is generally considered that virtual experiences that involve physical movements can enhance both spatial presence and mental immersion. In particular, using interfaces that recall hand gestures from daily life is poised to carry a higher sensation (De Paolis & De Luca, 2020). But our findings confirm the hesitation expressed in earlier works that found hand-tracking modules to be perceptively more difficult for performance (Caggianese et al., 2018; Ricca et al., 2020; Voigt-Antons et al., 2020). This could be because of unfulfilled user expectations about the capabilities of hand-tracking (Myers et al., 2019) spurlock-2019. There are several factors to consider when comparing hand-tracking and handheld-controller input modalities. One important factor is the accuracy and reliability of the tracking. Hand-tracking systems may sometimes have difficulty accurately tracking the user's hands, particularly if the user is making fast or complex gestures. There is no official hand-tracking accuracy rate specified for the Oculus Quest device but previous works have found it suitable for a wide range of applications Holzwarth et al. (2021) (Carnevale et al., 2022). Another common limitation with VR headsets (like the Oculus Quest) that provide inside-out tracking is having trouble detecting physical hands due to self-occlusion (Pacchierotti et al., 2016; Rehg & Kanade, 1995). Several subjects faced this issue while performing the task which hampered their experience. Handheld-controllers, on the other hand, offered a more precise and reliable means of input, as they are not reliant on the headset's tracking capabilities.

5.2. On actions and expectations

It is also important to consider the type of tasks that the user will be performing. In our case, the reach-grab-place task used in this experiment required the subjects to use different prehension types to grab the specific objects. We had considered that hand-tracking may be more suitable for the reach-grab-place task – due to the fine motor skills or precise hand movements involved – as it will allow users to use their own hands rather than relying on a physical controller. Here it is worthwhile to revisit the grips illustrated in (Fig. 3) of this document. The illustrations show six distinct ways in which the six respective objects are to be held. This is contrasted by the simplistic single "pinch" gesture currently available on the VR system. For example, lifting the paper clips using the *terminal opposition* grip had the best coincidence with the ”pinch” gesture supported by the VR headset. A close second was the *subterminal opposition* grip used for grabbing the two business cards. However, subjects had to use the same pinch gesture to grab and move the virtual mobile phone. This is a sharp departure from the real-world operation, or natural prehension, where the panoramic penta-digital grip is often used to lift objects of that shape. Isomorphic mapping is responsible for establishing a relationship between elements in one system (such as physical movements or actions) and corresponding elements in another system (such as virtual movements or actions) (Hutchins et al., 1985). This can be useful for creating a sense of immersion or presence in a VE but as we found out a mismatch could result in the opposite. Technological limitations with isomorphic mapping of real hands may contribute to lower scores. This is because in real-world conditions an activity like holding a mug would prompt an instantaneous, automatic response to the stimuli by the subjects. The grip mismatch in VR suddenly demanded more mental processing for the same task. Something that should have been *ready-at-hand* (transparent) was all at once *present-at-hand* (opaque) (Coyne, 1994). It is imaginable why such interactions might be perceived as counter-intuitive and why such operational mismatches might negatively affect the user-perceived realism or naturalness of the environment (Hameed & Perkis, 2021; Weber et al., 2021).

5.3. On usability and demand

One more factor to consider is the user’s level of comfort and ease of use. In our analysis, we found a significant difference between the handheld-controller and hand-tracking in terms of usability – the former was perceived as being more useable. We found that the controller-free modality continuously performed in the ”neutral” zone, which is not discouraging but did not receive the high user experience valuation we had hypothesized (Masurovsky et al., 2020). Overall, the controller-in-a-realistic-scenario (M2R1 variation) outperformed across the AttrakDiff dimensions showing that user experience draws a sensitive balance between achieving goals, maintaining desirability, and remaining instinctive at the same time.

The preference for the controller modality was most pronounced across the mental workload indices of NASA-TLX where users indicated a higher effort of the mental processing of information and individual reactions when using virtual hands. Mental-, physical-, and temporal-demands almost halved when users switched from virtual hands to controller-based interactions. This suggests that controller-based interactions had better intuitiveness, contrary to our initial hypothesis, and hand-tracking was perceived as significantly frustrating and demanding. This can also be attributed to learned digital literacies (Riecke et al., 2018), since all (N = 32) subjects had reported prior knowledge of controllers. It would appear that though closer-to-natural, hand-tracking still required some getting-used-to before requisite proficiency for the VR task could be met. So a learnability component may have influenced perception (Drew et al., 2018). Conversely, handheld-controllers offered a more familiar, stable, and ergonomic means of grabbing the objects; and, because subjects using controllers did not expect to grab virtual objects with natural prehension, this is why effects of expectation mismatch were also absent. The results indicate that unreliable behavior of the input modality can adversely impact user performance and overall experience.

5.4. Limitations

In our study we only assessed the wireless Oculus Quest 1 due to its affordability and wide access. This can be seen as a limitation. Also because this was not the most advanced, or current, inside-out system at the time. The use of a high performing and stable system may have provided a more responsive experience. Another noticeable limitation was the requisite digital literacy of the participants. Hand-tracking technology is still in its nascency. Familiarity with the technology, understanding of the interface, and comfort of use amongst users is still far limited compared to the ubiquitous handheld-controller devices. This limitation should be met as the technology becomes mainstream. Finally, and in retrospect, only 2 out of 6 grip types coincide with the gesture supported by the VR system. The number and/or variety of prehension types used for the task could be reduced for efficiency purposes. Especially because despite the simplicity of the reach-grab-place task, an obvious limitation was the availability of only a single selection and manipulation gesture, *pinch*, on the Oculus Quest system.

6. Conclusion

In summary, our study looked at the potential for common implementation of hand-tracking VR interface by comparing it to handheld-controller in a virtual reach-grab-place task representative of real-world motor performance. We also investigated if enhancing the visual realism (level of representation) of the environment alongside natural gestures improved subjective evaluations of presence, mental workload, and ease-of-use. The results of the statistical analysis show that visual realism had no effect on user performance and surprisingly nor did it have an effect on their subjective experience of the VE. Regarding interaction, we found that input modality did bear significant influence on user performance and overall experience. Subjects took lesser attempts at grabbing virtual objects using handheld-controllers and reported efficient completion times. With hand-tracking they took longer to complete the same task and reported higher perceived mental workload scores. There was a significant difference in the ease-of-use of the two modalities. Subjects found handheld-controllers to be more task-oriented and appealing compared to hand-tracking. Lastly, the subjective feeling of immersion, perceived realism, and engagement within the VE did not differ much across the four experimental variations. Our results do not support the hypothesis of higher naturalness and user experience for hand-tracking in its current state. However, as familiarity with hand-tracking increases and technical issues are progressively overcome, this may change. Just like touch interfaces were inferior to mouse point-and-click for many years before becoming commonplace. From a research perspective, it would be interesting to see if iterative improvements in hand-tracking technologies – enhanced scope and range of available gestures – may come to surpass handheld-controllers in the future. The results of this study make a good case for taking a closer look at performative and experiential aspects of gesture-based modalities. The authors are currently investigating the particularities of manipulation in hand-tracking with respect to its action possibilities, or affordances. For future research, the perception of object affordances within VR has been highlighted as an area for investigation, because understanding object manipulation from an affordance point-of-view can help achieve interfaces and mechanisms that are effective and efficient for VR interactions.

Author agreement statement

We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere. We confirm that the manuscript has been read and approved by all named authors and that there are no other people who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We understand that the Corresponding Author is the sole contact for the Editorial process. He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

References

Abdlkarim, D., Di Luca, M., Aves, P., Yeo, S. H., Miall, R. C., Holland, P., & Galea, J. M. (2023). A methodological framework to assess the accuracy of virtual reality hand-tracking systems: A case study with the Meta Quest 2. *Behavior Research Methods*, 1–12.

Abidi, M. H., Al-Ahmari, A., Ahmad, A., Ameen, W., & Alkhalefah, H. (2019). Assessment of virtual reality-based manufacturing assembly training system. *The International Journal of Advanced Manufacturing Technology, 105*, 3743–3759.

Alexandrovsky, D., Putze, S., Bonfert, M., Höffner, S., Michelmann, P., Wenig, D., Malaka, R., & Smeddinck, J. D. (2020). Examining design choices of questionnaires in vr user studies. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–21).

Argelaguet, F., Hoyet, L., Trico, M., & Lécuyer, A. (2016). The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE virtual reality (VR)* (pp. 3–10). IEEE.

Berki, B. (2020). Experiencing the sense of presence within an educational desktop virtual reality. *Acta Polytechnica Hungarica, 17*, 255–265.

Billger, M., Heldal, I., Stahre, B., & Renstrom, K. (2004). Perception of color and space in virtual reality: A comparison between a real room and virtual reality models. In *Human vision and electronic imaging IX* (pp. 90–98). SPIE.

Biocca, F. (2002). The evolution of interactive media. Narrative impact. *Social and cognitive foundations*, 97–130.

Blaga, A. D., Frutos-Pascual, M., Creed, C., & Williams, I. (2020). Too hot to handle: An evaluation of the effect of thermal visual representation on user grasping interaction in virtual reality. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–16).

Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2001). An introduction to 3-d user interface design. *Presence, 10*, 96–108.

Bowman, D. A., McMahan, R. P., & Ragan, E. D. (2012). Questioning naturalism in 3d user interfaces. *Communications of the ACM, 55*, 78–88.

Brunnström, K., Beker, S. A., De Moor, K., Dooms, A., Egger, S., Garcia, M. N., Hossfeld, T., Jumisko-Pyykkö, S., Keimel, C., Larabi, M. C., et al. (2013). *Qualinet white paper on definitions of quality of experience*.

Buttussi, F., & Chittaro, L. (2019). Locomotion in place in virtual reality: A comparative evaluation of joystick, teleport, and leaning. *IEEE Transactions on Visualization and Computer Graphics, 27*, 125–136.

Caggianese, G., Gallo, L., & Neroni, P. (2018). The vive controllers vs. leap motion for interactions in virtual environments: A comparative evaluation. In *International conference on intelligent interactive multimedia systems and services* (pp. 24–33). Springer.

Carnevale, A., Mannocchi, I., Sassi, M. S. H., Carli, M., De Luca, G., Longo, U. G., Denaro, V., & Schena, E. (2022). Virtual reality for shoulder rehabilitation: Accuracy evaluation of oculus quest 2. *Sensors, 22*, 5511.

Chellali, A., Mentis, H., Miller, A., Ahn, W., Arikatla, V. S., Sankaranarayanan, G., De, S., Schwaitzberg, S. D., & Cao, C. G. (2016). Achieving interface and environment fidelity in the virtual basic laparoscopic surgical trainer. *International Journal of Human-Computer Studies, 96*, 22–37.

Chen, F., Lin, Y. C., Chien, J. W., & Tsai, C. E. (2016). Virtual reality for digital user experience and interactive learning based on user satisfaction: A pilot study. In *2016 international conference on computational science and computational intelligence (CSCI)* (pp. 374–377). IEEE.

Chiu, P., Takashima, K., Fujita, K., & Kitamura, Y. (2019). Pursuit sensing: Extending hand tracking space in mobile vr applications. In *Symposium on spatial user interaction* (pp. 1–5).

Coyne, R. (1994). Heidegger and virtual reality: The implications of heidegger's thinking for computer representations. *Leonardo, 27*, 65–73.

Crocetta, T. B., de Araújo, L. V., Guarnieri, R., Massetti, T., Ferreira, F. H. I. B., De Abreu, L. C., & de Mello Monteiro, C. B. (2018). Virtual reality software package for implementing motor learning and rehabilitation experiments. *Virtual Reality, 22*, 199–209.

De Paolis, L. T., & De Luca, V. (2020). The impact of the input interface in a virtual environment: The vive controller and the myo armband. *Virtual Reality, 24*, 483–502.

Drew, M. R., Falcone, B., & Baccus, W. L. (2018). What does the system usability scale (sus) measure?. In *International conference of design, user experience, and usability* (pp. 356–366). Springer.

Feick, M., Kleer, N., Tang, A., & Krüger, A. (2020). The virtual reality questionnaire toolkit. In *Adjunct publication of the 33rd annual ACM symposium on user interface software and technology* (pp. 68–69).

Feix, T., Romero, J., Schmiedmayer, H. B., Dollar, A. M., & Kragic, D. (2015). The grasp taxonomy of human grasp types. *IEEE Transactions on human-machine systems, 46*, 66–77.

Felton, W. M. (2021). *Presence in virtual environments: Visual factors and measure convergence*. Ph.D. thesis. University of Idaho.

Fromberger, P., Meyer, S., Kempf, C., Jordan, K., & Müller, J. L. (2015). Virtual viewing time: The relationship between presence and sexual interest in androphilic and gynephilic men. *PLoS One, 10*, Article e0127156. https://doi.org/10.1371/journal.pone.0127156

Furmanek, M. P., Schettino, L. F., Yarossi, M., Kirkman, S., Adamovich, S. V., & Tunik, E. (2019). Coordination of reach-to-grasp in physical and haptic-free virtual environments. *Journal of NeuroEngineering and Rehabilitation, 16*, 1–14.

Graf, S., & Schwind, V. (2020). Inconsistencies of presence questionnaires in virtual reality. In *26th ACM symposium on virtual reality software and technology* (pp. 1–3).

Grimm, F., Naros, G., & Gharabaghi, A. (2016). Closed-loop task difficulty adaptation during virtual reality reach-to-grasp training assisted with an exoskeleton for stroke rehabilitation. *Frontiers in Neuroscience, 10*, 518.

Haar, S., Sundar, G., & Faisal, A. A. (2021). Embodied virtual reality for the study of real-world motor learning. *PLoS One, 16*, Article e0245717.

Hamblin, C. J. (2005). *Transfer of* training from virtual reality environments*. Wichita State University.

Hameed, A., & Perkis, A. (2021). A subjective and behavioral assessment of affordances in virtual architectural walkthroughs. *Applied Sciences, 11*, 7846.

Hameed, A., Perkis, A., & Möller, S. (2021). Evaluating hand-tracking interaction for performing motor-tasks in vr learning environments. In *2021 13th international conference on quality of multimedia experience (QoMEX)* (pp. 219–224). IEEE.

Hart, S. G. (2006). Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (pp. 904–908). Los Angeles, CA: Sage publications Sage CA.

Hassenzahl, M., Koller, F., Burmester, M., 2008. Der user experience (ux) auf der spur: Zum einsatz von www. attrakdiff. de. Tagungsband UP08 ..

Holderied, H. (2017). Evaluation of interaction concepts in virtual reality applications. *Informatik, 2017*.

Holzwarth, V., Gisler, J., Hirt, C., & Kunz, A. (2021). Comparing the accuracy and precision of steamvr tracking 2.0 and oculus quest 2 in a room scale setup. In *2021 the 5th international conference on virtual and augmented reality simulations* (pp. 42–46).

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1985). Direct manipulation interfaces. *Human-Computer Interaction, 1*, 311–338.

ITU-T. (1998). *Recommendation P.911. Subjective audiovisual quality assessment methods for multimedia applications*. https://handle.itu.int/11.1002/1000/4538.

ITU-T. (2018). *Recommendation P.809. Subjective evaluation methods for gaming quality*. https://handle.itu.int/11.1002/1000/13626.

ITU-T. (2020). *Recommendation P.919. Subjective test methodologies for 360° video on head-mounted displays*. https://handle.itu.int/11.1002/1000/14429.

Izard, S. G., Juanes, J. A., Peñalvo, F. J. G., Estella, J. M. G., Ledesma, M. J. S., & Ruisoto, P. (2018). Virtual reality as an educational and training tool for medicine. *Journal of Medical Systems, 42*, 1–5.

Jankowski, J., & Grabowski, A. (2015). Usability evaluation of vr interface for mobile robot teleoperation. *International Journal of Human-Computer Interaction, 31*, 882–889.

Jerald, J. (2015). *The VR book: Human-centered design for virtual reality*. Morgan & Claypool.

Kaliki, R. R., Davoodi, R., & Loeb, G. E. (2012). Evaluation of a noninvasive command scheme for upper-limb prostheses in a virtual reality reach and grasp task. *IEEE Transactions on Biomedical Engineering, 60*, 792–802.

Kamakura, N., Matsuo, M., Ishii, H., Mitsuboshi, F., & Miura, Y. (1980). Patterns of static prehension in normal hands. *American Journal of Occupational Therapy, 34*, 437–445.

van de Kamp, C., & Zaal, F. T. (2007). Prehension is really reaching and grasping. *Experimental Brain Research, 182*, 27–34.

Kapandji, I. A. (1987). *The physiology of the joint*. lower limb. 2.

Kilteni, K., Groten, R., & Slater, M. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments, 21*, 373–387.

Kim, M., Jeon, C., & Kim, J. (2017). A study on immersion and presence of a portable hand haptic system for immersive virtual reality. *Sensors, 17*, 1141.

Kim, H. K., Rattner, D. W., & Srinivasan, M. A. (2003). The role of simulation fidelity in laparoscopic surgical training. In *International conference on medical image computing and computer-assisted intervention* (pp. 1–8). Springer.

Lackey, S., Salcedo, J., Szalma, J., & Hancock, P. (2016). The stress and workload of virtual reality training: The effects of presence, immersion and flow. *Ergonomics, 59*, 1060–1072.

LaViola, J. J., Jr., Kruijff, E., McMahan, R. P., Bowman, D., & Poupyrev, I. P. (2017). *3D user interfaces: Theory and practice*. Addison-Wesley Professional.

Levin, M. F., Magdalon, E. C., Michaelsen, S. M., & Quevedo, A. A. (2015). Quality of grasping and the role of haptics in a 3-d immersive virtual reality environment in individuals with stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 23*, 1047–1055.

Liagkou, V., Salmas, D., & Stylios, C. (2019). Realizing virtual reality learning environment for industry 4.0. *Procedia CIRP, 79*, 712–717.

Liou, H. H., Yang, S. J., Chen, S. Y., & Tarng, W. (2017). The influences of the 2d image-based augmented reality and virtual reality on student learning. *Journal of Educational Technology & Society, 20*, 110–121.

Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). Serious games analytics: Theoretical framework. *Serious games analytics: Methodologies for performance measurement, assessment, and improvement, 3–29*.

Lougiakis, C., Katifori, A., Roussou, M., & Ioannidis, I. P. (2020). Effects of virtual hand representation on interaction and embodiment in hmd-based virtual environments using controllers. In *2020 IEEE conference on virtual reality and 3D user interfaces* (pp. 510–518). IEEE: VR).

Luro, F. L., & Sundstedt, V. (2019). A comparative study of eye tracking and hand controller for aiming tasks in virtual reality. In *Proceedings of the 11th ACM symposium on eye tracking research & applications* (pp. 1–9).

Masurovsky, A., Chojecki, P., Runde, D., Lafci, M., Przewozny, D., & Gaebler, M. (2020). Controller-free hand tracking for grab-and-place tasks in immersive virtual reality: Design elements and their empirical study. *Multimodal Technologies and Interaction, 4*, 91.

McMahan, R. P., Lai, C., & Pal, S. K. (2016). Interaction fidelity: The uncanny valley of virtual reality interactions. In *International conference on virtual, augmented and mixed reality* (pp. 59–70). Springer.

Mekbib, D. B., Han, J., Zhang, L., Fang, S., Jiang, H., Zhu, J., Roe, A. W., & Xu, D. (2020). Virtual reality therapy for upper limb rehabilitation in patients with stroke: A meta-analysis of randomized clinical trials. *Brain Injury, 34*, 456–465.

Merians, A. S., Poizner, H., Boian, R., Burdea, G., & Adamovich, S. (2006). Sensorimotor training in a virtual reality environment: Does it improve functional recovery poststroke? *Neurorehabilitation and Neural Repair, 20*, 252–267.

Myers, C. M., Furqan, A., & Zhu, J. (2019). The impact of user characteristics and preferences on performance with an unfamiliar voice user interface. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–9).

Nilsson, N. C., Nordahl, R., & Serafin, S. (2016). Immersion revisited: A review of existing definitions of immersion and their relation to different theories of presence. *Human Technology, 12*.

Nunnally, M. E., & Bitan, Y. (2006). Time to get off this pig's back?: The human factors aspects of the mismatch between device and real-world knowledge in the health care environment. *Journal of Patient Safety, 2*, 124–131.

Oculus, V. (2020). Hand tracking in unity | oculus developers. URL: https://developer.oculus.com/documentation/unity/unity-handtracking/.

Oculus, V. R. (2020). *Oculus device specifications | Oculus developers*. Inc.". URL: https://developer.oculus.com/resources/oculus-device-specs/.

Pacchierotti, C., Salvietti, G., Hussain, I., Meli, L., & Prattichizzo, D. (2016). The hring: A wearable haptic device to avoid occlusions in hand tracking. In *2016 IEEE haptics symposium (HAPTICS)* (pp. 134–139). IEEE.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

Palmas, F., Labode, D., Plecher, D. A., & Klinker, G. (2019). Comparison of a gamified and non-gamified virtual reality training assembly task. In *2019 11th international conference on virtual worlds and games for serious applications (VS-games)* (pp. 1–8). IEEE.

Pastel, S., Petri, K., Bürger, D., Marschal, H., Chen, C. H., & Witte, K. (2022). Influence of body visualization in vr during the execution of motoric tasks in different age groups. *PLoS One, 17*, Article e0263112.

Perez-Marcos, D., Chevalley, O., Schmidlin, T., Garipelli, G., Serino, A., Vuadens, P., Tadi, T., Blanke, O., & Millán, J.d. R. (2017). Increasing upper limb training intensity in chronic stroke using embodied virtual reality: A pilot study. *Journal of NeuroEngineering and Rehabilitation, 14*, 1–14.

Perkis, A., Timmerer, C., Baraković, S., Husić, J. B., Bech, S., Bosse, S., Botev, J., Brunnström, K., Cruz, L., De Moor, K., et al. (2020). *Qualinet white paper on definitions of immersive media experience (imex)*. arXiv preprint arXiv:2007.07032.

Pfeuffer, K., Mayer, B., Mardanbegi, D., & Gellersen, H. (2017). Gaze+ pinch interaction in virtual reality. In *Proceedings of the 5th symposium on spatial user interaction* (pp. 99–108).

Plooy, A. M., & Wann, J. P. (2000). Supporting natural prehension in virtual environments. In *Stereoscopic displays and virtual reality systems VII* (pp. 362–368). SPIE.

Pottle, J. (2019). Virtual reality and the transformation of medical education. *Future healthcare journal, 6*, 181.

Punako Jr, R., Thropp, S., . Grasping the future: Virtual hands control for fine motor tasks ..

Ragan, E. D., Bowman, D. A., Kopper, R., Stinson, C., Scerbo, S., & McMahan, R. P. (2015). Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. *IEEE Transactions on Visualization and Computer Graphics, 21*, 794–807.

Regazzoni, D., Rizzi, C., & Vitali, A. (2018). Virtual reality applications: Guidelines to design natural user interface. In *International design engineering technical conferences and computers and information in engineering conference*. American Society of Mechanical Engineers, Article V01BT02A029.

Rehg, J. M., & Kanade, T. (1995). Model-based tracking of self-occluding articulated objects. In *Proceedings of IEEE international conference on computer vision* (pp. 612–617). IEEE.

Ricca, A., Chellali, A., & Otmane, S. (2017). Study of interaction fidelity for two viewpoint changing techniques in a virtual biopsy trainer. In *2017 IEEE virtual reality (VR)* (pp. 227–228). IEEE.

Ricca, A., Chellali, A., & Otmane, S. (2020). Influence of hand visualization on tool-based motor skills training in an immersive vr simulator. In *2020 IEEE international symposium on mixed and augmented reality (ISMAR)* (pp. 260–268). IEEE.

Riecke, B. E., LaViola, J. J., Jr., & Kruijff, E. (2018). 3d user interfaces for virtual reality and games: 3d selection, manipulation, and spatial navigation. In *ACM SIGGRAPH 2018 courses* (pp. 1–94).

Sassatelli, L., Winckler, M., Fisichella, T., Dezarnaud, A., Lemaire, J., Aparicio-Pardo, R., & Trevisan, D. (2020). New interactive strategies for virtual reality streaming in degraded context of use. *Computers & Graphics, 86*, 27–41.

Schlesinger, G. (1919). The mechanical structure of the artificial limbs. In *Replacement links and work aids* (pp. 321–661). knight.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence: Teleoperators and Virtual Environments, 10*, 266–281.

Sheridan, T. B. (2016). Recollections on presence beginnings, and some challenges for augmented and virtual reality. *Presence: Teleoperators and Virtual Environments, 25*, 75–77.

Shin, Y., Lim, J., Kim, Y., Seo, D. G., & Ihm, J. (2022). Effects of virtual body-representation on motor skill learning. *Scientific Reports, 12*, Article 15283.

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3549–3557.

Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology, 109*, 431–433.

Sollerman, C. (1980). *The use of eight main hand grips in activities of daily living*. Thesis: Handens greppfunktion.

Steed, A., Pan, Y., Zisch, F., & Steptoe, W. (2016). The impact of a self-avatar on cognitive load in immersive virtual reality. In *2016 IEEE virtual reality (VR)* (pp. 67–76). IEEE.

Suznjevic, M., Mandurov, M., & Matijasevic, M. (2017). Performance and qoe assessment of htc vive and oculus rift for pick-and-place tasks in vr. In *2017 Ninth international conference on quality of multimedia experience (QoMEX)* (pp. 1–3). IEEE.

Sveistrup, H. (2004). Motor rehabilitation using virtual reality. *Journal of NeuroEngineering and Rehabilitation, 1*, 1–8.

Thorsteinsson, G. (2013). Developing an understanding of the pedagogy of using a virtual reality learning environment (vrle) to support innovation education. In *The Routledge International handbook of innovation education* (pp. 486–500). Routledge.

User interface design - UID GmbH". AttrakDiff. URL: http://www.attrakdiff.de/index-en.html.

Våpenstad, C., Hofstad, E. F., Langø, T., Mårvik, R., & Chmarra, M. K. (2013). Perceiving haptic feedback in virtual reality simulators. *Surgical Endoscopy, 27*, 2391–2397.

Viau, A., Feldman, A. G., McFadyen, B. J., & Levin, M. F. (2004). Reaching in reality and virtual reality: A comparison of movement kinematics in healthy subjects and in adults with hemiparesis. *Journal of NeuroEngineering and Rehabilitation, 1*, 1–7.

Voigt-Antons, J. N., Kojic, T., Ali, D., & Möller, S. (2020). Influence of hand tracking as a way of interaction in virtual reality on user experience. In *2020 twelfth international conference on quality of multimedia experience (QoMEX)* (pp. 1–4). IEEE.

Wang, Y., Wu, Y., Jung, S., Hoermann, S., Yao, S., & Lindeman, R. (2021). Enlarging the useable hand tracking area by using multiple leap motion controllers in vr. *IEEE Sensors Journal*.

Weber, S., Weibel, D., & Mast, F. W. (2021). How to get there when you are there already? Defining presence in virtual reality and the importance of perceived realism. *Frontiers in Psychology, 1538*.

Zhang, H. (2017). Head-mounted display-based intuitive virtual reality training system for the mining industry. *International Journal of Mining Science and Technology, 27*, 717–722.

Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Panton, O. N. M., & Atkins, M. S. (2012). Workload assessment of surgeons: Correlation between nasa tlx and blinks. *Surgical Endoscopy, 26*, 2746–2750.

ISBN 978-82-326-7950-8 (printed ver.)
ISBN 978-82-326-7949-2 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

NTNU
Norwegian University of
Science and Technology