frontiers | Research Topics

Beyond audiovisual: novel multisensory stimulation techniques and their applications

Edited by
Justyna Świdrak, Grzegorz Pochwatko, Orestis Georgiou,
Domna Banakou, Sofia Seinfeld and Marta Matamala-Gomez

Published in
Frontiers in Virtual Reality

ISSN 1664-8714
ISBN 978-2-8325-5523-1
DOI 10.3389/978-2-8325-5523-1

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Beyond audiovisual: novel multisensory stimulation techniques and their applications

Topic editors
Justyna Świdrak — August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Spain
Grzegorz Pochwatko — Institute of Psychology, Polish Academy of Sciences, Poland
Orestis Georgiou — Ultraleap Ltd., United Kingdom
Domna Banakou — New York University Abu Dhabi, United Arab Emirates
Sofia Seinfeld — Open University of Catalonia, Spain
Marta Matamala-Gomez — University of Barcelona, Spain

Citation
Świdrak, J., Pochwatko, G., Georgiou, O., Banakou, D., Seinfeld, S., Matamala-Gomez, M., eds. (2024). *Beyond audiovisual: novel multisensory stimulation techniques and their applications*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-5523-1

Table of contents

frontiers | Frontiers in Virtual Reality

TYPE Editorial
PUBLISHED 24 September 2024
DOI 10.3389/frvir.2024.1491854

OPEN ACCESS

EDITED AND REVIEWED BY
Anthony Steed,
University College London, United Kingdom

*CORRESPONDENCE
Justyna Świdrak,
justyna.swidrak@gmail.com

RECEIVED 05 September 2024
ACCEPTED 16 September 2024
PUBLISHED 24 September 2024

CITATION
Świdrak J, Georgiou O, Banakou D, Matamala-Gomez M, Pochwatko G and Seinfeld S (2024) Editorial: Beyond audiovisual: novel multisensory stimulation techniques and their applications.
Front. Virtual Real. 5:1491854.
doi: 10.3389/frvir.2024.1491854

Editorial: Beyond audiovisual: novel multisensory stimulation techniques and their applications

Justyna Świdrak[1,2]*, Orestis Georgiou[3], Domna Banakou[4], Marta Matamala-Gomez[5,6], Grzegorz Pochwatko[2] and Sofia Seinfeld[7]

[1]Event Lab, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain, [2]Institute of Psychology Polish Academy of Sciences, Warsaw, Poland, [3]Ultraleap Ltd., Bristol, United Kingdom, [4]Interactive Media, Arts and Humanities Division, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, [5]Brainvitge Cognition and Brain Plasticity Unit, Department of Cognition, Development, and Educational Psychology, University of Barcelona, Barcelona, Spain, [6]University School of Health and Sport (EUSES), University of Girona, Girona, Spain, [7]Open University of Catalonia, Barcelona, Spain

KEYWORDS

multisensory stimulation, virtual experience, touch, olfaction, gustatory stimulation, vestibular stimulation, navigation and spatial orientation, virtual nature

Editorial on the Research Topic
Beyond audiovisual: novel multisensory stimulation techniques and their applications

1 Introduction

The field of immersive virtual reality (VR) has rapidly evolved, driven by advancements in 3D graphics and accessible high-performance hardware. It is a multimodal technology based on multisensory integration. While much of the existing VR research has focused on audiovisual cues, humans perceive the world through a variety of senses, including touch, smell, taste, and less commonly considered senses like equilibrioception. Despite their importance, these additional sensory modalities have often been overlooked in VR development. This intentionally broad Research Topic aims to address that gap. It covers novel or underutilized sensory stimulation techniques and their applications in VR. To that end, we sought to collect manuscripts that push the boundaries of traditional VR by incorporating these diverse sensory inputs to enhance the immersive experience.

The present editorial summarizes the ten selected papers on multimodal VR applications that cover themes including navigation and spatial orientation, haptic feedback and touch, gustatory and olfactory stimuli, and therapeutic applications. We highlight significant technological advancements, methodological insights, and practical applications. Collectively, they underscore the significance and potential of multisensory integration in VR. In addition, these developments demonstrate the transformative potential to create more engaging, realistic, and effective experiences. Consequently, they pave the way for enhanced VR applications across education, entertainment, and healthcare.

2 Themes and content of the Research Topic

2.1 Enhancing navigation and spatial orientation

These studies emphasize the importance of combining physical movements and sensory inputs to enhance spatial orientation and user experience.

Adhikari et al. investigated the HeadJoystick, an embodied leaning-based flying interface, and its effects on performance and user experience during a 3D navigational search task in VR. Their findings suggest that leaning-based interfaces can offer a more intuitive and engaging VR navigation experience.

Kirollos and Herdman explored how the brain resolves conflicts between visual and vestibular inputs during self-motion perception using caloric vestibular stimulation (CVS) paired with visual stimuli in VR. They found that visual and vestibular cues are integrated equally by the nervous system to reduce perceptual uncertainties.

Takahashi et al. investigated the feasibility of using percutaneous electrical stimulation (PES) of ankle tendons to induce sensations of anteroposterior and lateral body tilt. Their technique presents a non-invasive, cost-effective method to provide realistic somatosensory feedback.

2.2 Haptic feedback and touch

These studies highlight the role of touch in enhancing presence, embodiment, and performance in VR.

Boban et al. examined the influence of active haptic feedback on the perception of finger movements and the dominance of visual cues. They concluded that active haptic feedback could override visual dominance, leading to improved accuracy in finger movement perception, and thus, enhance the realism and reliability of sensory feedback.

Desnoyers-Stewart et al. investigated the impact of performer-facilitated touch on presence and embodiment in immersive VR performances. The study found that real human touch significantly enhanced participants' sense of presence and embodiment, making the VR experience more engaging and emotionally impactful.

Sawahata et al. explored the effects of combining auditory and electrostatic force stimuli on visual field guidance in 360° VR. They demonstrated that combining these two modalities improved performance. The results prove that subtle haptic cues are effective in guiding visual attention and enhance user engagement in VR.

2.3 Olfactory and gustatory stimuli

These studies demonstrate how smells and tastes can alter perceptions, reduce stress, and enhance therapeutic outcomes. They showcase the potential of cross-modal correspondences.

Wu et al. explored the impact of ambient colours in VR on taste perception. Their findings suggest that ambient colours can indeed alter taste perceptions in VR. Results provide valuable insights for food-related VR applications and enhancing the overall user experience through cross-modal correspondence.

Lopes and Falk systematically reviewed the effectiveness of multisensory digital nature exposure, including olfactory stimuli, in reducing stress and anxiety. The review highlights the potential of integrating olfactory cues with audio-visual VR to enhance therapeutic outcomes and advocates for more standardized methodologies in future research.

2.4 Therapeutic applications

Freedman et al. presented the use of olfactory stimuli over VR exposure therapy for a combat veteran with PTSD. The integration of olfaction helped the patient recall and reprocess traumatic memories more effectively, consequently reducing the symptoms. This highlights the potential of multisensory VR in enhancing emotional processing and memory reconsolidation in PTSD treatment.

De Jesus Junior et al. tested the feasibility and effectiveness of a 3-week program with VR natural scenes, sounds, and scents. They found significant reductions in PTSD and depressive symptoms, cognitive improvements, and increased heart rate variability.

3 Broader implications

The selected studies introduce several novel hardware and software solutions that push the boundaries of multimodal VR. For instance, Adhikari et al. presented the HeadJoystick, an innovative leaning-based flying interface that enhances navigation by providing intuitive and embodied control for a more natural and engaging user experience. Similarly, Sawahata et al. explored the use of electrostatic force stimuli as a form of haptic feedback in 360° VR environments. This approach not only provides subtle yet effective guidance for visual attention but also demonstrates the feasibility of non-contact haptic feedback, opening new avenues for VR interaction design.

The papers also contribute to significant methodological advancements, offering robust experimental designs and comprehensive reviews that enhance our understanding of multisensory integration in VR. For example, Kirollos and Herdman methodological approach in investigating visual-vestibular integration can serve as a model for future studies aiming to explore the interplay between different sensory modalities in VR. Additionally, Lopes and Falk systematic review of the effects of multisensory digital nature exposure identified key methodological gaps that need to be addressed in future research.

These studies highlight the wide-ranging practical applications of multimodal VR applications, demonstrating the technology's broad potential in therapeutic, educational, and entertainment contexts. For instance, De Jesus Junior et al. and Freedman et al. studies shed light on the power of olfactory stimuli in VR exposure therapy for PTSD for both affective regulation and memory reconsolidation, offering a powerful tool for mental health practitioners. In terms of education and entertainment, Wu et al. study on cross-modal correspondence between ambient color and taste perception in VR provides insights that can be leveraged to create more engaging and immersive culinary VR applications.

3.1 Impact on VR quality

Papers in this Research Topic consistently show that multisensory VR leads to a stronger sense of presence, embodiment, and emotional engagement, which enhances the overall quality of VR experiences. For example, Desnoyers-Stewart et al. research on performer-facilitated touch in immersive performances highlights how real human touch can enhance the sense of presence and emotional connection in VR. This is particularly crucial in performance arts or therapeutic settings. Additionally, the innovative subtle haptic feedback introduced by Sawahata et al. enhanced task performance and created a more cohesive and immersive VR experience. Such improvements are essential for VR applications in education, where a strong sense of presence and engagement can significantly enhance learning outcomes.

4 Conclusion

This Research Topic has brought together diverse studies that collectively advance our understanding of the multimodal applications of VR to provide multisensory interventions. The findings underscore the importance of considering the full spectrum of human senses when designing immersive VR experiences. They provide valuable insights into making VR more engaging, realistic, and impactful. Importantly, the contributions made by the selected papers offer a solid foundation for continued exploration and innovation in the field. In summary, they promise a future where VR can more effectively replicate and augment real-world experiences across various applications.

Author contributions

JS: Conceptualization, Validation, Writing–original draft, Writing–review and editing. OG: Conceptualization, Writing–original draft, Writing–review and editing. DB: Writing–original draft, Writing–review and editing. MM-G: Writing–original draft, Writing–review and editing. GP: Writing–original draft, Writing–review and editing. SS: Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. JŚ is supported by XR-PAIN Grant Agreement No. 101070533 and META-TOO Grant Agreement No. 101160266. OG's contribution was supported from the EU Horizon 2020 research and innovation programme under grant agreement No 101017746, Touchless.

Conflict of interest

Author OG was employed by Ultraleap Ltd.

The remaining 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.

frontiers in Virtual Reality

ORIGINAL RESEARCH
published: 23 September 2021
doi: 10.3389/frvir.2021.730334

Lean to Fly: Leaning-Based Embodied Flying can Improve Performance and User Experience in 3D Navigation

*Ashu Adhikari[1], Abraham M. Hashemian[1], Thinh Nguyen-Vo[1,2], Ernst Kruijff[1,3], Markus von der Heyde[1,4] and Bernhard E. Riecke[1]**

[1]iSpace Lab, School of Interactive Arts and Technology, Burnaby, BC, Canada, [2]Microsoft Corporation, Seattle, WA, United States, [3]Institute of Visual Computing, Bonn-Rhein-Sieg University of Applied Sciences, Sankt Augustin, Germany, [4]VdH-IT, Weima, Germany

OPEN ACCESS

Edited by:
Doug A. Bowman, Virginia Tech, United States

Reviewed by:
Jon Kelly, Iowa State University, United States
Marco Fyfe Pietro Gillies, Goldsmiths University of London, United Kingdom

Correspondence:
Bernhard E. Riecke
ber1@sfu.ca

Specialty section:
This article was submitted to Virtual Reality and Human Behaviour, a section of the journal Frontiers in Virtual Reality

Received: *24 June 2021*
Accepted: *31 August 2021*
Published: *23 September 2021*

Citation:
Adhikari A, Hashemian AM, Nguyen-Vo T, Kruijff E, Heyde Mvd and Riecke BE (2021) Lean to Fly: Leaning-Based Embodied Flying can Improve Performance and User Experience in 3D Navigation. Front. Virtual Real. 2:730334. doi: 10.3389/frvir.2021.730334

When users in virtual reality cannot physically walk and self-motions are instead only visually simulated, spatial updating is often impaired. In this paper, we report on a study that investigated if HeadJoystick, an embodied leaning-based flying interface, could improve performance in a 3D navigational search task that relies on maintaining situational awareness and spatial updating in VR. We compared it to Gamepad, a standard flying interface. For both interfaces, participants were seated on a swivel chair and controlled simulated rotations by physically rotating. They either leaned (forward/backward, right/left, up/down) or used the Gamepad thumbsticks for simulated translation. In a gamified 3D navigational search task, participants had to find eight balls within 5 min. Those balls were hidden amongst 16 randomly positioned boxes in a dark environment devoid of any landmarks. Compared to the Gamepad, participants collected more balls using the HeadJoystick. It also minimized the distance travelled, motion sickness, and mental task demand. Moreover, the HeadJoystick was rated better in terms of ease of use, controllability, learnability, overall usability, and self-motion perception. However, participants rated HeadJoystick could be more physically fatiguing after a long use. Overall, participants felt more engaged with HeadJoystick, enjoyed it more, and preferred it. Together, this provides evidence that leaning-based interfaces like HeadJoystick can provide an affordable and effective alternative for flying in VR and potentially telepresence drones.

Keywords: locomotion interface, spatial orientation, navigational search, 3D navigation, leaning-based interfaces, virtual reality, spatial updating

1 INTRODUCTION

Spatial updating is a largely automated mental process of establishing and maintaining the spatial relationship between ourselves and our immediate surroundings as we move around Wang (2016). That is, as we move around through our environment and self-to-object relationships constantly change in non-trivial ways, our mind helps us to remain oriented by automatically updating our spatial knowledge of where we are with respects to relevant nearby objects in our surroundings. This ability allows us to navigate and interact with our immediate environment almost effortlessly Wang and Spelke (2002), Loomis and Philbeck (2008), McNamara et al. (2008). Spatial updating can also support complex activities like driving, climbing, diving, flying, or playing sports.

While spatial updating is mostly automatic or even obligatory (i.e., hard to suppress) during natural walking, it cannot be deliberately triggered when merely imagining self-motions Rieser (1989), Presson and Montello (1994), Farrell and Robertson (1998), Wang (2004). Similarly, spatial updating is impaired if self-motions are only visually simulated in virtual reality (VR) and people are not physically walking, especially when reliable landmarks are missing Klatzky et al. (1998). This has been demonstrated by comparing physical walking with a head-mounted display (HMD) to hand-held controller operated locomotion in VR Klatzky et al. (1998), Ruddle and Lessels (2006), Riecke et al. (2010). Moreover, a large percentage of participants completely fail to update rotations that are not physically performed but only visually simulated in VR Klatzky et al. (1998), Riecke (2008). This illustrates how critical it is to support reliable and automatic spatial updating in VR through, e.g., more embodied interaction and locomotion methods that can tap into such automatized and low-cognitive-load mechanisms.

While physical walking in VR is often considered the "gold standard" and can reliably elicit automatic spatial updating with low cognitive load, it is often not feasible due to restrictions on the available free-space walking area and/or safety concerns Steinicke et al. (2013). Moreover, walking does not allow for full 3D (flying or diving) locomotion, where there is currently no comparable or "gold standard" locomotion interface. To address this gap, we investigate in this study if an embodied leaning-based flying interface can help improve users' navigation performance in a 3D navigational search tasks that requires spatial updating, as well as improve other usability, performance, and user experience aspects in comparison to a commonly used dual-thumbstick flying interface. Recent research indicates that more embodied interfaces such as leaning-based interfaces can indeed improve task performance in a ground-based navigational search task, and almost reach the performance levels of physical walking Nguyen-Vo et al. (2019). However, it remains an open research question if such benefits of leaning-based interfaces would generalize to full 3D locomotion (flying) where an additional degree of freedom (DoF) needs to be controlled. If such embodied and affordable 3D locomotion interfaces could indeed improve navigation tasks relying on spatial updating, this could have substantial benefits for a variety of scenarios and use cases in both VR and immersive telepresence (UAV/drone) flying, as spatial updating is essential for reducing cognitive load during locomotion and thus leaving more cognitive resources for other tasks. These scenarios include training, disaster or emergency response management, embodied virtual tourism, or flying untethered (as no hand controllers are needed). With the advancement in both VR and drone technologies, affordable use cases across educational, commercial, health, and recreational sectors lie ahead.

To tackle this challenge, we conducted a user study to compare HeadJoystick, an embodied leaning-based flying interface adapted from Hashemian et al. (2020) (discussed in detail in **Section 2.1**), with Gamepad, a standard controller-based interface. We compared these interfaces in a novel 3D (flying) navigational search task. This task is a 3D generalization of a standard paradigm used to assess spatial updating and situational awareness (and other supporting aspects such as the user's ability to maneuver and forage effectively, and memorize which targets they have already visited vs. not etc.) in ground-based VR or real-world navigation Ruddle (2005), Ruddle and Lessels (2006), Ruddle and Lessels (2009), Riecke et al. (2010), Ruddle et al. (2011b), Fiore et al. (2013), Ruddle (2013), Nguyen-Vo et al. (2018), Nguyen-Vo et al. (2019). Further, we investigated if using HeadJoystick could help to reduce motion sickness and task load. Finally, we triangulated our finding through a post-experiment questionnaire and open-ended interviews.

2 RELATED WORKS

2.1 Locomotion in VR

In VR, hand-held controllers cannot provide physical self-motion cues that would normally accompany real-world locomotion. Since these non-visual cues, such as vestibular and proprioceptive cues, are missing, they cannot support the visual self-motion cues provided by the HMD, making it challenging to provide an embodied and compelling sensation of self-motion (vection) for the user Riecke and Feuereissen (2012), Lawson and Riecke (2014). This lack of non-visual and embodied self-motion cues has also been shown to impair performance in navigational search tasks requiring spatial updating Ruddle and Lessels (2009), Riecke et al. (2010) and spatial tasks such as directional estimates Chance et al. (1998), Klatzky et al. (1998), homing Kearns et al. (2002), pointing Waller et al. (2004), Ruddle et al. (2011b), Ruddle (2013) and estimation of distance traveled Sun et al. (2004). Moreover, missing body-based sensory information has also been shown to increase cognitive load Marsh et al. (2013), Nguyen-Vo et al. (2019) and motion sickness Aykent et al. (2014), Lawson (2014).

To provide at least some of these essential body-based cues, a variety of systems have been proposed and investigated, including large omnidirectional treadmills for ground-based locomotion and full-scale VR flight simulators Groen and Bles (2004), Ruddle et al. (2011a), Krupke et al. (2016), Perusquía-Hernández et al. (2017). Although these simulators provide a more believable experience of walking/flying using vestibular/proprioceptive sensory cues, the cost and maintenance needs of the equipment, complicated setups, required extensive safety measures, and weight and space requirements of some designs make them unfeasible for general VR home users.

VR researchers have designed leaning-based locomotion interfaces as a low-cost alternative that provide embodied control system and partial body-based sensory information. In these interfaces, leaning or stepping away from the center towards the desired direction instantiates a virtual motion in that direction. User studies have shown promising results for ground locomotion, such as improvement in spatial perception and orientation Harris et al. (2014), Kruijff et al. (2015), Nguyen-Vo et al. (2019), the sensation of self-motion, i.e., vection Kruijff et al. (2016), Riecke (2006), immersion Marchal et al. (2011), presence Kitson et al. (2015a), engagement Kitson et al. (2015a), Kruijff et al. (2016), Harris et al. (2014) and reduced cognitive load Marsh et al. (2013). Leaning-based interfaces (or a variation of stepping away from the center instead of just leaning) have also been adapted to 3D (flying) locomotion with similar effects. As we are mainly concerned with 3D locomotion in this experiment, we discuss these studies below in detail.

2.1.1 Flying in Real/Virtual Environments With 2DOF Leaning-Based Interfaces

Below we discuss several relevant leaning-based flying interfaces that allow users to control two DoF. While this is not sufficient for full control of 3D flight (which requires four DoF), they provide useful insights and inspiration.

Schulte et al. (2016) presented upper-body leaning-based flying interfaces using either the Kinect or Wii Balance Board. Both interfaces rely on the (novel) metaphor of riding a dragon. Leaning in the sagittal and coronal planes controls the dragon's pitch and combined yaw and roll, respectively. Though it travels at a constant speed, a hand gesture with Kinect temporarily accelerates the speed as well.

Miehlbradt et al. (2018) suggested a similar upper-body leaning-based interface. Using Kinect, the user's torso motion is used to perform five distinct behaviors (constant forward motion, right-banked turn (roll), left-banked turn (roll), upward pitch, and downward pitch). Users' performance (accuracy) with the leaning-based interface was better than a joystick and comparable to Birdly, a commercial mechanical interface for flying like a bird in VR Rheiner (2014).

Rognon et al. (2018) designed an upper-body soft exoskeleton, FlyJacket, that controls a fixed-wing drone flying at a constant speed. Participants use an HMD to view the real-world unmanned aerial vehicle (UAV) perspective, and control the pitch and roll through their torso leaning using an inertial measurement unit (IMU). Though FlyJacket showed no significant performance improvement compared to a standard two-thumbstick remote controller (RC), participants found it to be more natural, more intuitive, and less uncomfortable.

2.1.2 Flying in Real/Virtual Environments With Four DoF Leaning-Based Interfaces

Higuchi and Rekimoto (2013) developed a system, flying-head, that synchronizes a human head with UAV motions. Users see the UAV's camera feed through an HMD, and control the UAV's horizontal movement by walking around, elevation by crouching, and orientation by physically rotating. In a user study, their interface was found to be better than a joystick in completion time, accuracy, ease of control, ease of use and enjoyment. However, the interface implements a position control paradigm (1:1 mapping of the user's head and UAV position). This makes long distance navigation of the UAV impractical as its movement is limited to the user's head movement in the real world.

Cirio et al. (2009)'s and Marsh et al. (2013)'s design solves that problem by a concept of hybrid position/rate control. In this design, there is a 1:1 zone (positional mapping) where user can move freely and perform everyday action like bending and ducking. If the user gets out of the zone then the interface applies velocity in the direction they crossed the threshold. Marsh et al.'s study showed that using this interface in place of controllers lessens the cognitive load. In Cirio et al.'s study, participants completed the task faster compared to "freezing at the boundary and going back to the center" or redirected walking. However, this interface was only useful when users need to travel from one point in space to another in a relatively straight line and they struggled when navigation included constant twists and turns. This is also seen in Circo et al.'s study where the deviation from the ideal path is maximum for this interface.

Pittman et al. (2014) proposed "Head-Translation" and "Head-Rotation" with a very small 1:1 zone. In Head-Translation and Head-Rotation, a user controls the UAV's velocity (both magnitude and direction) in the horizontal plane by physically moving in the desired direction or tilting their head in the desired direction, respectively. In both interfaces, standing on tiptoe or squatting changes the elevation of the UAV, while returning the body to its original position halts the UAV. As they use velocity (i.e., rate) control, their interface supports sharp turns. However, it is not possible to perform actions like ducking and rolling. Among the six flying interfaces compared, Wiimote, a hand-held controller, showed the shortest completion time for passing through the waypoints. It also yielded better ratings for predictability, ease of use, and comfort. However, participants used a monitor instead of HMD with Wiimote. Excluding the Wiimote condition, the participants preferred Head-Rotation the most. However, participants had a hard time locating their original heading with Head-Translation and drifted away from their starting position.

Xia et al. (2019) also developed a VR telepresence UAV system with velocity control instead of position control. Similar to "Head-Translation," a user controls the UAV's velocity in the horizontal plane by moving in the desired direction. Their interface updates the reference point to mitigate the problem of drifting users. Whenever the user gets far away from the reference point, stepping in the opposite direction of the UAV flight automatically updates the stepped back position as the new reference point. The user no longer needs to keep track of the origin. However, during prolonged use, the reference point can keep moving away from the center of the tracked space and eventually out of the tracking space.

Hashemian et al. (2020) developed a seated leaning-based interface, "HeadJoystick," with a virtual quadcopter model. In their model, a user freely rotates the swivel chair to control the simulated rotation. The user leans in the direction they want to navigate. Further, they attach a tracker to the back of the chair to account for any difference between the head's resting position and the chair's center. The implementation, based on the tracker's orientation, updates the reference point as the chair rotates. This allows the user to rotate freely without worrying about the initial reference point. Assessing the interface in a maneuvering task, Hashemian et al. concluded that HeadJoystick improved both user experience and performance. They found the leaning-based interface to perform better than hand-held controllers in terms of accuracy, precision, ease of use, ease of learning, usability, long term use, presence, immersion, a sensation of self-motion, workload, and enjoyment.

To summarize, the studies mentioned above show that leaning-based interfaces can be a low-cost and relatively easy alternative for providing embodied control in VR. Hashemian et al.'s iteration of a flying leaning-based interface addresses the shortcomings of previous designs (except for actions like ducking and rolling; actions not required in this study's task). However, Hashemian et al. only used a fast maneuvering (waypoint travel) task Hashemian et al. (2020), and there seems to be no prior research investigating human spatial updating ability and situational awareness using embodied flying interfaces. Interfaces designed for maneuvering should support high precision of motion without compromising speed, while interfaces made for exploration and search should support spatial knowledge acquisition and knowledge gathering by freeing cognitive resources Bowman et al. (2004). So, both kind of travels are important but require the interfaces to support different kinds of motion. This motivated us to design and conduct this study which will shed light on whether HeadJoystick is suitable for only maneuvering tasks, or it can support navigational search and the underlying automatic spatial updating processes as well.

2.2 Navigational Search

Navigational search is one of the established tasks for investigating that rely on spatial updating and situational awareness Ruddle and Jones (2001), Lessels and Ruddle (2005), Ruddle (2005), Ruddle and Lessels (2006), Ruddle and Lessels (2009), Riecke et al. (2010), Ruddle et al. (2011b), Fiore et al. (2013), Nguyen-Vo et al. (2018)|, Nguyen-Vo et al. (2019). It is a complex spatial task with high ecological validity as it is equivalent to a person walking around a cluttered room looking for target objects Ruddle and Jones (2001). Ruddle and Lessels introduced a variant of a navigational search task in a series of experiments studying spatial updating as a key mechanism underlying task performance Lessels and Ruddle (2005), Ruddle and Lessels (2006), Ruddle and Lessels (2009). In their version, participants were located in a virtual rectangular room with a regular arrangement of 33 pedestals. Sixteen of those pedestals had a box on top of them, and half of those boxes contained a hidden object inside. The objective was to collect all of those eight hidden objects while minimizing revisits to previously visited boxes. To do this, participants need to be able to maneuver and forage effectively, and remember which targets they have already visited or not. Especially when there are no reliable landmarks that support re-orientation (as in our study), navigational search tasks critically rely on participants' ongoing situational awareness and more specifically their ability to spatially update where the various boxes are with respect to themselves as they constantly move around the area: once they get lost or loose track of where the already-visited-boxes are with respect to their own position and orientation (i.e., failure to spatially update), their performance will drop noticeably. That is, without a locomotion method affording automatic spatial updating, such landmark-free navigational search tasks cannot be performed effectively, making spatial updating a necessary (but not sufficient) condition for effective task performance (see chapter IV of Riecke, 2003).

Lessels and Ruddle (2005) showed that the task was trivial to perform when walking in the real world. Even when the field of view (FOV) was restricted to ($20 \times 16°$) and thus much smaller than the FOV of current HMDs, performance for real world walking was not significantly reduced. However, when the task was performed in VR, performance was significantly reduced whenever visual cues provided via HMD were not accompanied by real walking, both in a real-rotation and visual-only condition Ruddle and Lessels (2006), Ruddle and Lessels (2009).

Later, Riecke et al. (2010) pointed out that in Ruddle and Lessels' setup, navigators could use the room's geometry, a rectangular arrangement of the pedestals and the regular orientation of the objects to maintain global orientation. To avoid these confounds, they removed the surrounding room, removed the pedestals without the boxes, refrained from using a landmark-rich environment, and randomly positioned and oriented all objects for each trial in their experiment. With this modified experimental design, participants performed substantially better when they were allowed to physically rotate compared to visual-only simulation. Physical walking provided additional (but smaller) performance benefits.

Nguyen-Vo et al. (2018) showed that if a participant can walk out of the array of boxes and look at the whole scene, they could memorize the overall layout of the boxes and plan their trajectory. Studies also suggest that even a single viewing of the layout can help a user retain spatial orientation knowledge, including relative distances, directions, and scale Shelton and McNamara (1997), Zhang et al. (2011). This implied that participants just needed to memorize a pre-planned trajectory instead of needing to gradually build up their spatial knowledge as they navigated, especially if they could see the layout from an advantageous position in a fully lit room. To ensure continuous spatial updating in participants, Nguyen-Vo et al. (2019) later experimented in a dark virtual environment with a virtual headlamp attached to the avatar's head. The virtual lamp illuminated only half of the play area and prevented participants from ever seeing the overall layout and all boxes at once. Using this experimental design, Nguyen-Vo et al. compared four levels of translational cues and controls (none, upper-body leaning while sitting, whole-body leaning while standing/stepping, physical walking) accompanied by full rotational cues in all conditions. Their findings show that even providing partial body-based translational cues can help to bring performance to the level of real walking, whereas just using the hand-held controller significantly reduced both performance and usability.

In summary, the navigational search task has gone through numerous iterations with each iteration addressing previously found confounds. Successful and effective completion of the task critically relies on spatial updating as discussed above. Here, we build on and expand on this task, by for the first time including vertical locomotion in a navigational search task to study full 3D locomotion, similar to what drones and many computer games provide.

3 MOTIVATION AND GOAL

To close this gap in the literature, we are in this paper mainly concerned with investigating if leaning-based flying interfaces like HeadJoystick can not only improve maneuvering ability Hashemian et al. (2020) compared to the standard 2-thumbstick flying interfaces, but also improve navigation in a task reliant on automatic spatial updating, which is critical for effective and low-cognitive-load navigation and situational awareness Rieser (1989), Presson and Montello (1994), Klatzky et al. (1998), Farrell and Robertson (1998), Riecke et al. (2007). Further, we want to ground the applicability of the interface by studying its impact on motion sickness and task load, as well as diverse aspects of user experience and usability. Reduced cognitive/task load can be an indicator of improved spatial updating, as automatic spatial updating by it's very definition will automatize part of the spatial orientation challenges and thus reduce task load (TLX mental load, effort and frustration). Similarly, if an interface induces high motion sickness (disorientation, dizziness) then it can adversely impact spatial updating.

3.1 RQ1: Does HeadJoystick Improve Navigational Search Performance Compared to a Hand-Held Controller for 3D Locomotion?

In the HeadJoystick interface, the upper body is leaning in the direction of the simulated motion. This partial body movement of leaning provides minimal vestibular cues that are at least somewhat consistent with the virtual translation. This partial consistency is both spatial (i.e., leaning movements provide some proprioceptive and vestibular cues that are aligned with the acceleration direction in VR) and temporal, in that leaning directly controls the simulated self-motion without additional delay, thus providing almost immediate feedback. These (although limited) embodied translational cues have been shown to help reduce disorientation in 2D navigational search Nguyen-Vo et al. (2019). Hence, we hypothesize that they should yield improved task performance (by more effectively supporting spatial updating) for the HeadJoystick in 3D navigation as well (H1).

3.2 RQ2: Does HeadJoystick Help to Reduce Motion Sickness in 3D Navigation?

When physically stationary individuals view compelling visual representations of self-motion without any matching non-visual cues, it can cause unease and induce motion sickness Hettinger et al. (1990), Cheung et al. (1991), Riecke et al. (2015). This can cause illness in the user or even incapacitate them, limiting the utility of VR Hettinger and Riccio (1992). Further, sensory conflict is most prominent during the change in velocity (acceleration/deceleration) Bonato et al. (2008), Keshavarz et al. (2015).

Previous studies for ground-based locomotion have also shown that if virtual locomotion is accompanied by matching body-based sensory information similar to real-world locomotion, it can help to reduce motion sickness Aykent et al. (2014), Lawson (2014). However, the literature indicates mixed results for partial body-based sensory information. Some ground-based locomotion studies reported no significant difference in motion sickness between leaning-based interfaces and hand-held controllers Marchal et al. (2011), Hashemian and Riecke (2017a), while others reported significant reductions of motion sickness with a leaning-based interface Nguyen-Vo et al. (2019). Further, as far as the authors know, the literature does not provide a definitive answer on whether the benefits of such implementation translate to flying locomotion. Rognon et al. hypothesized increased motion sickness with the remote controller to explain their results, but do not have explicit measurements. In Pittman and LaViola's study, participants using Wiimote reported significantly less motion sickness than those using head-rotation and head-translation, but Wiimote did not use an HMD. Further, only five out of 18 participants reported more than 10% of total SSQ score after the experiment. In Hashemian et al. (2020)'s study, participants found a significant difference (again, change of <10% of total SSQ score) in motion sickness between real-rotation with leaning-based translation and controller-based translation and rotation conditions; however, implementing real-rotation with both leaning-based and controller-based translation did not produce a change that was statistically significant.

Despite these conflicting findings, HeadJoystick is designed for providing at leat some vestibular and proprioceptive cues that aid the visual self-motion perception provided by the HMD Hashemian et al. (2020). Handheld controllers like a Gamepad cannot provide these physical self-motion cues. To change the velocity in HeadJoystick the user has to physically lean in the direction of acceleration, thus providing at least some vestibular self-motion cues in the correct direction and thus reducing the visual-vestibular cue conflict. Thus, we hypothesize HeadJoystick should reduce motion sickness and thus potentially allow for more extended headset usage (H2). We performed a planned contrast to see if there is any trend in change in motion sickness even if both interfaces should produce minimal motion sickness.

3.3 RQ3: How do HeadJoystick and Gamepad Affect the Overall User Experience and Usability in 3D Navigation?

For the HeadJoystick interface, the simulated motion is consistent with the direction of the upper body. So, any point in the space can be reached by freely leaning towards that direction. However, each thumbstick on a Gamepad is constrained to control only two DoF. So, even with real-rotation, to control three degrees of translation simultaneously the user needs to proportionately combine inputs from those two thumbsticks to travel in the desired 3D direction. Alternatively, users could only use one thumbstick at a time and alternate their input to the thumbsticks, and keep switching their plane of movement until they reach the target. Hashemian et al. (2020)'s study also showed that participants found HeadJoystick to be easier to learn and use than Gamepad. Thus, we hypothesize that HeadJoystick should be more intuitive to use and learn, and users should be able to use HeadJoystick more effectively, even without any previous exposure (H3).

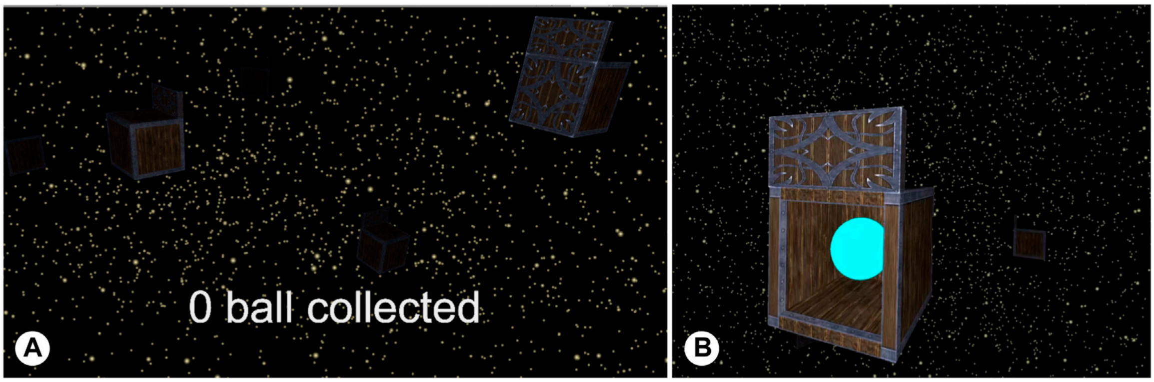


FIGURE 1 | (A) A snapshot of a participant's view of the play area (6 m × 3 m cylinder) with the closed boxes. The light attached to the avatar's head lit half of the play area, 4.24 m. Boxes became dimly lit as their distance increased from the player and stopped being visible when they were further than 4.24 m. **(B)** The target, a blue ball, seen after approaching the box.

In addition to the three specific aspects mentioned above, we are also interested in more generally exploring how the two interfaces affect user experience and usability. In Hashemian et al. (2020)'s study, participants rated HeadJoystick as providing overall better user experience and usability than Gamepad. We hypothesize that these findings can be replicated even in a different environment with a different task (navigational search based on spatial updating, instead of maneuvering).

4 METHODS

4.1 Participants

From 25 users whom participated in our experiment, we excluded three participants and performed analysis with remaining 22 participants (10 female), 19–32 years old (M = 24.0, SD = 3.70). 15 of them casually/regularly played video-games on a computer or a gaming console; 13 of them had used 3D navigation with video games, 3D modeling or flight simulator; and 15 of them had used a HMD before. Among the three excluded participants, one experienced motion sickness during the study and dropped the experiment. The other two excluded participants showed unusually high SSQ scores. Although we did not observe anything unusual during the experiment, we realized they reported high SSQ scores even before the start of the study. Due to this discrepancy and unreliability, we excluded their data as well. The studies had approval of the SFU Research Ethics Board (#2018s0649).

4.2 Virtual Environment and Task

As our main goal was to investigate how well different interfaces support participants' spatial updating and situational awareness, we carefully designed the virtual environment (**Figure 1**) and task to avoid potential confounds reported in prior works (see **Section 2.2**). Specifically, as our main focus was to investigate and compare locomotion interfaces and spatial orientation/updating performance, we carefully avoided all landmarks or global orientation cues. The experimental task was, apart from the changes described below, similar to the ground-based navigational search task in Nguyen-Vo et al. (2019)'s experiment, but generalized for 3D locomotion (flying) [1]. To generalize for 3D, we tested different shapes and sizes of the play area prior to the main experiment. If the play area was too small, a user could quickly collect all eight balls without revisiting the boxes. If the virtual area was too big, they got easily lost in the vast void space with no global orientation cues (landmarks). We also observed that participants would keep going out of the play area when it was spherical. As a result of iterative pilot testing, we determined a cylindrical virtual area 6 m in diameter and 3 m in height would be a good fit for our experiment.

During each trial, sixteen boxes were randomly placed within this area, with eight of the boxes containing a blue ball (see **Figure 1**). The remaining eight boxes served as decoys. The participants' objective was to efficiently collect as many balls as possible. In our navigational search task, participants started each trial from the center of the cylinder. A trial ended when participants found all eight balls or the trial ran for 5 min. We chose to limit the trial length to reduce motion sickness.

Participants were explicitly told that the criteria for efficiency were the number of balls collected, the total distance traveled, and the number of revisits. Since it was possible to complete the game by collecting all the balls before 5 min, we also recorded the completion time.

4.3 Interaction

To check if there was a ball inside a box, participants needed to approach it from its front side, indicated by an additional banner (**Figure 1B**). The box automatically opened when participants' viewpoints were close (within 90 cm from the box's center) and facing the front side (within ±45° from the box's forward vector). To prevent the accidental collection of the balls, the user needed to keep the box open for one second. The user was alerted through

[1]https:www.youtube.com/watch?v=xzTR_8sfZXA

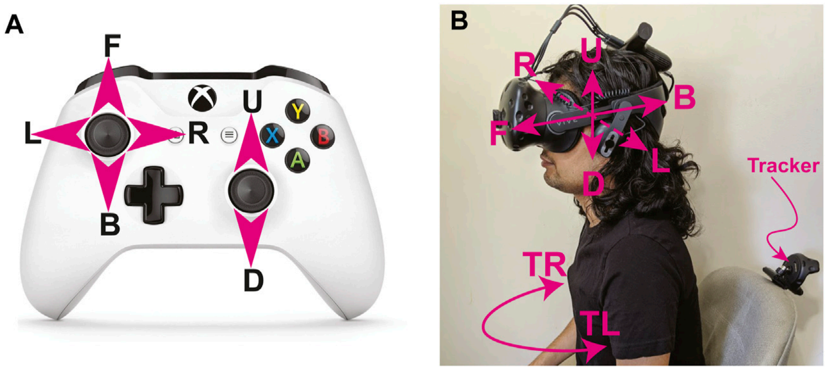


FIGURE 2 | (A) In Gamepad, the left thumbstick controls the horizontal velocity (**F**orward, **B**ackward, **L**eft, **R**ight) and the right thumbstick controls the vertical velocity (**U**p, **D**own) **(B)** In HeadJoystick, the user's head position controls the velocity. In both cases, physical rotation (**T**urning **R**ight, **T**urning **L**eft) is applied.

a ticking sound as the box opened, and it was followed by a "ding" sound for collection.

As colliding with the boxes and any subsequent physics simulation would disorient the user and even induce motion sickness, we switched off collision detection for the boxes and the user could pass through them. However, to prevent participants from peeking into the boxes from the other sides, the ball became visible only when a box was approached from the front side.

4.4 Locomotion Modes

Hand-based controllers are still the most common interfaces for navigation in VR, especially when physical walking is not feasible. To choose among the hand-based controllers for the experiment, we compared the Vive controller that came with the headset and a gamepad. Through our pilot testing, we learned that the participants use their thumbs for controlling both kinds of controllers and release their thumbs to come to a halt. As a trackpad has no physical feedback that indicates the center, participants had difficulty providing proper input once they released their thumb, and as a result took time to adjust their input. A gamepad has thumbsticks loaded with springs that force the thumbsticks to come back to their center when released. Because of this, the user can locate the center much more quickly. Further, a gamepad's thumbsticks are similar in design to the most common remote controller for drones. So, we chose to compare the HeadJoystick interface with a gamepad. Further, it adds comparability with Hashemian et al. (2020)'s original paper which proposed the HeadJoystick interface.

For both the Gamepad and HeadJoystick conditions, participants rotated the swivel chair they were sitting on to control the simulated rotations in VR. However, they translated in different manners. We chose to include only the interfaces that allow physical rotation because the importance of rotation in spatial updating has already been proved multiple times Klatzky et al. (1998), Riecke et al. (2010). Further, implementing physical rotations is no longer an issue, as HMDs are becoming increasingly wireless, and therefore have no cables to be entangled.

4.4.1 Gamepad Interface

For the Gamepad interface, the left control stick controlled horizontal translation velocities as illustrated in **Figure 2**. The right control stick controlled upward/downward translation speeds. Although physical rotation controlled yaw, for simplicity, we will refer to this interface as the Gamepad throughout the paper.

4.4.2 HeadJoystick Interface

In the HeadJoystick interface, head position determined the translation. The interface calibrated the zero-point before each use. Moving the user's head in any particular direction from that zero-point made the player move in the same direction. The distance of the head from the zero-point determined the speed of the virtual motion. To stop the motion, the user had to bring their head back to the zero-point. As a subsequent result, leaning forward and backward caused the user to move forward and backward, leaning left or right caused sideways motions, stretching their body up or slouching down created upward or downward motions, and coming back to the center stopped the motion. In this kind of interface, many prior implementations include a small 1:1 zone (also known as idle, dead, or neutral zone) where the physical head motion correspond to 1:1 mapped virtual motion, to allow users to more easily break or be stationary. However, during the pilot testing for this study, we observed that our exponential curve (relating head deflection to virtual translation speed, see **Section 4.5**) is fairly flat around the zero point and was sufficient for the user to easily slow down and collect the balls. Hence, we opted to not include a 1:1 zone. This also helped to reduce the amount users had to lean to travel with a faster speeds.

4.5 Motion Control Model

The velocity calculation is based on a scaled exponential function, similar to the function for a smooth translation proposed by LaViola et al. (2001)'s study on leaning based locomotion.

$$F = \alpha e^{-\beta|\vec{head}} \cdot \vec{V}_{up}| \quad (1)$$

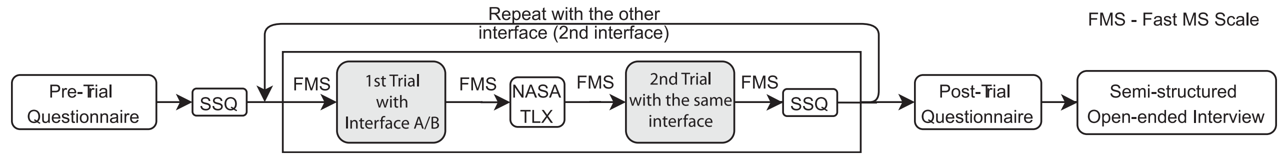


FIGURE 3 | Experimental procedure.

where α is the maximum speed factor, β controls the steepness of the exponential curve, $\overrightarrow{head}$ is a vector indicating the user's head orientation, and $\overrightarrow{V}_{up}$ is the vertical vector. Exponential implementation creates a smooth transition. It has been successfully implemented in other 2D interfaces Hashemian and Riecke (2017a), Nguyen-Vo et al. (2019). The same method also provides smooth translation in 3D when the projection of head orientation onto the plane $|\overrightarrow{head} \cdot \overrightarrow{V}_{up}|$ is replaced just with $|\overrightarrow{head}|$. It has successfully been implemented in Hashemian et al.'s study. Please consult the appendix of Hashemian et al. (2020) for a complete description of the HeadJoystick and its underlying mathematical model.

4.6 Experimental Design and Procedure

In this experiment, we compared the performance of the HeadJoystick and Gamepad interface. We collected the users' behavioral data while they were performing the tasks. After completing the trials, we asked them to fill out a questionnaire and performed a semi-structured open-ended interview. We deployed a 2-blocked, repeated measure experimental design. All participants performed the navigational search task twice for each interface, totaling four trials and thus up to 20 min of VR exposure in total. The order of the interfaces was counter-balanced to account for the order effects and maturation effects.

The overall procedure is illustrated in **Figure 3**. After reading and signing the informed consent form, participants filled out a pre-experiment questionnaire before starting the experiment, asking about their age, gender and previous experience with video games and HMDs. Then, they were guided through the tasks and tried out both interfaces before the actual experiment started. Before they began the experiment, they filled out the Simulator Sickness Questionnaire (SSQ) Kennedy et al. (1993). They started with one of the two interfaces. They performed two trials with the first interface. Participants were asked to take off the Vive Headset after the first trial and fill out NASA's Task Load Index (TLX) Hart and Staveland (1988) to reduce the potential for motion sickness. After completing the second trial, they filled out the SSQ again. The questionnaires were strategically placed between each trial to provide a short break between the trials. They repeated the same procedure including two trials with the interface they had not used in the first two trials.

Further, to assess motion sickness issues, we asked them to estimate their current state of motion sickness before and after each trial. They rated their motion sickness on a scale of 0–100. A rating of "0" meant "I am completely fine and have no motion sickness symptoms" and "100" meant "I am feeling very sick and about to throw up." Based on their scale, we recommended that they go ahead with the trial, take a longer break or drop the experiment. This scale was adapted from the Fast Motion Sickness Scale (FMS), which goes from 0–20 Keshavarz and Hecht (2011).

Before switching the interface, participants were asked to take a minimum 5 min break, including the time required to fill out the questionnaires. In addition to taking a mandatory break, participants were encouraged to take a short walk or drink water.

After completing all four trials, they completed a post-study survey questionnaire (detailed in **Section 5.2.3**) and responded verbally to semi-structured open-ended interview questions. The whole study took, on average, about 1 h to complete.

5 RESULTS

5.1 Behavioral Measures

Six quantitatively measured behavioral data types are summarized and plotted in **Figure 4**. Improved spatial updating for a given interface would be expected to increase the number of balls collected, and reduce task completion time and the number of revisits needed. Similarly, if an interface better supports spatial awareness, this might yield an increase in participants' head and body rotation, and reduce travel distance to reach the same number of targets. All six measures were analyzed using 2 × 2 × 2 repeated-measures ANOVAs with the independent variables **interface** (HeadJoystick vs. Gamepad), **repetition** (first vs. Second trial) and order of the interface assignment, **group** (GamepadFirst vs. HeadJoystickFirst). Since neither repetition, group, any interaction with the group, or the interaction between interface and repetition showed any significant effects for any of the dependent variables (all $p's > 0.05$), we only report the main effects of interface below. Unless stated otherwise, all test assumptions for ANOVA were confirmed in each case, $p < 0.05$ was considered a significant effect and its double $p < 0.10$ was considered a marginally significant effect.

5.1.1 Participants Collected Significantly More Balls When Using HeadJoystick

Figure 4A. Participants collected all eight balls in 31 out of 44 trials with HeadJoystick and 26 out of 44 trials with Gamepad. All

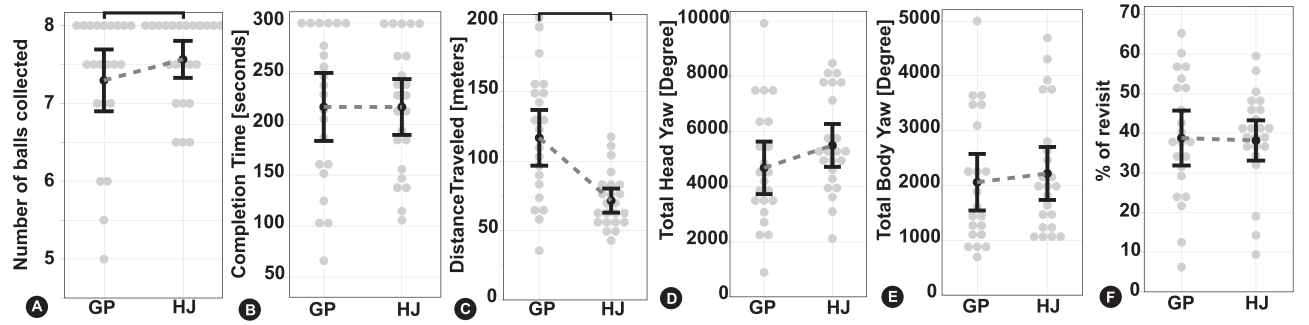


FIGURE 4 | Behavioral data plot for Gamepad (GP) and HeadJoystick (HJ) with an overall mean (black dots), individual participants' average (gray dots) and error bars at 95% Confidence Intervals (CI).

participants were able to collect at least six balls with HeadJoystick and at least four balls with Gamepad. On average, participants were able to collect more balls when using HeadJoystick (M = 7.61, SD = 0.655) than Gamepad (M = 7.30, SD = 1.01), $F(1, 42) = 4.51$, $p = .040$, $\eta_p^2 = .097$.

5.1.2 Task Completion Time did not Differ Between Interfaces

Figure 4B. Participants reached the time limit of 5 min in 13 trials when using the HeadJoystick vs. 18 trials when using the Gamepad. The fastest participant finished the task in 69 s with HeadJoystick and 48 s with Gamepad. Repeated measures ANOVA showed no significant difference in completion time between HeadJoystick (M = 214 s, SD = 76.5 s) and Gamepad (M = 217 s, SD = 84.6 s), $F(1, 42) = .119$, $p = .732$, $\eta_p^2 = .003$.

5.1.3 Participants Travelled Significantly Less While Using HeadJoystick

Figure 4C. Participants travelled from 33.5 to 143.9 m with HeadJoystick and from 30.8 to 277.4 m with Gamepad. ANOVA showed that participants overall travelled significantly less with the HeadJoystick (M = 72.3 m, SD = 28.3 m) than the Gamepad (M = 116.8 m, SD = 56.9 m), F (1, 42) = 25.4, $p < 0.001$, $\eta_p^2 = .378$.

5.1.4 HeadJoystick Marginally Increased Overall Head Rotations, but not Body Rotations

Figures 4D,E. We recorded the users' body rotation (rotation of the chair) and head rotation (rotation of HMD) because it would inform us if either of the interfaces restricted or reduced looking around and thus potentially hindered situational awareness. The accumulated body rotation for HeadJoystick (M = 2,255°, SD = 1,360°) and Gamepad (M = 2061°, SD = 1,260°) did not differ statistically, F (1, 42) = 0.782, p = 0.382, $\eta_p^2 = .018$. The accumulated head rotation, however, showed a marginally significant effect F (1, 42) = 3.88, p = 0.056, $\eta_p^2 = .085$, indicating marginally larger accumulated head rotation for HeadJoystick (M = 5520°, SD = 2410°) compared to Gamepad (M = 4670°, SD = 2490°).

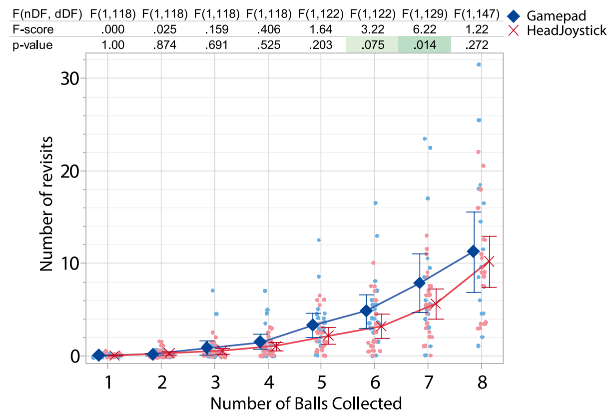


FIGURE 5 | Total target (boxes) revisited by the participants as they collected new balls. Light shades of dots represent the mean values per participant. ◆ (blue) and X (red) indicate the mean for Gamepad and HeadJoystick interface respectively, CI = 95%. Table above shows the planned contrast between the interfaces as the number of revisits accumulated with number of balls collected.

5.1.5 There was no Significant Difference Between % of Revisit (Ratio of Revisited Boxes to the Total Number of Boxes Visited)

Figure 4F. Only five participants (two with Gamepad, three with HeadJoystick) had no revisits to the target boxes before the trial completed. Some participants travelled slowly and visited only a few boxes. Others travelled quickly and visited as many boxes as possible. Since the total number of revisits depends on the total number of targets visited by the participants, we analyzed the ratio of revisited boxes to the total number of boxes visited by the participants. The mean % revisits was not significantly different, F (1, 42) = 0.018, p = 0.894, $\eta_p^2 < .001$, between the HeadJoystick (M = 38.3, SD = 18.7) and Gamepad (M = 38.8, SD = 18.8). We also analyzed how the number of revisits progressed as participants collected more balls. As seen in **Figure 5**, as participants collected more balls and travelled more within the environment, the number of revisits increased at a different rate for HeadJoystick and Gamepad. Starting from no difference for

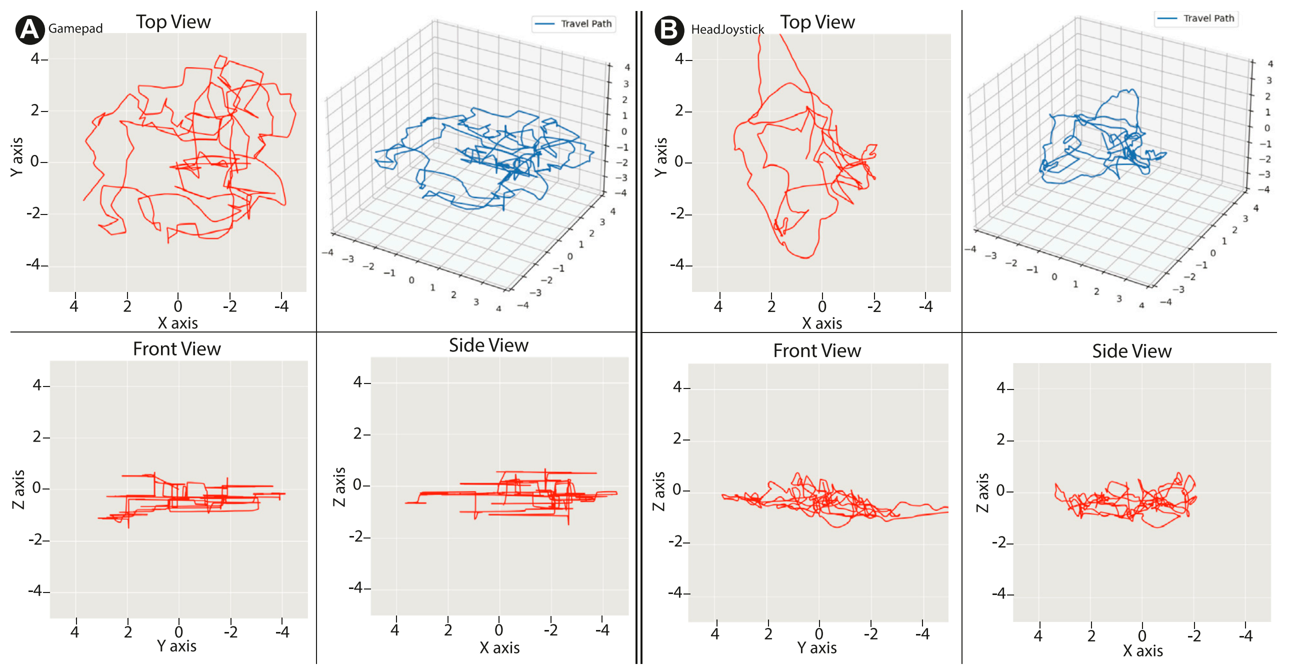


FIGURE 6 | Representative travel path (isometric, top, front and side) for **(A)** Gamepad and **(B)** HeadJoystick from a randomly selected participant.

the first collected ball, Gamepad took marginally significant more revisits ($p < 0.10$) to collected the sixth ball and significantly more ($p < 0.05$) revisits to collect the seventh ball. However, the above data uses all the revisits from all the trials. When we analyzed the data selecting only the participants who successfully collected all eight balls for both interfaces, the trend continued but the difference was less prominent.

We also recorded the travel path of the trials to investigate potential behavioral difference between the interfaces during navigation. Since putting the travel path of all the users and trials in a single graph created a dense path plot with impossible to distinguish travel instances, we show representative travel paths for Gamepad and HeadJoystick from a randomly selected participant (**Figure 6**). As the figure illustrates, with the Gamepad, participants restricted themselves to controlling no more than two translational DoF at a time, while with HeadJoystick, they controlled all available DoFs simultaneously. This is indicated by the straight horizontal and vertical lines with almost perpendicular turns with Gamepad (front and side views, **Figure 6A**) and curved paths with HeadJoystick in all projections (**Figure 6B**). Plots of almost all travel paths of individual trials showed similar trends and are submitted as a **Supplementary Material** for reference.

5.2 Subjective Ratings

5.2.1 Motion Sickness

5.2.1.1 Simulation Sickness Questionnaire

The **time** of SSQ measurement (participant's SSQ score before-0, after completing the trials with the first interface-1, and after completing the trials with the second interface-2) was one of the independent variables (within-subject factor). The order of assignment of the first interface (**group**-HeadJoystickFirst/GamepadFirst) was the second independent variable (between-subject factor). We chose to interpret the data with the time of SSQ measurement rather than the interfaces themselves because motion sickness accumulates over time. We performed a two-way mixed ANOVA using those two factors. Greenhouse-Geisser correction was applied whenever the assumption of sphericity was violated. As discussed in **Section 3**, we also compared Pre-Experiment SSQ scores to SSQ scores after using the first and second interfaces as planned contrasts with Bonferroni correction, summarized in **Table 1**. Finally, we tested the correlation between the distance travelled and motion sickness as participants travelled significantly more with Gamepad. However, linear correlation analysis showed that motion sickness did not correlate with travelled distance for either of the interfaces (all $p's > 5\%$).

Each trial produced only minimal motion sickness on average, and the average SSQ score after the experiment was 16.8% (or 39.7 in the SSQ scale from 0 to 235.32). The highest SSQ scores reached by any participant was 44.4% (104.7). We can see from **Figure 7** that for an average participant, when they used Gamepad as their first interface (blue line), SSQ total and its sub-scales increased from Pre-Experiment to after using the Gamepad, and it stayed at the same level or even decreased after switching to HeadJoystick. For an average participant using HeadJoystick as their first interface (red line), not only did SSQ total and its sub-scales increase after using the HeadJoystick, it further continued increasing after switching to Gamepad. Inferential statistical analysis done below shows the same result.

TABLE 1 | LSMeans Contrast was used to compare the Pre-Experiment SSQ score with SSQ score after Gamepad use and HeadJoystick use. Bonferroni correction was used for doubling the comparison. Significant differences ($p \leq 2.5\%$) are highlighted in green, and a lighter shade of green indicates marginally significant results ($p \leq 5\%$).

	GamepadFirst				HeadJoystickFirst			
	Pre vs Gamepad		Pre vs HJ		Pre vs Gamepad		Pre vs HJ	
	F(1,40)	p	F(1,40)	p	F(1,40)	p	F(1,40)	p
Total	5.62	.023	3.08	.087	14.2	<.001	3.11	.086
Nausea	5.91	.020	2.63	.113	8.14	.007	1.61	.212
Oculomotor	4.48	.041	4.48	.041	11.6	.002	2.18	.147
Disorientation	4.10	.050	1.30	.261	18.3	<.001	4.82	.034

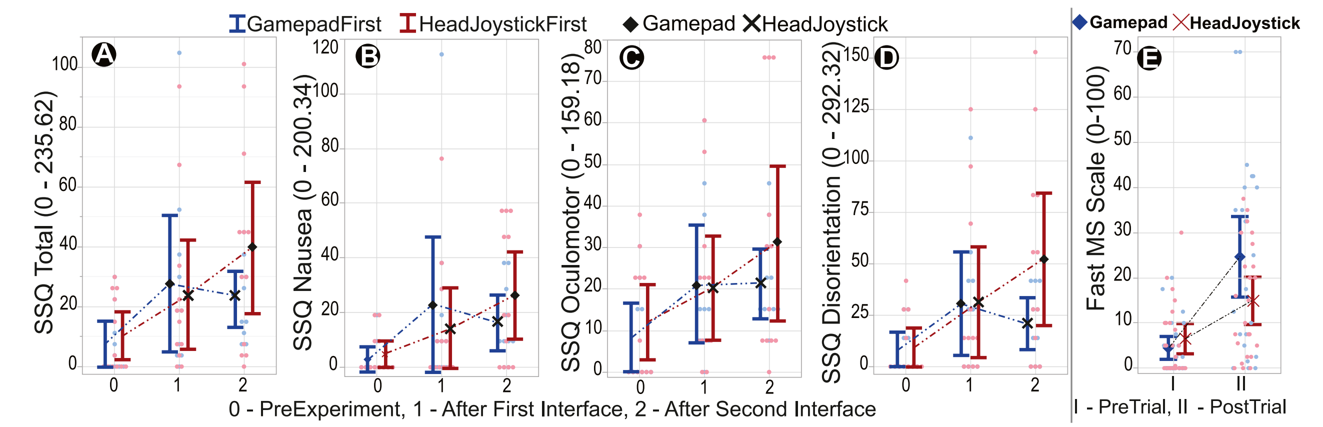


FIGURE 7 | User's self-report of motion sickness using SSQ questionnaire before the experiment started ("Pre-Experiment"), after using the first interface, and after using the second interface. Blue and red lines indicate an average participant using Gamepad and HeadJoystick as their first interface respectively. ◆ = Gamepad interface, X = HeadJoystick interface, *CI* = 95%.

5.2.1.1.1 Total SSQ Scores Increased Significantly After Using Gamepad and Increased Marginally After Using HeadJoystick

Figure 7A. ANOVA revealed a main effect of time (Pre-Experiment, after the first interface, and after the second interface), $F(1.85, 36.9) = 8.05$, $p = .001$, $\eta_p^2 = .287$. There was no effect of group (GamepadFirst/HeadJoystickFirst), $F(1, 20) = .457$, $p = .507$, $\eta_p^2 = .022$ as well as no interaction between time × group $F(1.85, 36.9) = 1.71$, $p = .197$, $\eta_p^2 = .079$. LSMeans (Least Squares Means) contrast showed that even with Bonferroni correction ($p < 0.025$ is significant) in the GamepadFirst group, the total SSQ score increased significantly from Pre-Experiment ($M = 7.48$, $SD = 10.7$) to Gamepad use ($M = 27.7$, $SD = 31.8$). However, the SSQ score dropped after switching to HeadJoystick ($M = 22.4$, $SD = 13.1$) as the second interface, and was no longer significantly higher than the Pre-Experiment score (blue line in **Figure 7A**). In contrast, for the HeadJoystickFirst group when using HeadJoystick for their first trial ($M = 24.0$, $SD = 28.7$) their motion sickness increased only non-significantly from Pre-Experiment ($M = 10.3$, $SD = 12.6$). When they switched from HeadJoystick to Gamepad, their motion sickness ($M = 39.6$, $SD = 34.5$) shot up and it was significantly higher than their Pre-Experiment scores (red line in **Figure 7A**).

5.2.1.1.2 Participants Got Significantly Nauseous After Using Gamepad While there was no Significant Change With HeadJoystick

Figure 7B. ANOVA revealed a main effect of time, $F(1.79, 35.8) = 5.67$, $p = .007$, $\eta_p^2 = .221$. However, there was no effect of group, $F(1, 20) = .031$, $p = .861$, $\eta_p^2 = .002$ as well as no interaction between time × group $F(1.79, 35.8) = 1.40$, $p = .260$, $\eta_p^2 = .065$. LSMeans contrast showed that in the GamepadFirst group, the nausea score increased significantly from Pre-Experiment ($M = 2.86$, $SD = 6.44$) to Gamepad use ($M = 22.9$, $SD = 34.6$), $F(1, 40) = 5.91$, $p = 0.020$. However, switching to HeadJoystick as the second interface reduced nausea scores ($M = 16.2$, $SD = 14.3$) and they were no longer significantly higher than the Pre-Experiment scores (blue line in **Figure 7B**). Similarly, for the HeadJoystickFirst group, nausea scores after their first trial with the HeadJoystick ($M = 14.3$, $SD = 23.2$) were not significantly elevated compared to their Pre-Experiment nausea scores ($M = 4.77$, $SD = 7.6$). When they switched from HeadJoystick to Gamepad, their nausea score increased ($M = 26.2$, $SD = 25.1$) and were significantly higher than their Pre-Experiment scores (red line in **Figure 7B**).

TABLE 2 | NASA Task Load demands for both interfaces are analyzed with ANOVA. Significant differences ($p \leq 5\%$) are highlighted in green, and a lighter shade of green indicates marginally significant results ($p \leq 10\%$).

Measures (%)	1st Interface				2nd Interface				Time (Within)			Group (First Interface/ Between)			Time x Group[1]		
	GamePad		HeadJoystick		Gamepad		HeadJoystick										
	Mean (M)	Standard Error (SE)	Mean (M)	Standard Error (SE)	Mean (M)	Standard Error (SE)	Mean (M)	Standard Error (SE)	F(1,20)	p	η_p^2	F(1,20)	p	η_p^2	F(1,20)	p	η_p^2
Total Task Load	49.8	12.7	57.3	12.2	64.0	18.1	56.6	12.1	4.50	.047	.184	2.07	.166	.094	.001	.973	<.001
Mental Demand	44.5	23.2	39.7	28.9	56.1	25.2	24.3	18.6	.072	.792	.004	3.00	.099	.130	6.70	.018	.251
Physical Demand	8.40	8.15	5.42	8.16	18.33	22.6	45.9	30.6	18.6	<.001	.482	6.38	.020	.242	4.43	.048	.181
Temporal Demand	15.8	19.7	44.5	17.1	19.4	20.6	11.5	13.1	8.48	.009	.298	9.96	.005	.332	4.24	.053	.175
Performance	31.0	16.9	20.4	10.4	18.3	14.0	20.7	12.8	2.47	.132	.110	2.28	.147	.102	1.09	.310	.051
Effort	25.6	16.0	35.1	21.8	44.4	21.5	46.4	22.2	8.96	.007	.309	.266	.612	.013	1.30	.268	.061
Frustration	24.2	24.8	27.1	26.6	35.3	26.6	21.1	21.1	.158	.695	.008	1.00	.329	.048	.766	.392	.037

[1] It is equivalent to the main effect of Interface (within) when the factors of ANOVA are Interface and Group

5.2.1.1.3 Participants had Oculomotor Issues After Using Gamepad While There were Mixed Results With HeadJoystick

Figure 7C. ANOVA revealed a main effect of time, $F(1.99, 39.8) = 7.72, p < .001, \eta_p^2 = .279$, indicating an overall increase over trials as shown in **Figure 7C**. However, there was no effect of group, $F(1, 20) = .383, p = .543, \eta_p^2 = .019$ as well as no interaction between time × group $F(1.99, 39.8) = .853, p = .434, \eta_p^2 = .041$. LSMeans contrast showed for the GamepadFirst group a marginally significant increase in oculomotor issues (eye strain, blurred vision, difficulty focusing, etc.) from Pre-Experiment (M = 8.34, SD = 11.6) to Gamepad use (M = 21.2, SD = 19.8) (note: with Bonferroni correction, significant results require $p <$ 0.05). Switching to HeadJoystick as the second interface did not change the ratings (M = 21.2, SD = 11.7) and the difference remained marginally significant from the Pre-Experiment score (blue line in **Figure 7C**). In contrast, for the HeadJoystickFirst group there was no significant increase in oculomotor issues from Pre-Experiment (M = 12.0, SD = 14.25) to their first trial (M = 20.2, SD = 23.2). When they switched from HeadJoystick to Gamepad, their ratings increased (M = 30.1, SD = 29.3) and were significantly higher than their Pre-Experiment ratings, (red line in **Figure 7C**).

5.2.1.1.4 Participants had Disorientation Issues After Using Gamepad While There were Mixed Results With HeadJoystick

Figure 7D. ANOVA revealed a main effect of time, $F(1.92, 38.3) = 7.77, p < .002, \eta_p^2 = .280$ and marginal tinteraction between time × group $F(1.92, 38.3) = .280, p = .075, \eta_p^2 = .041$. However, there was no effect of group, $F(1, 20) = .948, p = .342, \eta_p^2 = .045$. LSMeans contrast showed that for the GamepadFirst group, the increase in disorientation was marginally significant from Pre-Experiment (M = 8.35, SD = 11.7) to Gamepad use (M = 30.6, SD = 35.2), F (1, 40) = 4.10, p = 0.050. Switching to HeadJoystick as the second interface decreased disorientation (M = 20.9, SD = 17.7) to a level that was no longer significantly different from Pre-Experiment scores, F (1, 40) = 1.30, p = 0.261 (blue line in **Figure 7D**). For the HeadJoystickFirst group, disorientation scores increased marginally from Pre-Experiment (M = 9.28, SD = 14.9) to their first trial using the HeadJoystick (M = 31.3, SD = 42.4), F (1, 40) = 4.82, p = 0.034. When they switched from HeadJoystick to Gamepad, their disorientation ratings increased (M = 52.2, SD = 50.8) and were now significantly higher than their Pre-Experiment ratings, F (1, 40) = 18.3, $p <$ 0.001 (red line in **Figure 7D**).

5.2.1.2 Fast MS Scale

5.2.1.2.1 Participants Reported Higher FMS Increase After Using Gamepad than After Using HeadJoystick

Figure 7E. Participants' self-reported motion sickness score (Scale: 0–100) given before and after each trial was analyzed using 2-factor repeated measures ANOVA. The results show that self-reported motion sickness score increased overall from before to after a trial, $F(1, 21) = 30.0, p < .001, \eta_p^2 = .588$ (Before: $M = 5.43$, $SD = 7.23$| After: $M = 19.3$, $SD = 18.2$). However, as illustrated in **Figure 7E**, the interface × time interaction was also significant, $F(1, 21) = 21.1, p < .001, \eta_p^2 = .501$, indicating the degree of motion sickness increase from pre-to post-trial was larger for the Gamepad (Pre: M = 4.66, SD = 619 — Post: M = 25.0, SD = 21.7, a 436% increase) than for the HeadJoystick (Pre: M = 6.21, SD = 8.22 — Post: M = 13.6, SD = 12.1), where motion sickness only increased by 119%.

5.2.2 Task Load

The final weighted score as well as individual six sub-scores from the NASA Task Load Index (TLX) for the two factors, **time** (first vs. second interface), within-subject factor) and interface order (participant **group**: GamepadFirst or HeadJoystickFirst, between-subject factor) were analyzed with two-way mixed ANOVAs, with statistical results, means, and standard errors summarized in **Table 2**. We chose to analyze the TLX with time rather than interface as a main factor because it considers the effect of switching from HeadJoystick to Gamepad and vice versa. Further, this 2 × 2 ANOVA has factors with only two levels each. Therefore, the interaction between time and group (time × group) is equivalent to the main effect of interface in an ANOVA analysis with interface as one of the factors. To make the results' interpretation more easily understandable and comparable, we have scaled each measurement to 0–100.

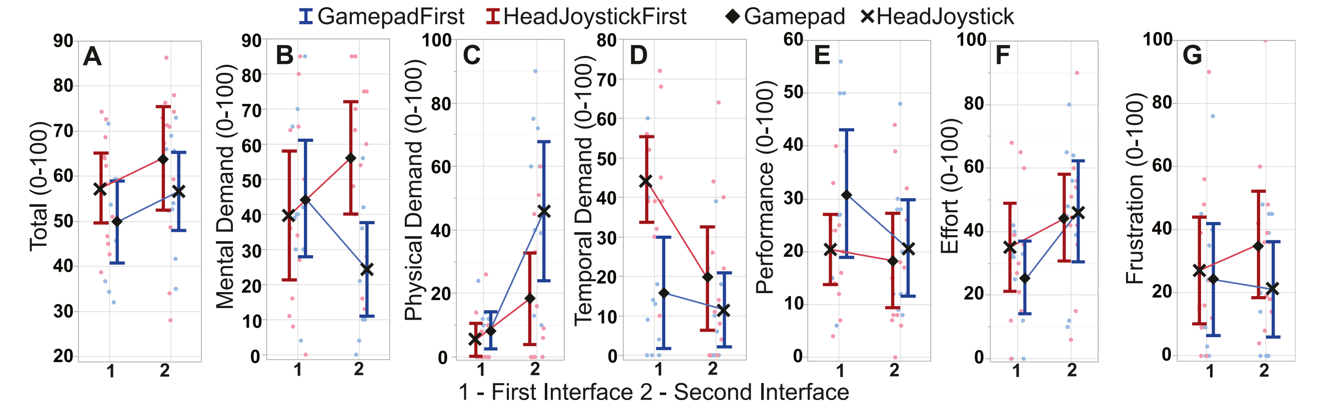


FIGURE 8 | NASA TLX in %, after using the first interface then second interface. The blue line and red line indicate participants using Gamepad and HeadJoystick respectively, for their first trial. ◆ = Gamepad interface, X = HeadJoystick interface, CI = 0.95.

5.2.2.1 The Participants Felt the Task was Overall More Demanding With the Second Interface Irrespective of the Group

Figure 8A and **Table 2**. However, there was no significant effect of group and no significant interaction between time and group for total NASA TLX scores. This equivalently means that there was no significant main effect of the interface.

5.2.2.2 Participants Felt HeadJoystick was Mentally Less Demanding

Figure 8B. The analysis did not show a main effect of time or group. However, there was a significant interaction between time and group. Irrespective of the group, the mental demand with the first interface was around 50%. However, when the participants switched from Gamepad to HeadJoystick they found the mental demand to be significantly reduced, whereas in the group that switched to Gamepad from HeadJoystick, the mental demand ratings went significantly up for the second interface. This is corroborated by the significant overall effect of interface, with significantly higher mental demand ratings for the Gamepad ($M = 50.8$, $SD = 24.5$) than HeadJoystick ($M = 32.7$, $SD = 5.43$), $F(1, 20) = 6.70$, $p = .018$, $\eta_p^2 < .001$.

5.2.2.3 Participants Felt HeadJoystick was Physically More Demanding

Figure 8C. The analysis showed a main effect of time as well as group. The second interface was rated as more physically demanding ($M = 30.9$, $SD = 29.5$) than the first interface ($M = 6.77$, $SD = 8.11$) and the GamepadFirst group found the task to be more physically demanding ($M = 27.2$, $SD = 4.47$) than the HeadJoystickFirst group ($M = 11.9$, $SD = 4.08$). There was also an interaction between time × group. Thus, the physical demand of the HeadJoystick ($M = 23.8$, $SD = 29.4$) was rated higher than that of Gamepad ($M = 13.8$, $SD = 18.0$), $F(1, 20) = 4.43$, $p = .048$, $\eta_p^2 = .181$.

5.2.2.4 Participants Found Gamepad to Marginally Decrease Temporal Demand (Time Pressure)

Figure 8D

As with the physical demand, there was a main effect on time as well as group. In general, the second interface ($M = 15.8$, $SD = 17.7$) had lower temporal demand than the first interface ($M = 31.5$, $SD = 23.1$). In particular, the group that switched from HeadJoystick to Gamepad reported lower time pressure registering a marginally significant interaction; i.e., Gamepad ($M = 17.8$, $SD = 19.8$) had marginally lower temporal demand than HeadJoystick ($M = 29.5$, $SD = 22.6$).

5.2.2.5 The Participants Felt the Task Needed More Effort in their Second Trial Irrespective of the Interface

Figure 8F. The second interface ($M = 45.3$, $SD = 21.3$) had significantly higher effort ratings than the first interface ($M = 30.8$, $SD = 19.6$). There was no significant difference between the groups or interaction between time and group.

There were no statistically significant main effects or interactions on performance or frustration **Figures 8E,G**.

5.2.3 Post-Experiment Questionnaire and Interview

Participants filled out a post-experiment questionnaire with 22 questions for each of the two interfaces, addressing different aspects like usability and performance, motion sickness, comfort, and immersion. The ratings were compared using t-tests, or Wilcoxon signed-rank tests whenever the assumption of normality was violated.

Figure 9 summarizes descriptive and inferential statistics. As seen from the plot, while participants did not have a strong opinion about Gamepad for the majority of the statements (most averages were near 5, neither agree nor disagree), for HeadJoystick they had a stronger positive opinion (positive statements) or stronger negative opinion (negative statements). Compared to the Gamepad, the HeadJoystick interface was rated as **easier to learn**, **easier to use**, gave more **control**, and was more **enjoyable** and **preferable**. It also made them **less motion sick** than Gamepad while increasing **immersion** and **vection**. That is, all significant effects were in favour of the HeadJoystick over the Gamepad. Both interfaces were judged to provide a comfortable sitting position (Gamepad, $M = 6.83$, $SD = 0.551$| HeadJoystick, $M = 7.92$, $SD = 0.394$), although the Gamepad provided a slightly (but only marginally significantly) more comfortable **sitting posture**.

GamePad Mean (M)	GamePad Standard Error (SE)	HeadJoystick Mean (M)	HeadJoystick Standard Error (SE)	Statistical Test (paired t-test/ Wilcoxon signed-rank)	Statements
6.82	0.583	8.68	0.357	Z = -2.62, p = .005	The interface was easy to learn
5.14	0.625	7.64	0.495	Z = -2.64, p = .008	The interface was easy to use
4.82	0.794	6.77	0.664	Z = -1.79, p = .074	I could easily navigate around the play area
5.00	0.736	7.32	0.413	Z = -2.67, p = .007	I had precise control of my movements
5.45	0.664	7.55	0.346	Z = -2.38, p = .017	The efficiency and stability of the interface for motion control is generally high
6.41	0.620	4.45	0.533	Z = -2.56, p = .010	* The task difficulty of keeping the right distance was high
6.45	0.553	4.73	0.596	t(21) = 2.17, p = .042	* Task difficulty of staying on intended path was high
6.14	0.703	7.27	0.373	Z = -1.31, p = .191	I could easily concentrate on the task
4.95	0.749	7.86	0.385	Z = -2.72, p = .007	I enjoyed using the interface
5.36	0.719	7.55	0.404	Z = -2.30, p = .022	The overall usability of the interface is high
4.36	0.788	7.86	0.433	Z = -2.70, p = .007	Overall, I prefer this interface
4.64	0.903	3.27	0.649	Z = -2.12, p = .034	* I felt dizzy
4.57	0.953	2.82	0.602	Z = -2.33, p = .020	* I felt sick
6.41	0.720	8.36	0.339	Z = -2.45, p = .014	I had a strong sensation of self-motion with the interface
5.36	0.663	7.95	0.319	t(21) = -4.21, p<.001	The interface helped me to be involved and engaged
5.36	0.673	6.32	0.604	t(21) = -2.63, p = .016	I was completely captivated by the virtual world
4.91	0.775	5.50	0.732	Z = -1.32, p = .189	I was no longer aware of my real environment
8.09	0.410	6.91	0.599	Z = -1.91, p = .057	My sitting posture was comfortable
6.32	0.601	5.68	0.608	t(21) = .794, p = .436	My muscles were relaxed
4.91	0.624	6.86	0.438	t(21) = -2.26, p = .034	Using the Interface was comfortable
4.09	0.797	5.59	0.572	Z = -1.42, p = .155	I think I would like to use the interface regularly
4.27	0.851	5.73	0.593	Z = -1.25, p = .210	I could imagine using the interface for longer time periods

Note: * Denotes statement with undesirable quality

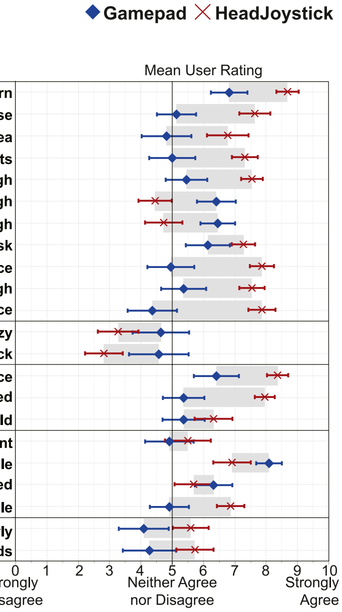


FIGURE 9 | User rating regarding statements about usability and preference, motion sickness, immersion, comfort, and long-term use. Green and lighter shade of green respectively indicate significant ($p \le 5\%$) and marginally significant ($p \le 10\%$) differences in favor of HeadJoystick. Lighter shade of red indicates marginally higher ratings in favor of Gamepad ($p \le 10\%$). Note: * denotes statements with undesirable quality (reversed scale). Effectively, the green highlights in those statements with * indicate that HeadJoystick had lower task difficulty and made participants less motion sick. ◆ = Gamepad interface, × = HeadJoystick interface, *CI* = 95%.

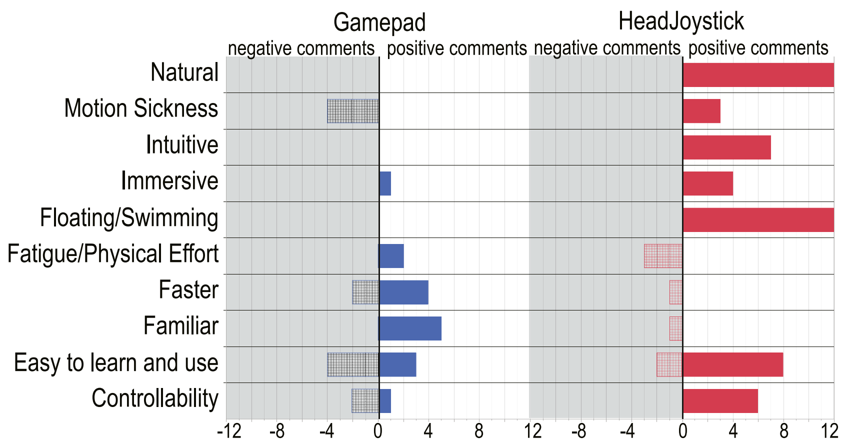


FIGURE 10 | Thematic count of participants' response in post-experiment interview. The *x*-axis represents the number of participants mentioning the themes on the *y*-axis. Bars to the left (gray area) correspond to the number of negative comments about that topic whereas bars to the right (white background) represent the number of positive comments.

We performed semi-structured open-ended interviews with the participants after they completed the post-experiment questionnaire to get more insight into their choices and underlying reasons for those choices. 16 out of 22 participants mentioned in the interview that they preferred HeadJoystick over Gamepad. As seen in **Figure 10**, the recurrent themes among the participants for preferring the HeadJoystick over Gamepad were that HeadJoystick made them **less sick**, it was **easier to learn and use**, it was **intuitive**, it provided better **controllability**, it felt **natural**, there was a stronger sense of **self-motion (floating/swimming)**, and the virtual environment felt more **immersive**. Even though these participants preferred HeadJoystick over Gamepad, they still felt Gamepad had the advantages of familiarity, HeadJoystick would be fatiguing after a long use, and "*it would be hard to stand still with HeadJoystick*" {P13}.

Five participants preferred Gamepad over HeadJoystick. The recurrent themes among the participants for preferring the Gamepad were **familiarity**, **ease of control**, **faster** and **less physical effort**. However, even among those who preferred Gamepad over HeadJoystick, some mentioned that the HeadJoystick made them less sick and a few appreciated the novel approach to VR locomotion.

Among the listed thematic counts in **Figure 10**, motion sickness turned out to be such a predominant concern with Gamepad that when we asked, "*How was your experience*?" as the first question after the experiment, several participants immediately responded:

> "*Felt like throwing up after using Gamepad*" {P03}
>
> "*Fun with the HeadJoystick but got dizzying after using Gamepad*" {P18}
>
> "*Gamepad was terrible … made me almost sick*" {P13}
>
> "*…Gamepad made me sick …* " {P25}

As we wanted to understand why they preferred one interface over the other we had also specifically asked "*What did you like about the locomotion interfaces*?", the minimal cognitive load and intuitiveness of the HeadJoystick was one of the most consistent responses.

> "*HeadJoystick is more intuitive. You don't really have to learn to use it.*" {P17}
>
> "*There was no cognitive load … with the HeadJoystick, motion was intuitive and [I] could concentrate more in the task.*" {P13}
>
> "*Head one [HeadJoystick] is more intuitive. It feels like I am swimming.*" {P23}

As for the Gamepad, participants liked that they were "*familiar*" {P16, P22} with its mechanics. However, they were split between "*easy to use*" {P12} due to its familiarity and "*difficult*" {P25} due to confinement in a single plane, i.e., "*moved in either vertical direction or moved in the horizontal plane*" {P01} as well as the apparent disjunction between "*two different kind of movements*" {P09, P23}, i.e., physical rotation and controller translation.

Other reasons for preferring one interfaces over the others included enjoyment, better control, naturalness and required effort:

> *"Because [HeadJoystick] was easier to learn. More enjoyable-feels like flying in VR."* {P05}
>
> *"[HeadJoystick] gives more control and [is] more precise."* {P09}
>
> *"[HeadJoystick] is more matched to the body. Felt similar to scuba diving."* {P14}
>
> *"HeadJoystick is easy to control and felt more immersive."* {P15}
>
> *"The Vive did not fit perfectly. It was also heavy. So, I was wary about moving properly for the HeadJoystick.* {P06}
>
> *"Gamepad made me dizzy, but still, it required less effort."* {P18}

We also asked specific question regarding their strategies with both interfaces. When we asked, "*Did you use any strategies*? *Were they different for the different interfaces*?", many participants indicated using different search strategies depending on which interface they used. For Gamepad, they tended to first search horizontally, went up or down and then searched on that new level and so on, while avoiding motions that involved all three translational degrees of freedom. For example, P12 stated that they "*stayed in [the] same level, searched there then changed altitude*". At the same time, for HeadJoystick they "*looked around in circle*" {P13}. Participants' descriptions of their travel strategies of moving in distinct planes for the controller but more fluidly through 3D space with the HeadJoystick mirrors the plots of their trajectories in **Figure 6**.

6 DISCUSSION

This paper presents the first study exploring the effect of partial body-based self-motion cues, in the form of a leaning-based interface, on spatial orientation/updating while flying in virtual 3D space. Currently, flying is typically achieved through low-fidelity interfaces like a gamepad, joystick, keyboard, or point-and-click teleportation, or through high-fidelity interfaces with actuators or motors, like motion platforms McMahan et al. (2011). We explore a relatively novel flying interface (HeadJoystick) that tries to bring together the advantages of both low- and high-fidelity interfaces: it is embodied, inexpensive, easy to set up, provides at least minimal translational motion cueing and full rotational cues, and is capable of controlling all four DoFs needed for full flight control, as discussed in more detail in Hashemian et al. (2020).

Though past studies have shown that leaning-based interfaces can improve spatial perception and orientation in ground-based (2D) locomotion Harris et al. (2014), Nguyen-Vo et al. (2019), leaning-based interfaces with the capability to control all four DoFs needed for full flight control have not been scrutinized in these contexts. We discuss the findings of our experiment in the context of our main research questions below.

6.1 RQ1: Does HeadJoystick Improve Navigational Search Performance Compared to a Hand-Held Controller for 3D Locomotion?

Participants collected significantly more balls with HeadJoystick, while being more efficient, i.e., travelling less distance. These quantitative findings indicate improved navigational search performance for the HeadJoystick. We propose that the better performance of HeadJoystick can be mainly attributed to two factors. First, participants mention a strong sense of self-motion in our study when using the HeadJoystick. They felt they were "*actually floating*" {P19} and moving in the space. The importance of non-visual and embodied self-motion in a variety of spatial orientation and updating tasks has been shown through a number of studies Chance et al. (1998), Klatzky et al. (1998), Kearns et al. (2002), Sun et al. (2004), Waller et al. (2004), Ruddle and Lessels (2009), Riecke et al. (2010), Ruddle et al. (2011b), Ruddle (2013), Harris et al. (2014), Riecke et al. (2015), as discussed in **Section 2.1**. Further, in Nguyen-Vo et al. (2019)'s study, the interfaces providing partial body-based sensory information (in particular, leaning-based minimal translational self-motion cues) performed significantly better compared to controller-based locomotion in a task heavily reliant in spatial updating, even though real rotation was applied in all cases (standing and sitting interfaces). Together with these earlier findings, the current study confirms that the improved navigational search performance of a well-designed leaning-based interface for ground-based locomotion Nguyen-Vo et al. (2019) do indeed generalize to 3D locomotion (flying). As we argued in **Section 2.2**, spatial updating is a necessary prerequisite for effective navigational search. This suggests that the improved performance for the HeadJoystick might be related to it better supporting automatic spatial updating, although further research is needed to disambiguate spatial updating from other potential underlying mechanisms.

Second, the Gamepad and HeadJoystick interfaces used in the current study facilitated different kinds of motions and subsequently led to different locomotion trajectories and potential underlying search strategies-at least when the vertical direction was involved. For instance, HeadJoystick afforded changing the travel direction in all three axes with a single head motion while Gamepad required combination of two thumbsticks. Participants stated in the post-experiment interview that they searched on different horizontal levels with the Gamepad and searched in a more spiraling fashion with HeadJoystick. These statements are corroborated by the difference in travel paths between the interfaces illustrated in **Figure 6** and the **Supplementary Material**. With Gamepad the participants' search pattern reflected a switching between the motion control between two hands (and thumbsticks), i.e., they typically did not make use of the bi-manual control of motion to go directly towards a selected target and change height at the same time. Hence, when seen from above the foraging paths look qualitatively similar, whereas the side view highlights the switch between horizontal and vertical locomotion. However, with the HeadJoystick, their movements seemed to be less restricted to any plane or axis. This could be based on participants being reluctant to control multiple degrees of freedom simultaneously when using the Gamepad: participants mostly controlled just one DoF at a time (e.g., forward translation) in combination with full body rotations with their head mostly facing forward. In contrast, HeadJoystick seems to much better facilitate traveling in any direction and controlling all three translational degrees of freedom simultaneously as evident from participant quotes and differences in movement trajectories. This assumption is supported by the travel path shown in **Figure 6** and the **Supplementary Material**, where the top-down views for both interfaces are fairly similar, but show a stark contrast in the side/front views. Future research is needed to investigate if these different interfaces might lead to different foraging strategies (such as different orders in which the boxes where visited), or if participants use overall fairly similar foraging strategies but separate horizontal and vertical locomotion due to different affordances of the interfaces. The observed locomotion trajectories seem to suggest that the order of boxes visited did not depend a lot on the locomotion interface, at least for the first few boxes. However, as the task became successively harder with less balls left to find, participants using the gamepad seemed to travel further and revising more boxes, potentially because they got more easily lost and situational awareness and spatial updating ability was reduced. Further research is needed to further investigate these aspects and more specifically address underlying strategies and processes.

In sum, the above analysis suggests that the standard two-thumbstick controller creates a mapping problems between the interface input and the resulting effect (simulated self-motion in VR), especially when used for 3D (flying) locomotion. This is supported by participant feedback, e.g., {P01} mentioned that *"With gamepad I moved in either vertical direction or moved in the horizontal plane. With HeadJoystick it was a combination. So, it felt much easier"*. Motivated in part by the observed challenges in controlling multiple degrees of freedom with separate hands/thumbsticks in the current gamepad condition, in a recent spatial updating study we implemented a controller-directed steering mechanism where the direction of the hand-held controller indicates the direction they want to travel Adhikari et al. (2021). However, that study showed that even adding some level of embodiment (the controller direction) and not requiring 2-handed operation was still not enough for it to support the same spatial updating performance as observed with leaning-based interfaces. We propose that the additional vestibular/proprioceptive cues provided by the head/trunk movement of leaning-based interfaces might help reduce the intersensory cue conflict in VR when users cannot physically walk, and more effectively trigger compelling sensations of self-motion and support automatic spatial updating, thus improving navigation performance while reducing cognitive load. Future carefully designed studies are needed to test these hypotheses and for example disambiguate contributions of spatial updating from other potential underlying factors.

Note that using the HeadJoystick lead to a marginally significant increase in overall head (but not body) rotations, suggesting that participants might be looking around more and/or looking more into their periphery. This might contribute to an increased situational awareness with HeadJoystick.

The behavior of being confined to planes or axes is quite common for input devices that separate control dimensions. Thereby, studies of user performance of 3D tasks with 2D input devices have shown performance decrements, especially for selection and manipulation tasks Zhai (1998). Basically, six DoF need to be mapped to controllers that usually only afford two DoF (e.g., a micro-joystick). While precision (accuracy) may be higher (an issue we will reflect on below), trials tend to take longer than with 3D input devices as different control axes have to be controlled independently and serially (e.g., in case of a mouse). Even when multiple axes can be controlled in parallel by being able to use two controllers at once with two hands (or fingers, e.g., by using a gamepad), analysis of the movement paths of such tasks still show jagged patterns (a sign of non-optimal movement paths), similar to our typical paths in **Figure 6**. Even more so, control also requires coordination between both hands. The comparison between HeadJoystick and Gamepad shows that HeadJoystick performs better for larger (course) movements as evident from our task (quickly traveling from one location to another while maintaining spatial orientation). Surprisingly, fine grained movement has also been shown to be better with HeadJoystick, compared to Gamepad, as shown in Hashemian et al. (2020)'s maneuvering task (travelling through narrow tunnels). While some of these benefits have also been shown for specialized 3D desktop input devices (in contrast to handheld "free-air" input devices), these devices tend to require much training to precisely control them. In contrast, users of HeadJoystick did improve performance over time (steady learning slope), yet not as drastically as shown for 3D desktop input devices with a steep slope over multiple usage sessions, where users started with low performance Zhai (1998).

Despite these interface-specific locomotion strategies there was no significant difference in task completion time or revisit percentage to the boxes. The completion time is influenced by the auto-termination of the program after 5 min, terminating 13 HeadJoystick trials and 20 Gamepad trials. As for the revisits, the participants complained that the task was *"difficult"* {P14} and that they easily *"lost track"* {P1, P20, P25} with both interfaces. Removing any global orientation cues dropped the percentage of perfect trials (no revisits) in HMD with real walking from 90% in Ruddle and Lessels' studies Ruddle and Lessels (2006, 2009) to 13.9% in Riecke Riecke et al. (2010)'s study. Further limiting overall visibility such that the whole layout of boxes could never be seen at once but had to be integrated during locomotion further decreased the percentage of perfect trials. There were no perfect trials without revisit in Nguyen-Vo et al. (2019)'s ground based locomotion study, and just 5.68% in this study (5 out of 88 trials). Thus, similar to the ground-based locomotion study, we did not find any difference between the interfaces for revisits. However, we can see a trend of increasing revisits with the Gamepad compared to HeadJoystick as seen in **Figure 5**. As the data have been split for eight different balls, the data is too scarce to provide reliable statistical conclusions, though. Further study concentrating on how interfaces' performance changes with increasing difficulty could shed more light into the matter.

6.2 RQ2: Does HeadJoystick Help to Reduce Motion Sickness?

Table 1 shows a clear advantage for the HeadJoystick in comparison to Gamepad interface. Gamepad caused a significant increase in motion sickness (for total SSQ and Nausea sub-scale) independent of whether it was the first or second interface. When Gamepad was used as the second interface, oculomotor and disorientation issues also significantly increased compared to Pre-Experiment scores. Further, oculomotor ($p = 0.041$) and disorientation ($p = 0.050$) issues showed a marginal increase compared to Pre-Experiment scores (that would be significant if we had not applied the fairly conservative Bonferroni correction), even with the Gamepad as first interface. Although HeadJoystick caused a marginal increase in oculomotor issues when it was the first interface and a marginally significant increase in disorientation when it was the second interface, there was no overall significant increase in overall motion sickness score or any of the sub-scales. In sum, there was a consistent increase in overall and some of the SSQ subscores for the Gamepad, and only marginal increases in oculomotor and disorientation scores for the HeadJoystick. This confirms our hypothesis 2.

The single-item Fast Motion-Sickness Scale (FMS) also showed a distinct difference in motion sickness between Gamepad and HeadJoystick. With a more straight-forward rating system of "0-I am completely fine and have no motion sickness symptoms" and "100-I am feeling very sick and about to throw up", the participants reported that their motion sickness increased significantly more with Gamepad than HeadJoystick. Further, one of the participants completed the tasks with the HeadJoystick but dropped it due to motion sickness after trying the Gamepad. Two more participants let us know after the trial that they were *"about to drop the trial"* {P03, P06} with the Gamepad before it auto-terminated due to the time limit.

These quantitative findings are corroborated by the participants' responses in open-ended interviews. Previous ground-based locomotion studies have shown that combining visual motion with body-based sensory information can help to make people less sick Aykent et al. (2014), Lawson (2014). However, the results have been mixed for leaning-based interfaces. Marchal et al. (2011)'s study (Joyman) and Hashemian and Riecke (2017a)'s study (SwivelChair) showed no significant difference in motion sickness between a leaning-based interface and a hand-based controller. Nguyen-Vo et al. (2019)'s study showed a significant motion sickness reduction for a standing interface (NaviBoard) as compared to a hand-held controller, but not for a sitting interface (NaviChair). Similarly, Hashemian et al. (2020)'s flying study (HeadJoystick) showed a significant reduction in motion sickness for real-rotation with leaning-based translation compared to controller-based translation and rotation conditions, but only varying the translation mechanism did not produce a change that was statistically significant. A closer look into Marchal et al.'s study show that a Likert scale of 7 was used to measure motion sickness, and both conditions barely caused any motion sickness (average of 6, where 7 meant no motion sickness symptoms). Further, Hashemian and Riecke's SwivelChair and Nguyen-Vo et al.'s Navichair were not compared with pre-experiment SSQ scores but only with SSQ scores after using Gamepad (HeadJoystick vs. Gamepad) Hashemian and Riecke (2017a), Nguyen-Vo et al. (2019). Hashemian et al.'s study compared the motion sickness of the interfaces by subtracting the pre-experiment values of SSQ from each interface (HeadJoystick-Pre-experiment vs. Gamepad-Pre-Experiment). In these three studies, the leaning-based interfaces

(SwivelChair, NaviChair and HeadJoystick) showed a trend towards lower motion sickness than the controller condition, but were shy of statistical significance. In our study, we performed a planned contrast to compare all the conditions with pre-experiment SSQ scores (HeadJoystick vs. Pre-experiment and Gamepad vs. Pre-experiment). Combining this comparison with the results from FMS, post-experiment questionnaires and participants' testimony shows that Gamepad makes participants significantly more motion sick than HeadJoystick, and that the SSQ might be too conservative a measure for registering those differences. Even in the meta-analysis of the effectiveness of SSQ, in spite of having a strong correlation between SSQ scores and participant drop out, the dropped-out participant still rated only a 39.63 total SSQ score on average out of 235.62 (16.8%) Balk et al. (2013). Thus, we recommend using the SSQ for comparing the pre-vs. post-use motion sickness scores for each interfaces, rather than just comparing post-use motion sickness scores among the interfaces.

Therefore, our study, along with Hashemian et al. (2020)'s study indicates that a partial body-based locomotion interface can help to reduce motion sickness not only for ground-based locomotion but also for flying locomotion. In addition, participants also reported a significantly stronger sensation of self-motion with the HeadJoystick in the post-experiment questionnaire **Section 5.2.3**. They described their HeadJoystick experience as *"natural"* in post-experiment interviews and compared it to *"swimming"* or *"floating,"* see **Figure 10**. This rejects the concern sometimes mentioned in the literature that increasing vection might also increase motion sickness Hettinger et al. (1990), Hettinger and Riccio (1992), Stoffregen and Smart (1998), Smart et al. (2002). Instead, our data show that increasing vection can, in fact, be accompanied by reduced motion sickness, e.g., if a carefully designed embodied interface like HeadJoystick is used. That is, vection is not in general a sufficient prerequisite or predictor for motion sickness, and often does not even correlate with motion sickness-see discussion in Keshavarz et al. (2015), Riecke and Jordan (2015).

6.3 RQ3: How do HeadJoystick and Gamepad Affect the Overall User Experience and Usability?

In terms of cognitive load, the participants felt HeadJoystick could be used with significantly lower mental demand than Gamepad. Users also rated HeadJoystick to be significantly easier to learn, easier to use, and more precise. Though some participants preferred Gamepad for its familiarity, others complained the apparent disjunction between physical rotation and controller translation. Previous studies have also documented participants complaining about the disjunction between the physical rotation and controller translation Hashemian et al. (2020), Hashemian and Riecke (2017b). All of these observations confirm our hypothesis 3, that HeadJoystick should be more intuitive to use and learn, and the user should be able to use HeadJoystick more effectively, even without any previous exposure.

Further, our participants highly and consistently preferred HeadJoystick over Gamepad for helping to reduce motion sickness, and for its ease of use, controllability, and learnability. They also found it more immersive, engaging and enjoyable. This mirrors the finding of previously discussed leaning-based flying interfaces Harris et al. (2014), Kitson et al. (2015b), Kruijff et al. (2016), Marchal et al. (2011) in **Section 2.1**. In Higuchi and Rekimoto's study, participants preferred the physical interface for ease of control, ease of use and enjoyment Higuchi and Rekimoto (2013). The participants in Pittman and LaViola's user study appreciated the Head-Rotation for being natural, intuitive and immersive Pittman et al. (2014).

Hashemian et al. (2020)'s study also compared two interfaces on a number of user experience and usability factors including enjoyment, immersion, vection, long-term use, daily use, ease of use and ease of learning. The participants found "HeadJoystick" to be better in all the criteria. Our findings echo most of these findings, except for long-term use where HeadJoystick in our study had a higher average but did not reach statistical significance. This suggests that leaning-based interfaces, if designed well, can provide a fairly clear affordance Riecke and Zielasko (2020) in multiple kind of environments and tasks, which can be further improved by providing a brief demonstration or showing a video of the interfaces as in Nguyen-Vo et al. (2019).

We also explicitly asked participants about their reason for preferring one interface over the other. The recurrent themes, as summarized in **Figure 10** and as also seen throughout the discussion participants, were mainly concerned with ease of learning and use, motion sickness, and familiarity. In particular, HeadJoystick was appreciated for being easy to learn and use, making them less motion sick, and providing a strong sensation of self-motion. Gamepad was preferred for familiarity, less physical effort and being faster.

One of the issues with leaning-based interfaces could be comfort and stability, especially when there is no backrest. Standing leaning-based interfaces or sitting interfaces without a backrest can make users wary of losing balance and falling Badcock et al. (2014), Kitson et al. (2015a). However, in Kitson et al.'s paper, when the leaning-based interface was used with a swivel chair having a backrest, it was rated as the most comfortable interface. HeadJoystick had significantly higher comfort ratings in our study too, and there were no concerns about stability or falling. However, participants still rated the sitting posture of Gamepad as marginally more comfortable compared to HeadJoystick, which might be related to the need to constantly adjust one's posture during movement changes with HeadJoystick. Similarly, participants reported significantly higher physical demand and marginally higher temporal demand for HeadJoystick than the Gamepad. All of these issues might be related to us trying to make the interface stable and safe. In order to create a stable and safe interface suitable for diverse participants with a wide range of physical features, we designed the interface to require relatively large leaning postures for high speed. Therefore, either the participants leaned excessively to achieve high velocity and experienced higher physical demand and uncomfortable seating posture, or they leaned moderately but travelled slower and experienced higher temporal demand. Further fine tuning the mapping, or even allowing for personalized or context-dependent speed mappings (e.g., allowing for faster speeds in large or outdoor spaces), could help to address these issues.

In sum, our main research hypotheses were all confirmed: compared to the Gamepad, the HeadJoystick improved navigational search performance and thus presumably also spatial updating (RQ1), helped to reduce motion sickness (RQ2), and resulted in improved user experience and usability ratings across all measures (RQ3).

7 CONCLUSION

The current study provides the first compelling experimental evidence that providing partial body-based self-motion cues through leaning/head-movements with full physical rotation can improve performance in a virtual flying navigational search tasks that requires users to maintaining situational awareness through continuous spatial updating. It also provides evidence that an interface with embodied control like the HeadJoystick can be intuitively learned and used effectively without any previous exposure or lengthy training or practice. Finally, it shows that partial body-based self-motion cues from leaning/head-movements can mitigate the conflict between visual information and vestibular cues observed in controller-based interfaces like a Gamepad, and thus arguably help to minimize motion sickness.

Whereas Hashemian et al. (2020) showed similar benefits of HeadJoystick over Gamepad for a gamified VR waypoint navigation (maneuvering) task, the current study shows that these benefits extend to a novel 3D navigational search task that requires spatial updating and building up and maintaining a mental representation of a large array of objects and thus situational awareness. Together, this suggests that our HeadJoystick locomotion interface can be useful in a wide range of 3D flying scenarios and applications, even if they require spatial updating and/or situational awareness, engaging novice users, or the need to minimize motion sickness. Moreover, as HeadJoystick implements four DoF control as in quadcopter drones, it also has the potential to be integrated with UAVs and used as a potentially more immersive, embodied, and intuitive control interface. A comprehensive future study combining all these isolated factors (maneuvering and spatial updating) together in a virtual task mimicking real world scenario can provide a better insight into this. In the future, we also want to investigate how the advantages of a leaning-based interface translate into immersive 3D telepresence scenarios in the context of flying experience, usability aspects, and performance measures.

In the current study we intentionally removed all landmarks to avoid related confounds and require continuous spatial updating. Future research could investigate how the observed differences between leaning- and controller-based interfaces might or might not generalize to other environments that include landmarks. A recent study suggests that an advantage of leaning-over controller-based interfaces indeed remains for a pointing-based spatial updating study even when a landmark-rich naturalistic city environment is used Adhikari et al. (2021), although that study only investigated ground-based locomotion.

Given that HeadJoystick's approach is compact and affordable, it can be easily integrated into existing systems without any additional cost or setup, provided a swivel chair is available. Though we attached the Vive tracker for detecting the center in a rolling chair, the interface also works efficiently without the tracker by fixing the chair in a place. Additionally, a pilot study showed that the interface could easily be used while standing, which might be more suitable for specific scenarios, or to provide more user engagement/movement abilities Zielasko and Riecke (2020). Further research could investigate this.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Simon Fraser University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AA, BR, and TN-V conceived of the presented idea. AA modified the existing experiment design and VR simulation from TN-O's previous work under BR's guidance. BR, MH, EK, and AH provided feedback throughout the study. AA collected and analyzed data. AA wrote the first draft of the manuscript. All authors discussed the results and contributed to the final manuscript.

FUNDING

This research was supported by Natural Sciences and Engineering Research Council (NSERC) (grant: #R611547).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2021.730334/full#supplementary-material

REFERENCES

Adhikari, A., Zielasko, D., Bretin, A., Heyde, M. v. d., Kruijff, E., and Riecke, B. E. (2021). "enIntegrating Continuous and Teleporting VR Locomotion Into a Seamless "HyperJump" Paradig," in *IEEE VR Workshop on Finding a Way Forward in VR Locomotion* (Lisbon, Portugal: IEEE), 1–2. Available at: http://ieeevr.org/2021/contribute/workshoppapers/#Finding-a-way-forward-in-VR-locomotion.

Aykent, B., Merienne, F., Guillet, C., Paillot, D., and Kemeny, A. (2014). Motion Sickness Evaluation and Comparison for a Static Driving Simulator and a Dynamic Driving Simulator. *Proc. Inst. Mech. Eng. D: J. Automobile Eng.* 228, 818–829. doi:10.1177/0954407013516101

Badcock, D., Palmisano, S., and May, J. (2014). "Vision and Virtual Environments," in *Handbook of Virtual Environments*. Editors K. S. Hale and K. M. Stanney. 2nd ed (Boca Raton, FL: CRC Press), 39–85. doi:10.1201/b17360-6

Balk, S. A., Bertola, M. A., and Inman, V. W. (2013). Simulator Sickness Questionnaire: Twenty Years Later. *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*. doi:10.17077/drivingassessment.1498

Bonato, F., Bubka, A., Palmisano, S., Phillip, D., and Moreno, G. (2008). Vection Change Exacerbates Simulator Sickness in Virtual Environments. *Presence: Teleoperators and Virtual Environments.* 17, 283–292. doi:10.1162/pres.17.3.283

Bowman, D., Kruijff, E., LaViola, J. J., Jr, and Poupyrev, I. P. (2004). *3D User Interfaces: Theory and Practice*. Boston: CourseSmart eTextbook Addison-Wesley.

Chance, S. S., Gaunet, F., Beall, A. C., and Loomis, J. M. (1998). Locomotion Mode Affects the Updating of Objects Encountered during Travel: The Contribution of Vestibular and Proprioceptive Inputs to Path Integration. *Presence.* 7, 168–178. doi:10.1162/105474698565659

Cheung, B., Howard, I., and Money, K. (1991). *Visually-Induced Sickness in Normal and Bilaterally Labyrinthine-Defective Subjects*. Washington DC: Aviation, space, and environmental medicine.

Cirio, G., Marchal, M., Regia-Corte, T., and Lécuyer, A. (2009). "The Magic Barrier Tape: a Novel Metaphor for Infinite Navigation in Virtual Worlds With a Restricted Walking Workspace," in Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology, 155–162.

Farrell, M. J., and Robertson, I. H. (1998). Mental Rotation and Automatic Updating of Body-Centered Spatial Relationships. *J. Exp. Psychol. Learn. Mem. Cogn.* 24, 227–233. doi:10.1037/0278-7393.24.1.227

Fiore, L. P., Coben, E., Merritt, S., Liu, P., and Interrante, V. (2013). "Towards Enabling More Effective Locomotion in VR Using a Wheelchair-Based Motion Platform," in Proceedings of the 5th Joint Virtual Reality Conference (Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, JVRC '13), 83–90. doi:10.2312/EGVE.JVRC13.083-090

Groen, E. L., and Bles, W. (2004). How to Use Body Tilt for the Simulation of Linear Self Motion. *J. Vestib. Res.* 14, 375–385.

Harris, A., Nguyen, K., Wilson, P. T., Jackoski, M., and Williams, B. (2014). "Human Joystick: Wii-Leaning to Translate in Large Virtual Environments," in Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry (New York, NY: ACM), 231–234.

Hart, S. G., and Staveland, L. E. (1988). Development of Nasa-Tlx (Task Load Index): Results of Empirical and Theoretical Research. *Adv. Psychol.* 52, 139–183. doi:10.1016/s0166-4115(08)62386-9

Hashemian, A., Lotfaliei, M., Adhikari, A., Kruijff, E., and Riecke, B. E. (2020). *HeadJoystick: Improving Flying in VR Using a Novel Leaning-Based Interface*. New York, NY: IEEE Transactions on Visualization and Computer Graphics.

Hashemian, A. M., and Riecke, B. E. (2017a). "Leaning-Based 360° Interfaces: Investigating Virtual Reality Navigation Interfaces with Leaning-Based-Translation and Full-Rotation," in International Conference on Virtual, Augmented and Mixed Reality (Springer), 15–32. doi:10.1007/978-3-319-57987-0_2

Hashemian, A. M., and Riecke, B. E. (2017b). Swivel-Chair: Evaluating Seated Full-Rotational Interfaces for Virtual Reality Navigation (Vancouver, Canada: Software for Augmented and Virtual Reality (SAVR)).

Hettinger, L. J., Berbaum, K. S., Kennedy, R. S., Dunlap, W. P., and Nolan, M. D. (1990). Vection and Simulator Sickness. *Mil. Psychol.* 2, 171–181. doi:10.1207/s15327876mp0203_4

Hettinger, L. J., and Riccio, G. E. (1992). Visually Induced Motion Sickness in Virtual Environments. *Presence: Teleoperators & Virtual Environments.* 1, 306–310. doi:10.1162/pres.1992.1.3.306

Higuchi, K., and Rekimoto, J. (2013). "Flying Head: a Head Motion Synchronization Mechanism for Unmanned Aerial Vehicle Control," in CHI'13 Extended Abstracts on Human Factors in Computing Systems (New York, NY: ACM), 2029–2038.

Kearns, M. J., Warren, W. H., Duchon, A. P., and Tarr, M. J. (2002). Path Integration From Optic Flow and Body Senses in a Homing Task. *Perception.* 31, 349–374. doi:10.1068/p3311

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *Int. J. aviation Psychol.* 3, 203–220. doi:10.1207/s15327108ijap0303_3

Keshavarz, B., Riecke, B. E., Hettinger, L. J., and Campos, J. L. (2015). Vection and Visually Induced Motion Sickness: How Are They Related?. *Front. Psychol.* 6, 472. doi:10.3389/fpsyg.2015.00413

Keshavarz, B., and Hecht, H. (2011). Validating an Efficient Method to Quantify Motion Sickness. *Hum. Factors.* 53, 415–426. doi:10.1177/0018720811403736

Kitson, A., Riecke, B. E., Hashemian, A. M., and Neustaedter, C. (2015a). "Navichair: Evaluating an Embodied Interface Using a Pointing Task to Navigate Virtual Reality," in Proceedings of the 3rd ACM Symposium on Spatial User Interaction (New York, NY: ACM), 123–126.

Kitson, A., Riecke, B. E., Hashemian, A. M., and Neustaedter, C. (2015b). "NaviChair," in Proceedings of the 3rd ACM Symposium on Spatial User Interaction (New York, NY, USA: Association for Computing Machinery), 123–126. doi:10.1145/2788940.2788956

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., and Golledge, R. G. (1998). Spatial Updating of Self-Position and Orientation during Real, Imagined, and Virtual Locomotion. *Psychol. Sci.* 9, 293–298. doi:10.1111/1467-9280.00058

Kruijff, E., Marquardt, A., Trepkowski, C., Lindeman, R. W., Hinkenjann, A., Maiero, J., et al. (2016). "On Your Feet!: Enhancing Vection in Leaning-Based Interfaces through Multisensory Stimuli," in Proceedings of the 2016 Symposium on Spatial User Interaction (New York, NY: ACM), 149–158.

Kruijff, E., Riecke, B., Trekowski, C., and Kitson, A. (2015). "Upper Body Leaning Can Affect Forward Self-Motion Perception in Virtual Environments," in Proceedings of the 3rd ACM Symposium on Spatial User Interaction (New York, NY, USA: Association for Computing Machinery), 103–112. doi:10.1145/2788940.2788943

Krupke, D., Lubos, P., Demski, L., Brinkhoff, J., Weber, G., Willke, F., et al. (2016). "Control Methods in a Supernatural Flight Simulator," in 2016 IEEE Virtual Reality (VR) (IEEE), 329. doi:10.1109/vr.2016.7504788

LaViola, J. J., Jr, Feliz, D. A., Keefe, D. F., and Zeleznik, R. C. (2001). Hands-Free Multi-Scale Navigation in Virtual Environments. *SI3D.* 1, 9–15. doi:10.1145/364338.364339

Lawson, B. D. (2014). Motion Sickness Symptomatology and Origins. *Handbook of Virtual Environments.*, 536–605. doi:10.1201/b17360-33

Lawson, B., and Riecke, B. E. (2014). "The Perception of Body Motion," in *Handbook of Virtual Environments: Design, Implementation, and Applications*. Editors K. S. Hale and K. M. Stanney. 2nd edn (Boca Raton: CRC Press), Ch. 7, 163–195.

Lessels, S., and Ruddle, R. A. (2005). Movement Around Real and Virtual Cluttered Environments. *Presence: Teleoperators & Virtual Environments.* 14, 580–596. doi:10.1162/105474605774918778

Loomis, J. M., and Philbeck, J. W. (2008). "Measuring Spatial Perception With Spatial Updating and Action," in *Embodiment, Ego-Space, and Action* (New York, NY, US: Psychology Press), 1–43. doi:10.4324/9780203809891-6

Marchal, M., Pettré, J., and Lécuyer, A. (2011). "Joyman: A Human-Scale Joystick for Navigating in Virtual Worlds," in 2011 IEEE Symposium on 3D User Interfaces (3DUI). doi:10.1109/3DUI.2011.5759212

Marsh, W. E., Kelly, J. W., Dark, V. J., and Oliver, J. H. (2013). Cognitive Demands of Semi-Natural Virtual Locomotion. *Presence: Teleoperators and Virtual Environments.* 22, 216–234. doi:10.1162/pres_a_00152

McMahan, R. P. (2011). *Exploring the Effects of Higher-Fidelity Display and Interaction for Virtual Reality Games. Ph.D. Thesis*. Blacksburg, Virginia: Virginia Tech.

McNamara, T. P., Sluzenski, J., and Rump, B. (2008). "Human Spatial Memory and Navigation," in *Learning and Memory: A Comprehensive Reference*. Editor J. H. Byrne (Oxford: Academic Press), 157–178. doi:10.1016/b978-012370509-9.00176-5

Miehlbradt, J., Cherpillod, A., Mintchev, S., Coscia, M., Artoni, F., Floreano, D., et al. (2018). Data-Driven Body-Machine Interface for the Accurate Control of Drones. *Proc. Natl. Acad. Sci. USA*. 115, 7913–7918. doi:10.1073/pnas.1718648115

Nguyen-Vo, T., Riecke, B. E., and Stuerzlinger, W. (2018). "Simulated Reference Frame: A Cost-Effective Solution to Improve Spatial Orientation in Vr," in 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (IEEE), 415–422. doi:10.1109/vr.2018.8446383

Nguyen-Vo, T., Riecke, B. E., Stuerzlinger, W., Pham, D.-M., and Kruijff, E. (2019). Naviboard and Navichair: Limited Translation Combined with Full Rotation for Efficient Virtual Locomotion. *IEEE Trans. Visualization Comput. Graph.* 27, 165–177. doi:10.1109/TVCG.2019.2935730

Perusquía-Hernández, M., Enomoto, T., Martins, T., Otsuki, M., Iwata, H., and Suzuki, K. (2017). "Embodied Interface for Levitation and Navigation in a 3d Large Space," in Proceedings of the 8th Augmented Human International Conference (Blacksburg, Virginia: ACM). doi:10.1145/3041164.30411734

Pittman, C., and LaViola, J. J., Jr (2014). "Exploring Head Tracked Head Mounted Displays for First Person Robot Teleoperation," in Proceedings of the 19th international conference on Intelligent User Interfaces (New York, NY: ACM), 323–328. doi:10.1145/2557500.2557527

Presson, C. C., and Montello, D. R. (1994). Updating After Rotational and Translational Body Movements: Coordinate Structure of Perspective Space. *Perception*. 23, 1447–1455. doi:10.1068/p231447

Rheiner, M. (2014). Birdly an Attempt to Fly. *ACM SIGGRAPH Emerging Tech.* 3, 1–3. doi:10.1145/2614066.2614101

Riecke, B. E., Bodenheimer, B., McNamara, T. P., Williams, B., Peng, P., and Feuereissen, D. (2010). "Do we Need to Walk for Effective Virtual Reality Navigation? Physical Rotations Alone May Suffice," in International Conference on Spatial Cognition (Springer), 234–247. doi:10.1007/978-3-642-14749-4_21

Riecke, B. E. (2008). Consistent Left-Right Reversals for Visual Path Integration in Virtual Reality: More Than a Failure to Update One's Heading?. *Presence: Teleoperators and Virtual Environments*. 17, 143–175. doi:10.1162/pres.17.2.143

Riecke, B. E., Cunningham, D. W., and Bülthoff, H. H. (2007). Spatial Updating in Virtual Reality: the Sufficiency of Visual Information. *Psychol. Res.* 71, 298–313. doi:10.1007/s00426-006-0085-z

Riecke, B. E., Feuereissen, D., Rieser, J. J., and McNamara, T. P. (2015). More Than a Cool Illusion? Functional Significance of Self-Motion Illusion (Circular Vection) for Perspective Switches. *Front. Psychol.* 6, 1174. doi:10.3389/fpsyg.2015.01174

Riecke, B. E., and Feuereissen, D. (2012). "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, 17–24.

Riecke, B. E. (2003). *How Far Can We Get with Just Visual Information? Path Integration and Spatial Updating Studies in Virtual Reality of MPI Series in Biological Cybernetics*. Berlin: Logos, Vol. 8.

Riecke, B. E., and Jordan, J. D. (2015). Comparing the Effectiveness of Different Displays in Enhancing Illusions of Self-Movement (Vection). *Front. Psychol.* 6, 713. doi:10.3389/fpsyg.2015.00713

Riecke, B. E. (2006). "Simple User-Generated Motion Cueing Can Enhance Self-Motion Perception (Vection) in Virtual Reality," in Proceedings of the ACM symposium on Virtual reality software and technology (ACM)), 104–107. doi:10.1145/1180495.1180517

Riecke, B. E., and Zielasko, D. (2020). "Towards an Affordance of Embodied Locomotion Interfaces in VR: How to Know How to Move?," in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (Atlanta, GA, USA: IEEE), 295–296. doi:10.1109/VRW50115.2020.00066

Rieser, J. J. (1989). Access to Knowledge of Spatial Structure at Novel Points of Observation. *J. Exp. Psychol. Learn. Mem. Cogn.* 15, 1157–1165. doi:10.1037/0278-7393.15.6.1157

Rognon, C., Mintchev, S., DellAgnola, F., Cherpillod, A., Atienza, D., and Floreano, D. (2018). Flyjacket: An Upper Body Soft Exoskeleton for Immersive Drone Control. *IEEE Robot. Autom. Lett.* 3, 2362–2369. doi:10.1109/lra.2018.2810955

Ruddle, R. A., and Jones, D. M. (2001). Movement in Cluttered Virtual Environments. *Presence: Teleoperators & Virtual Environments.* 10, 511–524. doi:10.1162/105474601753132687

Ruddle, R. A., and Lessels, S. (2006). For Efficient Navigational Search, Humans Require Full Physical Movement, but Not a Rich Visual Scene. *Psychol. Sci.* 17, 460–465. doi:10.1111/j.1467-9280.2006.01728.x

Ruddle, R. A., and Lessels, S. (2009). The Benefits of Using a Walking Interface to Navigate Virtual Environments. *ACM Trans. Computer-Human Interaction (Tochi).* 16, 5. doi:10.1145/1502800.1502805

Ruddle, R. A. (2005). "The Effect of Trails on First-Time and Subsequent Navigation in a Virtual Environment," in IEEE Proceedings. VR 2005. Virtual Reality, 2005 (IEEE), 115–122.

Ruddle, R. A. (2013). "The Effect of Translational and Rotational Body-Based Information on Navigation," in *Human Walking in Virtual Environments*. Editors F. Steinicke, Y. Visell, J. Campos, and A. Lécuyer (New York: Springer), 99–112. doi:10.1007/978-1-4419-8432-6_5

Ruddle, R. A., Volkova, E., and Bülthoff, H. H. (2011a). Walking Improves Your Cognitive Map in Environments That Are Large-Scale and Large in Extent. *ACM Trans. Comput.-Hum. Interact.* 18, 1–20. doi:10.1145/1970378.1970384

Ruddle, R. A., Volkova, E., Mohler, B., and Bülthoff, H. H. (2011b). The Effect of Landmark and Body-Based Sensory Information on Route Knowledge. *Mem. Cogn.* 39, 686–699. doi:10.3758/s13421-010-0054-z

Schulte, B., Bimberg, C. P., Hertel, J., Neu, F., Heidenreich, P., Bruder, G., et al. (2016). "Evaluation of Two Leaning-Based Control Methods in a Dragon Riding Interface," in Proceedings of the GI-Workshop VR/AR, 109–116.

Shelton, A. L., and McNamara, T. P. (1997). Multiple Views of Spatial Memory. *Psychon. Bull. Rev.* 4, 102–106. doi:10.3758/bf03210780

Smart, L. J., Jr, Stoffregen, T. A., and Bardy, B. G. (2002). Visually Induced Motion Sickness Predicted by Postural Instability. *Hum. Factors.* 44, 451–465. doi:10.1518/0018720024497745

Steinicke, F., Vissell, Y., Campos, J. L., and Lecuyer, A. (2013). *Human Walking in Virtual Environments*. Berlin: Springer.

Stoffregen, T. A., and Smart, L. J., Jr (1998). Postural Instability Precedes Motion Sickness. *Brain Res. Bull.* 47, 437–448. doi:10.1016/s0361-9230(98)00102-6

Sun, H.-J., Campos, J. L., and Chan, G. S. W. (2004). Multisensory Integration in the Estimation of Relative Path Length. *Exp. Brain Res.* 154, 246–254. doi:10.1007/s00221-003-1652-9

Waller, D., Loomis, J. M., and Haun, D. B. M. (2004). Body-Based Senses Enhance Knowledge of Directions in Large-Scale Environments. *Psychon. Bull. Rev.* 11, 157–163. doi:10.3758/bf03206476

Wang, R. F. (2004). Between Reality and Imagination: When Is Spatial Updating Automatic? *Perception & Psychophysics.* 66, 68–76. doi:10.3758/BF03194862

Wang, R. F. (2016). Building a Cognitive Map by Assembling Multiple Path Integration Systems. *Psychon. Bull. Rev.* 23, 692–702. doi:10.3758/s13423-015-0952-y

Wang, R. F., and Spelke, E. S. (2002). Human Spatial Representation: Insights From Animals. *Trends Cogn. Sci.* 6, 376–382. doi:10.1016/s1364-6613(02)01961-7

Xia, X., Pun, C.-M., Zhang, D., Yang, Y., Lu, H., Gao, H., et al. (2019). "A 6-dof Telexistence Drone Controlled by a Head Mounted Display," in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (IEEE), 1241–1242. doi:10.1109/vr.2019.8797791

Zhai, S. (1998). User Performance in Relation to 3d Input Device Design. *SIGGRAPH Comput. Graph.* 32, 50–54. doi:10.1145/307710.307728

Zhang, H., Mou, W., and McNamara, T. P. (2011). Spatial Updating According to a Fixed Reference Direction of a Briefly Viewed Layout. *Cognition.* 119, 419–429. doi:10.1016/j.cognition.2011.02.006

Zielasko, D., and Riecke, B. E. (2020). "Sitting vs. Standing in VR: Towards a Systematic Classification of Challenges and (Dis)Advantages," in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and

Workshops (VRW) (Atlanta, GA, USA: IEEE), 297–298. doi:10.1109/VRW50115.2020.00067

Conflict of Interest: TN-V was employed by Microsoft Corporation; MH was employed by VdH-IT.

The remaining 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: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Virtual Reality

ORIGINAL RESEARCH
published: 05 April 2022
doi: 10.3389/frvir.2022.800884

Sensation of Anteroposterior and Lateral Body Tilt Induced by Electrical Stimulation of Ankle Tendons

Nozomi Takahashi[1]*, *Tomohiro Amemiya*[1,2], *Takuji Narumi*[1], *Hideaki Kuzuoka*[1,2], *Michitaka Hirose*[3] and *Kazuma Aoyama*[2,3]*

[1]*The Graduate School of Information Science and Technology, The University of Tokyo, Bunkyo, Japan,* [2]*Virtual Reality Educational Research Center, The University of Tokyo, Tokyo, Japan,* [3]*Research Center for Advanced Science and Technology, The University of Tokyo, Tokyo, Japan*

While virtual reality technology enables users to walk on surfaces of various shapes in virtual environments, these experiences are often on a flat floor, and any discordance between visual and somatosensory information hampers the immersive experience. To resolve this issue, we have focused on the ankle joint angle as an essential cue in detecting the surface shape of the ground. To modulate the sensation of the ankle joint angle, we propose novel percutaneous electrical stimulation methods that stimulate four targeted ankle tendons: the tibialis anterior muscle tendon, the Achilles tendon, the peroneus longus muscle tendon, and the flexor digitorum longus tendon. Since electrically stimulating the elbow tendon is known to evoke reflexive hand movement, electrically stimulating the ankle tendon is expected to evoke a body tilt due to illusory changes in the ankle joint angle. In this study, we designed electrode configurations to stimulate the above four ankle tendons using a finite element analysis and investigated the effect of electrically stimulating the ankle tendons on the subjective sensation of body tilt and actual body sway through psychophysical experiments. The results revealed that applying this stimulation with our novel electrode configurations can induce a subjective sensation of body tilt and actual body sway in a direction opposite to the stimulated part.

Keywords: new haptic and tactile interaction, tendon electrical stimulation, proprioceptive sensation, golgi tendon organ, illusory movement

OPEN ACCESS

Edited by:
Thierry Duval,
IMT Atlantique Bretagne-Pays de la Loire, France

Reviewed by:
Vincent Levesque,
École de Technologie Supérieure (ÉTS), Canada
Hojin Lee,
Electronics and Telecommunications Research Institute (ETRI), South Korea

Correspondence:
Nozomi Takahashi
nozomi@cyber.t.u-tokyo.ac.jp
Kazuma Aoyama
aoyama@vr.u-tokyo.ac.jp

Specialty section:
This article was submitted to Technologies for VR, a section of the journal Frontiers in Virtual Reality

Received: *24 October 2021*
Accepted: *10 March 2022*
Published: *05 April 2022*

Citation:
Takahashi N, Amemiya T, Narumi T, Kuzuoka H, Hirose M and Aoyama K (2022) Sensation of Anteroposterior and Lateral Body Tilt Induced by Electrical Stimulation of Ankle Tendons. Front. Virtual Real. 3:800884. doi: 10.3389/frvir.2022.800884

INTRODUCTION

Owing to their recent price reduction, head-mounted displays (HMDs) are now more accessible to the public, subsequently increasing the popularity of virtual reality (VR) experiences. An HMD can simulate an enormous virtual environment. To take advantage of this feature, methods that help navigation through a virtual environment have been proposed (Boletsis, 2017). For example, a method that synchronizes body position provides an immersive experience by allowing users to move intuitively on their feet (Bruder and Steinicke, 2014). However, this method provides only visual information. When standing on sloped ground in the real world, ankle joint angles vary with ground tilt and induce somatosensory sensations. In comparison, VR activities are typically experienced on flat indoor floor, in which case the somatosensory and visual sensations may be inconsistent.

Such an inconsistency hampers the immersive experience. To resolve this issue, many studies have been conducted to present somatosensory sensations induced by the surface shape of the ground in a virtual environment. A popular approach is to use mechanical devices that deform shapes in a real environment with actuators to reproduce the surface shape of the ground in a virtual environment. For example, Schmidt et al. (2015) developed level-ups, a boot-shaped device whose height from the floor can be controlled by actuators. With this device, a discrete change in the ground height in a virtual environment can be replicated. In addition, Je et al. (2021) proposed Elevate, a deformable platform that users can walk on, that reproduces the continuous changes in the shape of the ground in the virtual environment using a pin array. Although these devices present an accurate ground surface shape by controlling the actuators, they are typically expensive, large, and heavy because they require complex structures. In addition, physically changing the shape of the ground around the feet during the use of HMDs can be dangerous.

In contrast to methods that modulate the physical environment using actuators, there is another approach that can induce a sensation in the ankle joint that corresponds to a standing on a sloped terrain in a virtual environment. While standing on a slope in a physical environment, the angle of the ankle joint changes by the same amount as the tilt of the ground to keep the body parallel to gravity (Sasagawa et al., 2009). Considering this fact, presenting a virtual sensation of the angle of the ankle joint may change users' perception of ground tilt without modulating the physical environment.

For presentation of a virtual sensation of an angle in the ankle joint, the illusory sensation of motion induced by tendon vibration is widely acknowledged. Goodwin et al. (1972) found that applying a vibration of 100 Hz to the tendons of the biceps or triceps increases the error of the estimated angle of the elbow joint. This phenomenon is presumed to be due to the activation of the muscle spindle (Seki et al., 2011). In some recent studies, tendon vibration has been utilised in VR experiences to present a sensation of body movement, with only small motion or even without any actual motion occurring (Hagimori et al., 2019; Tanaka et al., 2020). However, if tendon vibration is applied to ankle joints, the mechanical vibrators must be fixed to the tendons around the ankle, making it difficult for users to move smoothly in a standing posture. In addition, to strongly stimulate the tendon using vibration, large and heavy vibrators are required. To present a virtual sensation of an angle in the ankle joint without having to install a bulky physical device around the feet, we focused on transcutaneous electrical stimulation, which can induce various sensations by applying a weak current through surface electrodes attached to the skin. This technique has recently been utilised in the field of human–computer interaction and VR as a more inexpensive and simpler way of inducing sensation with a lighter and smaller device (Tamaki et al., 2016; Aoyama et al., 2017). Moreover, the extent of the sensation can be controlled by the intensity of applied current.

To present joint sensation, Kajimoto (2013) applied tendon electrical stimulation (TES) to the outside of the elbow joint and noted an illusory inward movement of the forearm. The author also suggested that the Golgi tendon organ (GTO), which detects muscle tension (Moore, 1984), was involved in this illusory movement. Based on this result, we hypothesise that applying TES to the tendons around the ankle (ankle TES) could induce an illusory sensation of an angle in the ankle joint, leading to a change in the perception of ground tilt. Moreover, the GTO is known to cause a Golgi tendon reflex, which relaxes the agonistic muscle and contracts the antagonistic muscle when it detects continuous muscle tension (Moore, 1984). This suggests that ankle TES would also induce the Golgi tendon reflex and cause an actual postural change while standing.

To investigate whether applying a current on the tendons around the ankle can stimulate the tendons, we previously conducted a psychophysical experiment that investigated the subjective intensity and direction of force sensation around the ankle and recorded the temporal pattern of the centre of pressure (CoP) (Takahashi et al., 2021). Note that the other research group has reported that the force sensation can be induced by TES (Takahashi et al., 2016). While the result showed that applying a current on the tibialis anterior muscle tendon and Achilles tendon can induce force sensation in a direction opposite to the stimulated part, which indicates that these tendons were stimulated by the current, the experiment had some limitations. First, we simply placed the electrodes on the targeted tendons; however, whether these electrodes stimulated the targeted tendons independently was unclear. Second, only limited tendons were selected as the target. Both the tibialis anterior muscle tendon and the Achilles tendon are responsible for the anteroposterior movement of the ankle, while the ankle can also be moved in a lateral direction by other tendons. Third, the result of the CoP was unreliable because of the inadequate process of recording the CoP; the measurement under a specific condition was conducted only once, and a baseline correction of obtained data was conducted based on the single data obtained at the beginning of the measurement.

In this study, we selected four targeted tendons according to the classification of ankle joint movement and developed novel stimulation configurations, based on the simulation of current density, to stimulate the targeted tendon under the electrodes most strongly among four targeted tendons. With these stimulation configurations, we evaluated the effect of ankle TES on the sensation of the angle of the ankle joint and the actual body by conducting psychophysical experiments with more elaborated methods.

STIMULATION DESIGN AND SIMULATION

Electrode Configuration Design

There are six main types of ankle joint movements: dorsiflexion, plantarflexion, inversion, eversion, internal rotation, and external rotation (Saltzman, 1997). Among these movements, dorsiflexion, plantarflexion, inversion, and eversion, which are movements in the sagittal and coronal planes, are considered to be involved in the detection of ground tilt. Therefore, ankle TES should be applied to tendons that are responsible for these movements to present a sensation of ground tilt in all directions. In our study, we selected the tibialis anterior

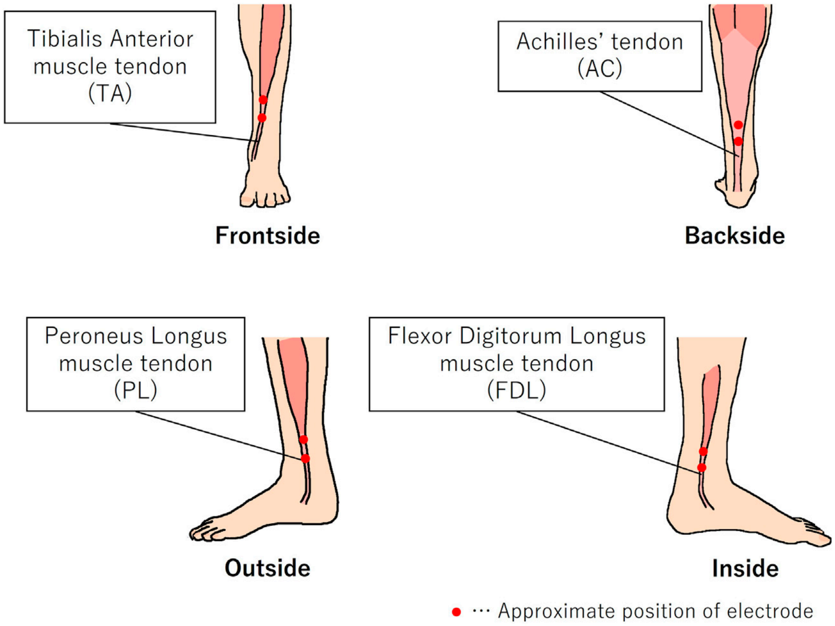


FIGURE 1 | Anatomical positions of the targeted tendons in the left leg. The red dots show the approximate positions of the electrodes placed for electrical stimulation.

muscle tendon (TA), Achilles tendon (AC), flexor digitorum longus muscle tendon (FDL), and peroneus longus muscle tendon (PL) as the stimulation targets. **Figure 1** shows the anatomical positions of the tendons. The TA is at the front side of the ankle and is responsible for dorsiflexion of the foot. The AC is at the back side of the ankle and is responsible for plantarflexion of the foot. The FDL is inside the ankle and is responsible for foot inversion. The PL is outside the ankle and is responsible for foot eversion.

To stimulate these tendons, electrodes were placed above them (red dots in **Figure 1**). Though it is desirable that these tendons are stimulated independently to control the direction of the presented sensation of ground tilt, these tendons are located close to each other. Hence, it is expected that the current applied on a targeted tendon also affects other tendons. However, if the tendon under the electrodes is stimulated most strongly among the four targeted tendons, the stimulation of the tendon under the electrodes is expected to have the largest effect on the sensation of an ankle joint angle. To confirm that each electrode configuration can stimulate the targeted tendon most strongly among the four targeted tendons, the current density distributions of the tendons, produced by the current applied by each electrode configuration, were simulated using the finite element method with a 3D model of the human leg.

Pre-Processing of 3D Model

A solid 3D male model (Zygote Media Group Inc.) was used for the simulation. This model was constructed based on magnetic resonance imaging and computed tomography and was consistent with anatomical knowledge. The model was imported into Scan IP (Simpleware, SYNOPSYS Inc.), and all parts of the model, except for the left leg, were cut out. Subsequently, the elements of the model were classified into six groups: bone, cartilage, nerve, vessel, skin, and muscle. In this process, the tendon was classified as muscle because the tendon and muscle are inseparable in the model. Since their conductivities are almost the same (tendon: 0.30 S/m, muscle: 0.35 S/m) (Laakso and Hirata, 2013), the integration is expected to have little effect on the result of the simulation of the distribution of current density. Moreover, there are some unnatural gaps in the model because the model does not contain blood or inner tissue. These gaps do not exist in the actual human body and could possibly make the result of the simulation inaccurate by preventing the propagation of the simulated current. Hence, these gaps were filled by the Boolean operation function of Scan IP. Thereafter, the filled regions inside the vessels were classified as blood, and the remaining regions were classified as inner tissue. Four pairs of electrodes were manually attached to the skin model (eight electrodes in total). On each targeted tendon, the upper and lower electrodes were treated as the anode and cathode, respectively. **Figure 2** shows an overview of the processed model and the position of the attached electrodes in the model.

Simulation of Current Density

The processed model was converted into NASTRAN format and imported into COMSOL Multiphysics 5.6 (COMSOL Inc.). Before the simulation, the conductivity values for each tissue were set, as listed in **Table 1**. These values were defined based on a previous study conducted by Laakso and Hirata (2013). The Laplace equation $\nabla \cdot (\sigma \nabla V) = 0$ (where V is the electrical potential, and σ is the conductivity) was solved by applying the following boundary conditions:

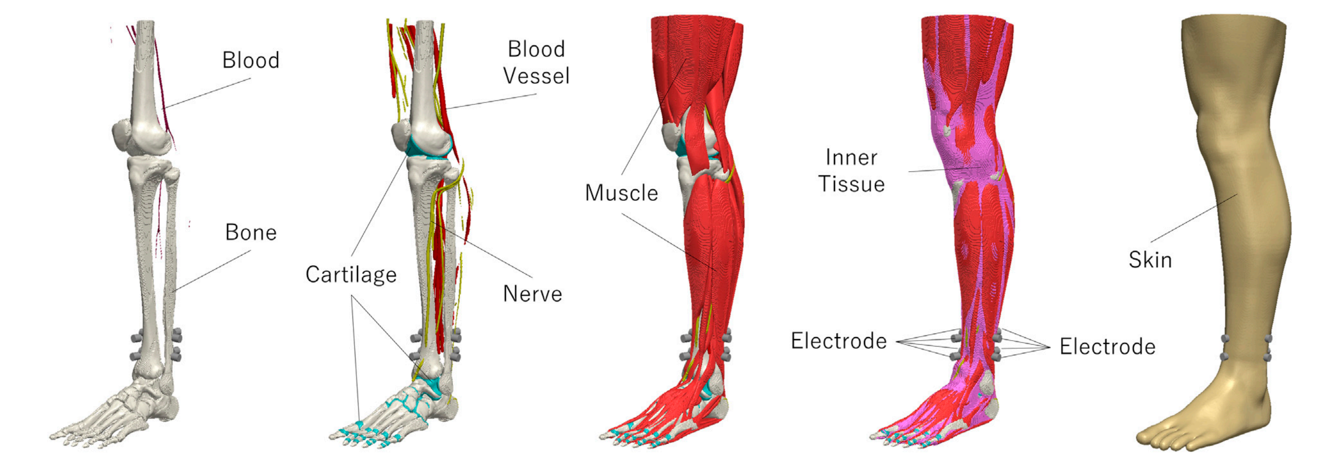


FIGURE 2 | Overview of the processed model of the left leg and position of electrodes.

TABLE 1 | Conductivity of each tissue in the model. S (siemens) is the SI derived unit of electric conductance.

Tissue type	Conductivity (S/m)
Bone	0.02
Cartilage	0.18
Nerve	0.03
Vessel	0.04
Skin	0.1
Muscle	0.35
Blood	0.7
Inner tissue	0.14
Electrode	0.3

1) The inward current density, *Jn,* was applied to the exposed surface of the anode (upper electrode).
2) A ground was applied to the exposed surface of the cathode (lower electrode).
3) All other external surfaces were treated as insulated.
4) The inward current density (*Jn*) for each anode was defined to adjust the value of the current to 4.0 mA, considering the superficial size of each electrode.

The simulation was repeated for the four target tendons: TA, AC, PL, and FDL.

Simulation Result

Figure 3 shows the simulated current density distribution of the muscles in each simulation, where red indicates a high current density (higher than 10 A/m^2), while blue indicates a low current density. **Table 2** shows the maximum current density distribution of the four targeted tendons in each simulation. As we expected, the table shows that each stimulation changes the current density distribution not only of the tendon under the electrodes but also of the other targeted tendons. However, the maximum current density distribution of the tendons under the electrodes is at least three times higher, and 14 times higher on average, than that of the other targeted tendons. Whether the extent of the effect of stimulation of targeted tendons other than that under the electrodes is unclear, these results indicate that the four proposed types of ankle TES configurations can stimulate the targeted tendon under the electrodes most strongly and the stimulation of the tendon is considered to have the largest effect on the sensation of an ankle joint angle. Hence, we adopted these electrode configurations to apply current in the following experiment.

MATERIAL AND METHODS FOR PSYCHOPHYSICAL EXPERIMENT

The purpose of this study was to demonstrate the effect of the simulated ankle TES configurations on the sensation of an ankle joint angle. A psychophysical experiment was conducted to achieve this purpose. For the psychophysical experiment, we recruited participants on whom five experimental conditions were applied, and a subjective sensation of body tilt and objective body sway were measured. The experimental protocol was approved by the local ethics research committee of the Graduate School of Information Science and Technology, University of Tokyo. Twelve healthy adults (11 males and 1 female) participated in the experiment, with an average age of 23.1 years.

Hypothesis

Kajimoto (2013) reported that applying TES to the outside of the elbow induces an illusory inward movement of the forearm, suggesting that TES induces illusory movement in a direction opposite to the stimulated part. In addition, if the GTO is assumed to be stimulated by TES, the Golgi tendon reflex, which relaxes the agonistic muscle and contracts the antagonistic muscle, might be induced, making participants unconsciously lean in the opposite direction to the stimulated body part owing to the reflex. Therefore, we have formulated the following hypotheses regarding the effect of our novel ankle TES configuration:

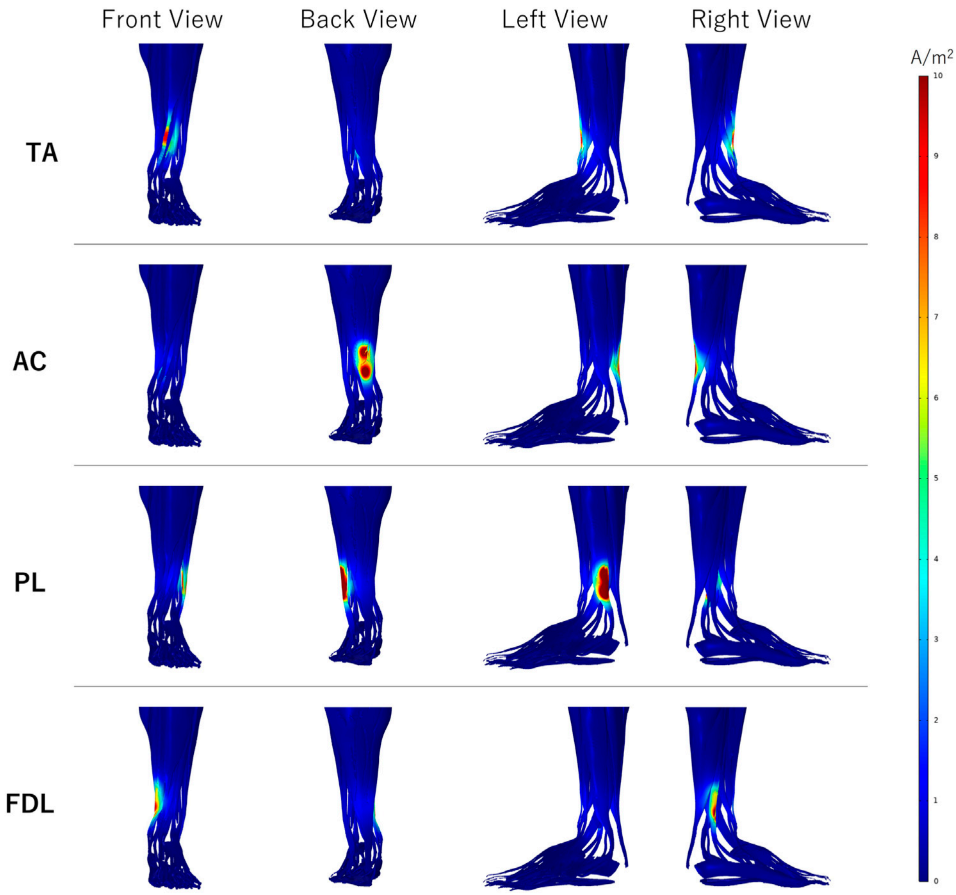


FIGURE 3 | Distributions of the current density in the simulations.

TABLE 2 | Maximum current density distribution on each targeted tendon in each simulation.

Target of simulation	Maximum current density on TA (A/m^2)	Maximum current density on AC (A/m^2)	Maximum current density on PL (A/m^2)	Maximum current density on FDL (A/m^2)
TA	10.007	0.9535	1.1052	1.2513
AC	0.92907	23.247	1.1397	1.771
PL	1.0612	1.5966	22.291	1.1751
FDL	3.3302	1.7378	0.83064	12.045

H1. Ankle TES induces a subjective sensation of body tilt solely in the direction opposite to the stimulated part.
H2. Ankle TES evokes body sway solely in the direction opposite to the stimulated part.

Experimental Conditions

Five experimental conditions were used in the experiment (**Table 3**). The front side (FS) condition stimulates the TA of both feet. The back side (BS) condition stimulates the AC of both feet. The left side (LS) condition stimulates the PL of the left foot and the FDL of the right foot. The right side (RS) condition stimulates the FDL of the left foot and PL of the right foot. The no-stimulation (NS) condition is a baseline condition in which no tendon is stimulated. Each experimental condition was repeated 10 times. **Figure 4** shows the position of the electrodes used under all conditions where current was applied. These electrodes were manually placed on each targeted tendon referring to the simulated electrodes configurations to stimulate each targeted tendon as independently as possible. Under all conditions, except for the NS condition, a bipolar square wave was used as the waveform of the current, and the frequency of the current was set to 80 Hz based on previous research that investigated the relationship between the current frequency and the intensity of force sensation induced by TES (Takahashi et al., 2016).

TABLE 3 | Experimental conditions.

Name of condition	Targeted tendon on the left foot	Targeted tendon on the right foot
Front side (FS)	TA	TA
Back side (BS)	AC	AC
Left side (LS)	PL	FDL
Right side (RS)	FDL	PL
No stimulation (NS)	No tendon	No tendon

The amplitude of the current wave was adjusted to the maximum value within the participant's uncomfortable sensation threshold for each stimulus position to make the effect of the stimulation clearer. At the beginning of the adjustment, the participant was asked to remove their shoes and socks and stand with their feet together, which was the same posture as that during experiment. Thereafter, 80 Hz square wave current was applied to the stimulus position. The amplitude of the current wave was increased in steps of 0.5 mA from 0 mA until the participant reported any uncomfortable sensations such as strong tactile sensation or vibration. Thereafter, the amplitude of the current wave was decreased in steps of 0.1 mA. When the participant reported that the uncomfortable sensation disappeared, the amplitude of the current wave at that time was decided as the value used in the experiment. The adjustment was repeated for each stimulus position. If the threshold of uncomfortable sensation exceeded 4 mA, the amplitude of the current was selected as 4 mA because the electrical stimulator used in the experiment could not generate larger than 4 mA of current. The value was determined by our pilot study that investigated the threshold of uncomfortable sensation with the same adjustment procedure as stated above.

Data Collection and Analysis for Subjective Sensation

To evaluate the effect of ankle TES on the sensation of the ankle joint angle, participants were asked to fill in a questionnaire regarding their subjective sensation of body tilt during stimulation. The following questions were posed:

Q1. Did you feel a sensation of lateral body tilt? (1: strong sensation of leftward body tilt, 4: no sensation of lateral body tilt, 7: strong sensation of rightward body tilt).
Q2. Did you feel a sensation of anteroposterior body tilt? (1: strong sensation of forward body tilt, 4: no sensation of anteroposterior body tilt, 7: strong sensation of backward body tilt).

Each question was answered by choosing a number from one to seven. For each question, 10 answers under the same experimental condition were averaged and treated as a representative score for each condition.

Data Collection and Analysis for Actual Body Sway

Based on the role of the GTO, actual body sway was expected to be induced by ankle TES. However, it was difficult to measure

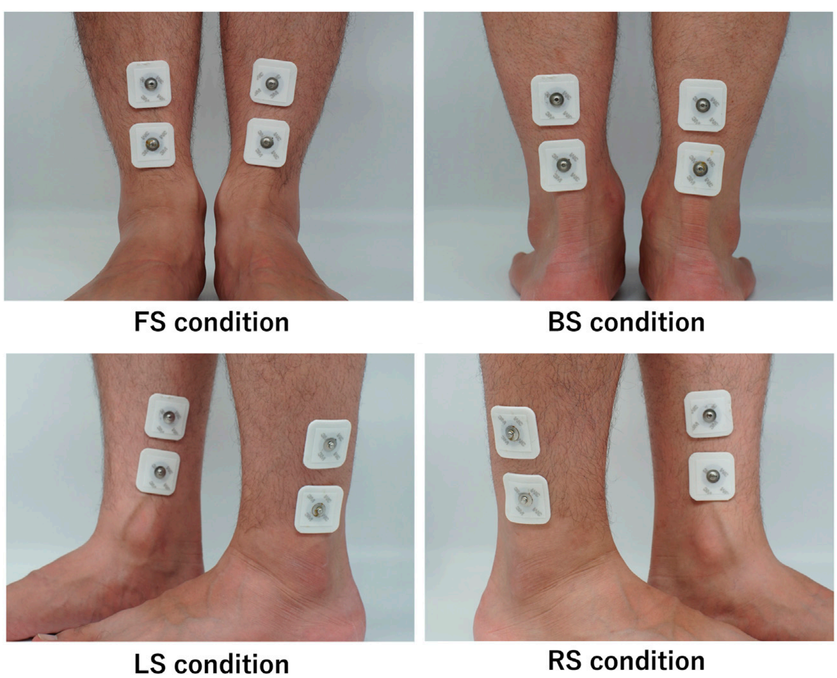


FIGURE 4 | Position of electrodes.

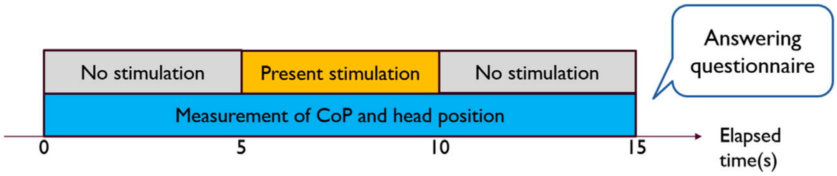


FIGURE 5 | Time sequence of tendon electrical stimulation and measurement of the centre of pressure (CoP) and head position.

body sway because the movement of every body part was involved. Therefore, we measured the CoP and head position as representative values of body sway. The Wii Balance Board (Nintendo Co., Ltd.) was used to measure the 2D position of the CoP, and Polhemus G4 (Polhemus, Inc.) was used to measure the 3D position of the head by attaching a sensor on top of the participant's head. In both devices, the rightward, forward, and upward directions of the participants were set as the positive direction of the lateral, anteroposterior, and vertical axes, respectively. In addition, these values were measured for 15 s at a frequency of 10 Hz. Therefore, 150 data points for the CoP and head position were obtained in a single measurement. In very rare cases, the data could not be stored on the PC correctly. In such cases, the broken data was replaced by the average of the data before and after the damaged data.

In each measurement, the initial position of the CoP was calculated by averaging 50 CoP data points from the first 5s of measurement. Baseline correction was then applied by subtracting the initial position from the entire CoP dataset. Thereafter, 10 corrected trial data under the same experimental condition were averaged and treated as representative data for that condition. The same process was applied to the head-position data. We estimated a latency of stimulus onset (offset) from the earliest (latest) time when the displacement of CoP kept exceeding the mean ± SD of the CoP under the NS condition.

Procedure

The experiment involved 50 trials in total (five experimental conditions × 10 repetitions). At the beginning of the experiment, the amplitude of the current applied in each experimental condition was adjusted according to the procedure described in *Experimental Conditions*. After the adjustment, these trials were conducted in a randomised order. In each trial, the CoP and head position were measured for 15 s, as shown in **Figure 5**. During the measurement, participants were asked to remove their shoes and socks and stand on the Wii Balance Board with their feet together, known as the Romberg's erect posture, with their eyes closed (Rogers, 1980). The ankle TES under one of the experimental conditions was applied for 5 s at a time of 5 s after the start of the trial. After the measurement, the participants were allowed to open their eyes and answer the questionnaire on a tablet (Surface Pro 7, Microsoft Corp.). To eliminate the influence of fatigue, the participants could rest at any time.

Statistical Analysis

The results of the questionnaire regarding the intensity of the subjective sensation of body tilt, the largest displacement of CoP, and the largest displacement of the head position were analysed by a many-to-one multiple comparison test, which compares a control condition (baseline, i.e., NS condition) with other conditions (FS, BS, LS, and RS conditions). We did not conduct a comparison of all possible pairs of conditions because our hypotheses mainly focus on the effect of ankle TES compared with a control condition (NS condition). To decide the method of a many-to-one multiple comparison test, the Shapiro-Wilk normality test was initially conducted to determine whether the data don't follow a normal distribution. If the assumption of normality was not violated in each condition, Dunnett's method, a parametric method of multiple comparison test, was conducted. Otherwise, Steel's method, a nonparametric version of Dunnett's method (Steel, 1959), was conducted. We did not conduct ANOVA F-test before either method of multiple comparison, following previous research (Hothorn, 2016).

RESULT OF PSYCHOPHYSICAL EXPERIMENT

The Amplitude of the Current Wave

The amplitude of the current wave of TES was determined for each foot and participant. The averaged amplitudes on a left foot in the FS, BS, LS, and RS conditions across participants were 3.17 ± 0.83, 2.82 ± 0.90, 3.00 ± 1.08, and 2.79 ± 0.87 mA, respectively. In contrast, the averaged amplitudes on a right foot in the FS, BS, LS, and RS conditions were 3.24 ± 0.99, 2.83 ± 1.08, 2.95 ± 1.13, and 3.04 ± 1.12 mA, respectively. Two-factor ANOVA was conducted on the amplitude of the current wave, with foot type (left and right) and experimental condition (FS, BS, LS, and RS) used as the independent factors. The result of ANOVA did not show any significant effect or interaction effect.

Questionnaire

Figure 6 shows the average questionnaire scores regarding the intensity of the subjective sensation of body tilt. The result of the Shapiro-Wilk normality test showed that the result of Q1in the NS, FS, and BS conditions and the result of Q2 in the NS, LS, and RS conditions do not follow a normal distribution ($p < 0.05$). Therefore, statistical analysis was performed using Steel's method for nonparametric multiple comparison, with the NS condition as the control. The results showed that the average score for Q1 was significantly higher under the LS condition and lower under the RS condition than under the NS condition (LS vs. NS: $p = 0.0075$, RS vs. NS: $p = 0.0089$) (**Figure 6A**). Moreover, the average score

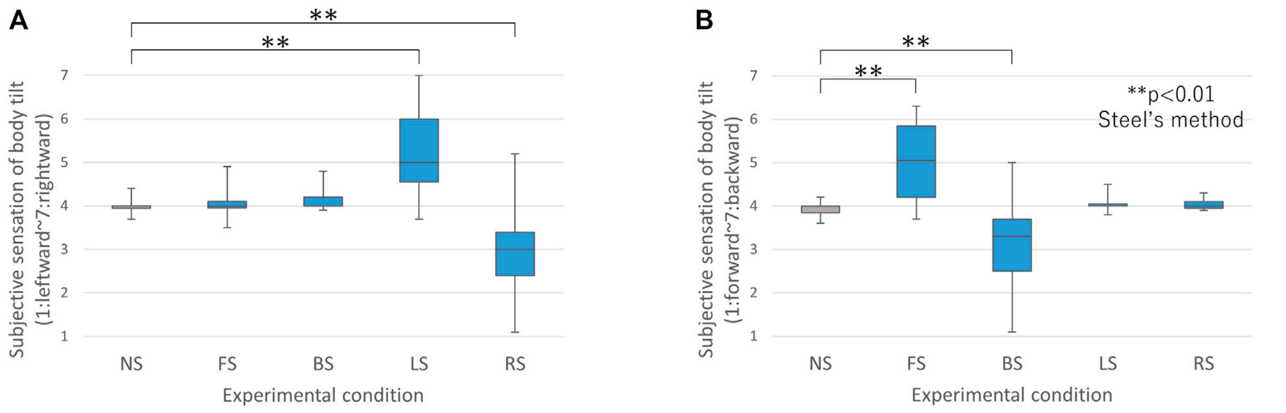


FIGURE 6 | Average of participants' answers on the questionnaire: **(A)** Result of the intensity of subjective sensation of lateral body tilt. 1 Indicates a strong sensation of leftward body tilt. 4 Indicates no sensation of lateral body tilt. 7 Indicates a strong sensation of rightward body tilt. **(B)** Result of the intensity of subjective sensation of anteroposterior body tilt. 1 Indicates a strong sensation of forward body tilt. 4 Indicates no sensation of anteroposterior body tilt. 7 Indicates a strong sensation of backward body tilt. (NS: No stimulation, FS: front side, BS: back side, LS: left side, or RS: right side).

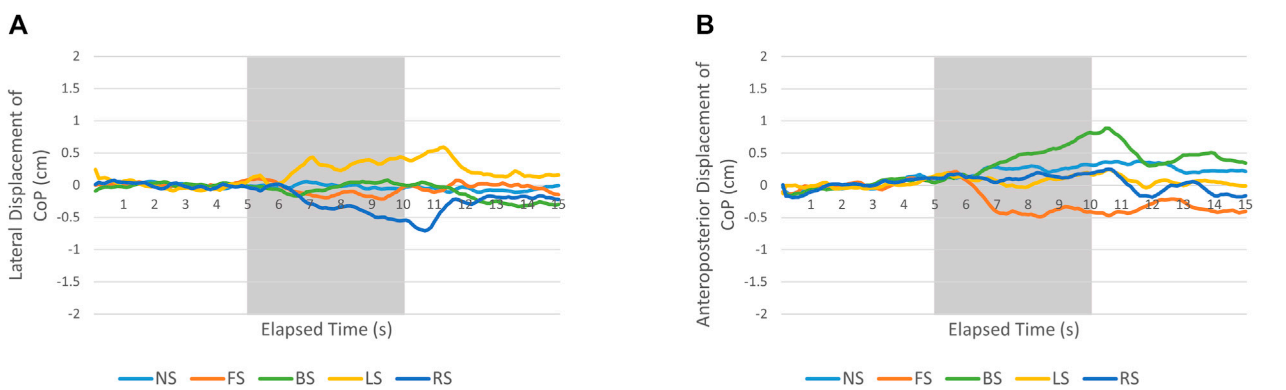


FIGURE 7 | Change in the CoP in the lateral direction **(A)** and anteroposterior direction **(B)** over time. The grey area indicates the interval in which electrical stimulation is applied. A positive lateral (anteroposterior) displacement indicates rightward (forward) motion. (NS: No stimulation, FS: front side, BS: back side, LS: left side, or RS: right side).

for Q2 was significantly higher under the FS condition and lower under the BS condition than under the NS condition (FS vs. NS: $p = 0.0011$, BS vs. NS: $p = 0.0029$) (**Figure 6B**).

CoP

Figure 7 shows the average temporal pattern of the CoP for each stimulation condition. In addition, the largest displacement of the CoP from the initial position during the presentation of stimulation in four directions (forward, backward, leftward, and rightward) was averaged among the participants (**Figure 8**). These figures show that the LS, RS, FS, and BS conditions evoked the largest rightward, leftward, backward, and forward CoP changes, respectively. Additionally, **Figure 8** shows the largest displacement of the CoP. The result of the Shapiro-Wilk normality test showed that the data of the largest forward displacement in the NS, FS, RS conditions, the largest leftward displacement in the FS and RS conditions and the largest rightward displacement in the NS, LS, and RS conditions do not follow a normal distribution ($p < 0.05$). Although the result did not show that the data of the largest backward displacement did not follow a normal distribution in each condition ($p > 0.05$), Steel's method was adopted in all directions to unify the method across all directions of CoP data. The results of the multiple comparison tests with Steel's method showed that the largest backward displacement of the CoP from the initial position was significantly greater under the FS and LS conditions than under the NS condition (FS vs. NS: $p = 0.0059$, LS vs. NS: $p = 0.035$) (**Figure 8B**). Moreover, marginally significant differences were found between the NS and BS conditions in the forward displacement, the NS and RS conditions in the leftward displacement, and the NS and LS conditions in the rightward displacement (BS vs. NS: $p = 0.0684$, RS vs. NS: $p = 0.0529$, LS vs. NS: $p = 0.0982$) (**Figures 8A,C,D**).

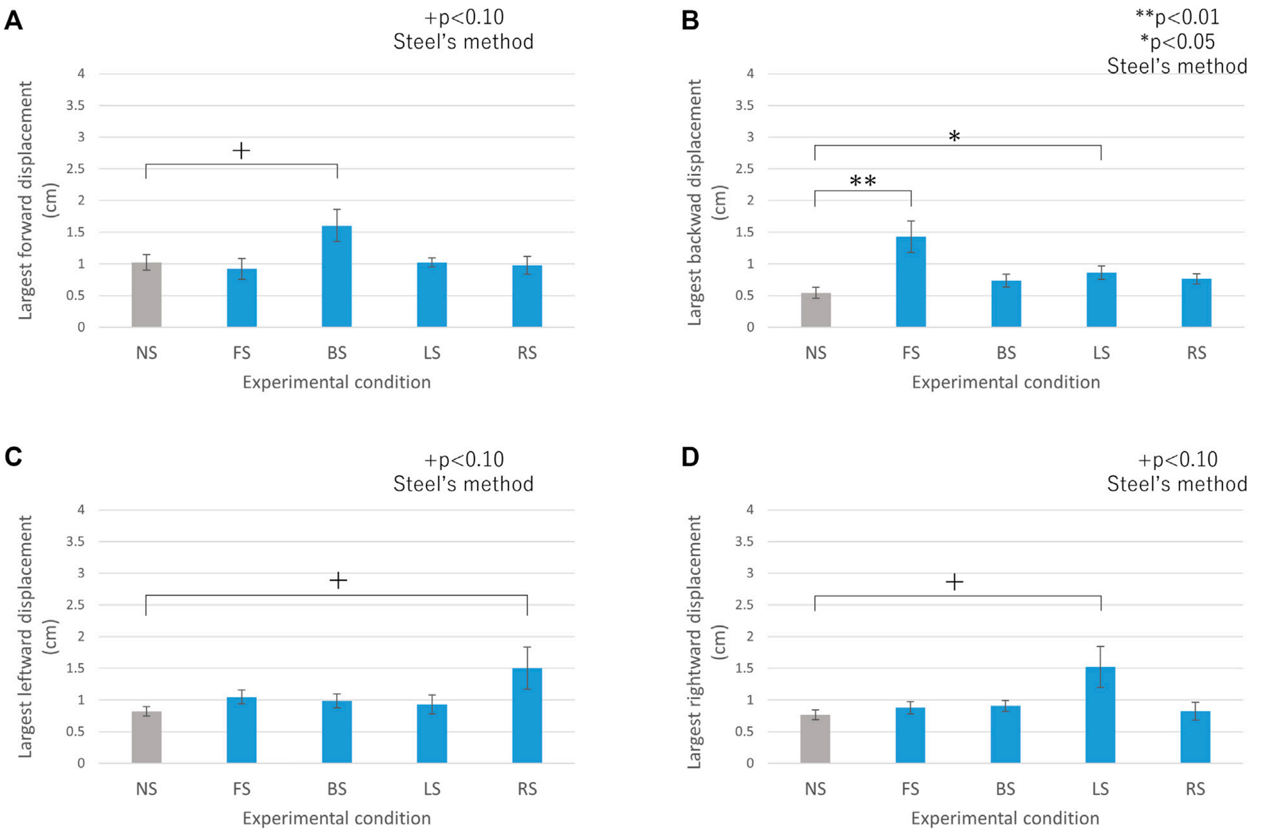


FIGURE 8 | Largest displacement of CoP from the initial position: **(A)** Forward displacement, **(B)** Backward displacement, **(C)** Leftward displacement, **(D)** Rightward displacement. (NS: No stimulation, FS: front side, BS: back side, LS: left side, or RS: right side).

TABLE 4 | Latency of the stimulus onset and offset of the displacement of CoP. "-" means that the latency is more than 5s.

Experimental condition	FS	BS	LS	RS
Direction of the displacement of CoP	Backward	Forward	Rightward	Leftward
Latency of the stimulus onset (s)	1.2	1.0	0.9	1.4
Latency of the stimulus offset (s)	—	1.9	—	—

In addition, the latencies of the stimulus onset and offset were calculated (**Table 4**). The former was defined as the earliest time during the stimulation when, afterwards, the displacement of the CoP exceeded the mean ± SD of the CoP in the NS condition for a duration of 1.3 s. The latter was defined as the earliest time after the stimulation when the displacement of the CoP became smaller than the mean ± SD of the CoP in the NS condition. The threshold of 1.3 s is half of the general period of body sway, which was calculated from the primary frequency peak of body sway reported in a previous study (Soames and Atha, 1982).

Head Position

The change in head position was averaged among the participants over time (**Figure 9**). In addition, the largest displacement of the head from its initial position during stimulation presentation in the four directions (forward, backward, leftward, and rightward) was averaged among the participants (**Figure 10**). These figures show that the LS and RS conditions evoked the largest rightward and leftward head position changes, respectively. The results of the Shapiro-Wilk normality test showed that the data of the largest forward displacement in the NS, FS, LS, and RS conditions, the largest backward displacement in the FS and LS conditions, the largest leftward displacement in the RS condition and the largest rightward displacement in the NS, LS, and RS conditions do not follow a normal distribution ($p < 0.05$). Therefore, statistical analysis was performed using Steel's method for a nonparametric multiple comparison with a control of the NS condition. The result showed that only the largest rightward displacement of the head from its initial position had a marginally significant difference between the NS and LS conditions ($p = 0.0774$) (**Figure 10D**).

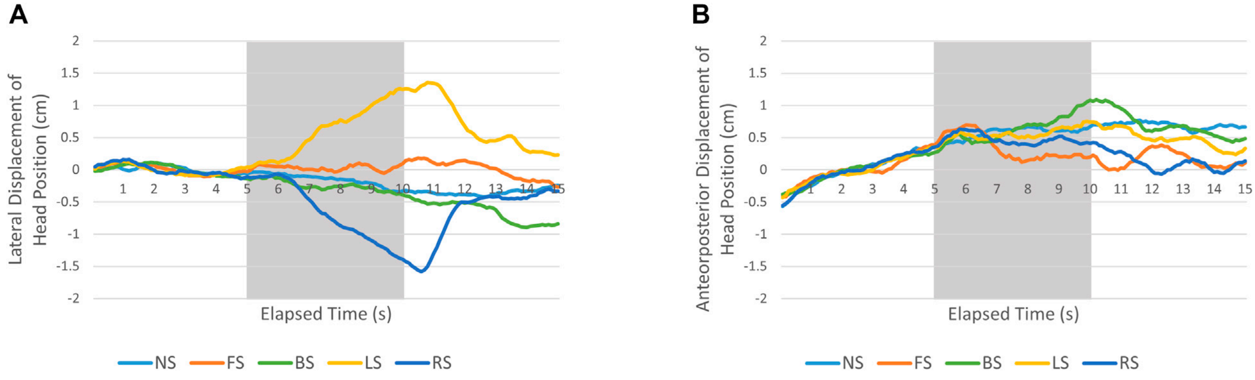


FIGURE 9 | Change in the head position in the lateral direction **(A)** and anteroposterior direction **(B)** over time. The grey area indicates the interval in which electrical stimulation is applied. (NS: No stimulation, FS: front side, BS: back side, LS: left side, or RS: right side).

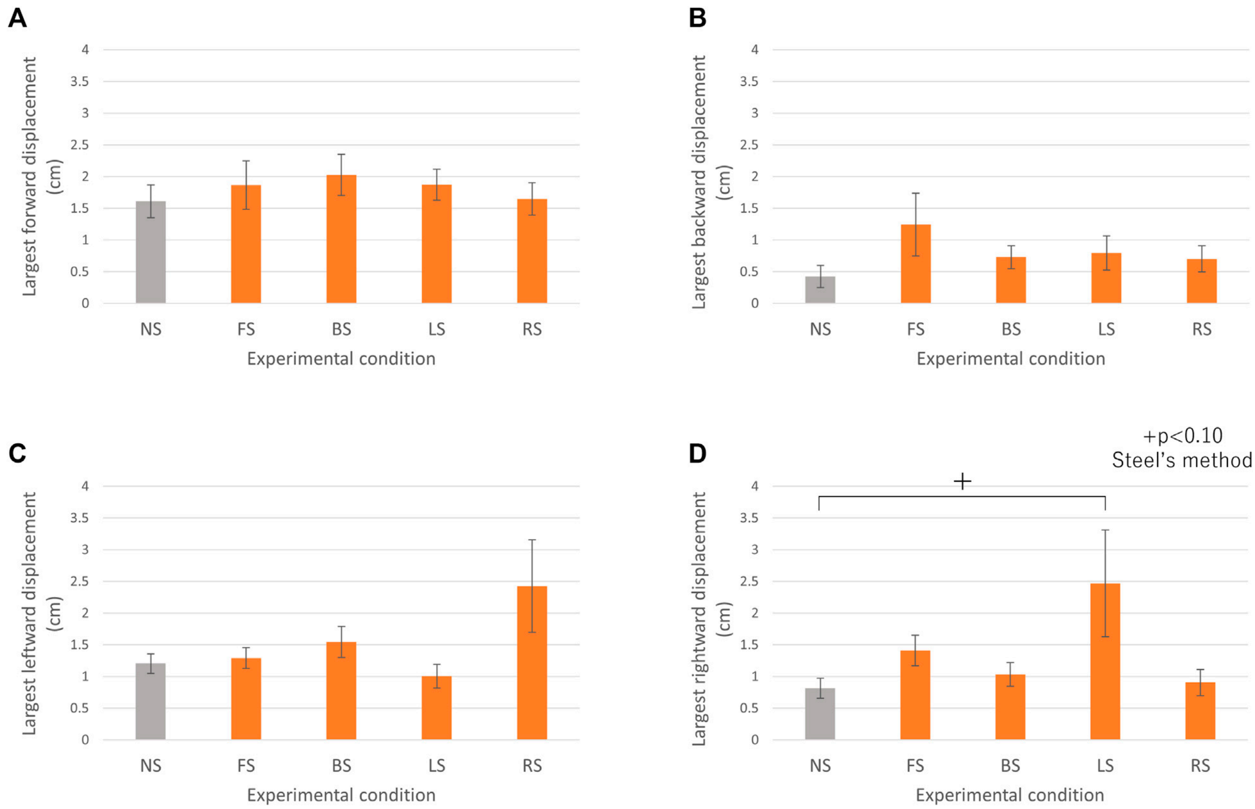


FIGURE 10 | Largest displacement of head position from the initial position: **(A)** Forward displacement, **(B)** Backward displacement, **(C)** Leftward displacement, **(D)** Rightward displacement (NS: No stimulation, FS: front side, BS: back side, LS: left side, or RS: right side).

DISCUSSION

The purpose of the experiment was to investigate the effect of a novel ankle TES configuration on the subjective sensation of body tilt and objective body sway. Before the experiment, we hypothesised that ankle TES induces a subjective sensation of body tilt and an objective body sway in a direction opposite to the stimulated part. In the experiment, the participants reported that they felt a sensation of body tilt when the ankle TES was applied to the four targeted tendons (TA, AC, PL, and FDL). In addition, we observed an exclusive body sway in the backward, forward, leftward, and rightward directions when applying ankle TES to the targeted tendons. These results clearly confirm our hypothesis. In the following section, we will have a detailed discussion on the interpretation of these results.

Subjective Sensation of Ankle Joint Angle Induced by Ankle TES

The results of the questionnaire indicate that the LS condition induced a subjective sensation of rightward body tilt, while the RS condition induced a subjective sensation of leftward body tilt (**Figure 6A**). Moreover, the FS condition induced a subjective sensation of backward body tilt, while the BS condition induced a subjective sensation of forward body tilt (**Figure 6B**). Overall, these results indicate that ankle TES can induce a subjective sensation of body tilt in a direction opposite to the stimulated part, and H1 is proven to be appropriate.

For the induced sensation of body tilt, we have considered two possible mechanisms. The first possible mechanism is related to the prediction of movement based on the force sensation. Our previous research revealed that ankle TES can induce force sensation at the ankle in a direction opposite to the stimulated part (Takahashi et al., 2021). If the actual force is applied on the ankle, the body will tilt in the same direction as the applied force. It is possible that participants predicted their body movement based on the sensation of force induced by ankle TES, inducing the subjective sensation of body tilt. The second possible mechanism is that the body sway was actually induced by the ankle TES, and participants simply answered based on this sway.

Relationship Between Simulation Results and Sensation of Body Tilt

We selected four muscle tendons, TA, AC, FDL, and PL, which were anatomically responsible for dorsiflexion, plantarflexion, foot inversion, and eversion, respectively, as the stimulation targets by ankle TES, and placed the electrodes on the targeted tendons. Our simulation results, using the finite element method (*Simulation Result*), confirmed that each electrode configuration stimulated the target tendon under the electrodes most strongly among the four targeted tendons. However, it also observed that the current density distributions of the other targeted tendons could be affected; e.g., for the case of current application on the FDL, **Table 2** shows that the maximum current density on the TA is approximately 28% of that on the FDL.

However, as discussed in *Subjective Sensation of Ankle Joint Angle Induced by Ankle TES*, four experimental conditions with stimulation (FS, BS, LS, and RS) induced a subjective sensation of body tilt only in each of the four directions that were significantly stronger in the experiment as shown in **Figure 6**. This result indicates that a sensation of body tilt was dominated by the targeted tendon under the electrodes which were thought to be most strongly stimulated even when some tendons were stimulated together. We assume that the effect of the amplitude of the current wave, which was determined to be the maximum within the uncomfortable sensation threshold of the participant, on the targeted tendons other than that under the electrodes would not be strong enough to induce a sensation of body tilt in four cardinal directions. In addition, even when a combination of stimulation on the FDL of one foot and the PL of the other for the lateral directions was used, it seems that the participants could categorise the direction of the sensation of body tilt (i.e., leftward or rightward tilt). Therefore, our proposed electrode configuration could selectively induce a sensation of body tilt in one of four directions.

Body Sway Induced by Ankle TES

The results of the multiple comparison tests showed that the largest displacement of the CoP tends to increase toward the backward, forward, rightward, and leftward directions under the FS, BS, LS, and RS conditions, respectively. In addition, **Figure 8** shows that the largest backward displacement of the CoP under the LS condition is significantly larger than that under the NS condition. **Table 2** shows that the maximum current density on TA, which is revealed to increase the backward displacement of the CoP, is the second largest when the current is applied on FDL. Hence, we assume that this is a result of the subtle stimulation of TA on the right foot under the LS condition. Furthermore, we assume that the reason why only the LS condition, and not the RS condition, increases backward displacement is the difference of dominant foot among participants. Based on previous research, suggesting that a non-dominant foot has a more essential role in balance control than a dominant foot (Shigaki et al., 2013), stimulation to the non-dominant foot is expected to have a greater effect on the displacement of the CoP compared with the dominant foot. To investigate the influence of the dominant foot, we additionally asked participants which foot they perceived to be their dominant foot and divided them into two groups according to their dominant foot. The average of the maximum displacement of the CoP in all experimental conditions and directions in the left-footed group (five participants, 1.12 cm) is larger than that in the right-footed group (seven participants, 0.92 cm); this difference might make the effect of the LS condition more dominant.

In the case of head position, **Figure 10** shows that the largest displacement of head position in a specific direction is greatest when the stimulation is applied to the opposite side of the ankle, which is the same trend as the CoP. However, the result of multiple comparison tests did not show a clear effect of ankle TES on the largest displacement of head position. We assume that the larger variance of head position compared with CoP is a leading cause of this result. **Table 5** shows that the average standard error of the largest displacement of head position in each direction is nearly or more than two times larger than that of the CoP. If the soles of the feet are regarded as the centre of body sway when standing, the head position is influenced by the angle change of every joint in the lower limbs, which is assumed to increase the variance in the head position.

Overall, although the effect of ankle TES on head position is unclear, the result of the CoP indicates that ankle TES induces body sway in the direction opposite to the stimulated part and it is rational to consider that H2 is appropriate.

Moreover, the results include temporal characteristics of the effect of ankle TES on body sway. **Table 4** shows that the

TABLE 5 | Average of standard error of the largest displacement of CoP and head position in each direction.

Value type	Average of standard error of the largest forward displacement (cm)	Average of standard error of the largest backward displacement (cm)	Average of standard error of the largest leftward displacement (cm)	Average of standard error of the largest rightward displacement (cm)
CoP	0.15	0.12	0.15	0.14
Head position	0.29	0.27	0.30	0.33

displacement of the CoP begins to increase in a direction opposite to the stimulated part about 1 s after starting the stimulation, and tends to maintain this displacement after the stopping of the stimulation. From these results, we can assume that the body sway caused by ankle TES has latencies of the stimulus onset and offset on the order of seconds. This is potentially due to the inertia of the human body. According to Newton's laws of motion, it takes a certain amount of time to cause body sway from a static upright position by the torque generated in the ankle joint. In addition, the motor response to the ankle TES does not appear to differ greatly across stimulation directions.

Illusory Sensation of Ankle Joint Angle Induced by Ankle TES

While the body sway is considered to be caused by the stimulation of GTO, as described in *Hypothesis*, we noted two possible mechanisms of the subjective sensation of body tilt: the result of the movement prediction based on the force sensation or the result of initially induced body sway. It is impossible to determine which mechanism is correct from the results of this experiment. In either case, however, the result of body sway indicates that the targeted tendons are stimulated by ankle TES and the illusory sensation of the ankle joint angle is induced.

LIMITATIONS AND FUTURE WORK

Multiple Stimulations of Different Tendons

Our experiment showed that ankle TES can induce a subjective sensation of body tilt in the direction opposite to the stimulated part. In this experiment, however, only one tendon in each leg was stimulated at a time. In addition, the direction of the sensation was limited to the four cardinal directions: leftward, rightward, forward, and backward. By stimulating multiple tendons in each leg simultaneously, we expect that the effects of ankle TES could be combined, and a sensation of body tilt in various directions can be induced.

Effect of Ankle TES on Perception of Ground Tilt

As mentioned above, the expected use of our proposed method is to change the perception of ground tilt by applying ankle TES. However, participants were asked about their subjective sensation of body tilt in the experiment, rather than their perception of ground tilt. Since this is a basic study to demonstrate the effectiveness of our novel ankle TES configuration on the sensation of the ankle joint angle, we wanted to clearly investigate the effect on the perception of ankle joint angle. We assumed that the perception of ground tilt was affected by the preconception that the angle of the ground does not change, since the experiment was conducted on a static floor, without any visual information. Therefore, we selected the subjective sensation of body tilt as a reliable parameter to evaluate the sensation of the ankle joint angle.

The experimental results showed that ankle TES induces a subjective sensation of body tilt, indicating that the illusory sensation of the ankle joint angle, which is essential for changing the perception of ground tilt, can be induced. Therefore, in the next experiment, we will apply the same stimulation on a floor that can change the angle of tilt and investigate whether the illusory sensation of the ankle joint angle can change the perception of ground tilt.

Temporal Characteristics of the Sensation Induced by Ankle TES

To make the effect of ankle TES consistent with visual information provided by an HMD, it is desirable that the sensation of the ankle joint angle be induced after applying ankle TES with minimum time delay and disappear immediately after stopping the stimulation. As discussed in *Relationship Between Simulation Results and Sensation of Body Tilt*, the body tilt caused by ankle TES appears to have a few seconds of onset and offset latencies. However, we did not investigate a temporal change in the subjective sensation of body tilt. Therefore, further research is required to make sure that the sensation is induced without inconsistency with visual stimuli.

Relationship Between Stimulus Intensity and Effect of Ankle TES

In the experiment, the amplitude of the current wave was adjusted at first and was not changed during the experiment. When presenting the sensation of ankle joint angle that is, consistent with a virtual environment, however, the extent of the effect of ankle TES should be controlled. Takahashi et al. (2016) showed that the amount of force sensation induced by electrically stimulating a wrist tendon can be controlled by the pulse height of the current. This result suggests that the extent of the effect of ankle TES can also be controlled by stimulus intensity. In order to make our proposed method more suitable for VR experiences, the relationship between stimulus intensity and the extent of the effect of ankle TES on the illusory sensation of ankle joint angle has to be investigated.

Application of Ankle TES to VR Experiences

During the application of ankle TES, some participants reported the sensation of their foot soles vibrating, which could be caused by stimulating nerves that dominate the haptic sensation. When we apply our proposed method to VR experiments, this sensation could make users uncomfortable and disturb their immersive experiences. Since we did not conduct a subjective evaluation of feelings of discomfort during the stimulation, it should be investigated in the future. Additionally, the safety of ankle TES is essential. As the result of our experiment shows, ankle TES causes actual body movements. These movements might reduce the stability of a user's body and increase the risk of falling over. To confirm the safety of ankle TES, the stability of body during the stimulation should also be investigated by conducting balance assessments such as Berg Balance Scale (BBS) or the Single-Leg Stance (SLS) (Flansbjer et al., 2012).

CONCLUSION

In this study, we developed a novel method that possibly induces a sensation of standing on a slope by applying ankle TES, and investigated the effect of ankle TES on the subjective sensation of body tilt and actual postural change. The experimental results revealed that ankle TES induces both a sensation of body tilt and actual body tilt in a direction opposite to the stimulated part, in the lateral and anteroposterior directions. In the future, we will investigate the effect of ankle TES on the perception of ground tilt instead of body tilt and implement a method that can induce a sensation of standing on a slope. By utilising our method in virtual environments that have rugged terrain, such as mountains or hills, and controlling the extent and direction of the illusory sensation of ground tilt according to the position of a user in a virtual environment, a much more immersive VR experience can be realised.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethics research committee of the Graduate School of Information Science and Technology, The University of Tokyo. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

NT, KA and TA proposed the method, and TA and TN designed the experiment. NT conducted the experiments and collected the data. KA, MH and HK conducted the simulations. NT, KA and TN analysed the data. NT, KA and TA wrote the manuscript. All the authors discussed the results and commented on the manuscript.

FUNDING

This study was supported by the JST, PRESTO (JPMJPR19J1), and JSPS KAKENHI (JP21H04883).

REFERENCES

Aoyama, K., Higuchi, D., Sakurai, K., Maeda, T., and Ando, H. (2017). "GVS RIDE: Providing a Novel Experience Using a Head Mounted Display and Four-Pole Galvanic Vestibular Stimulation," in ACM SIGGRAPH 2017 Emerging Technologies, Los Angeles, CA, July 30–August 3, 2017, 1–2.

Boletsis, C. (2017). The New era of Virtual Reality Locomotion: A Systematic Literature Review of Techniques and a Proposed Typology. *Multimodal Tech. Interaction* 1 (4), 24. doi:10.3390/mti1040024

Bruder, G., and Steinicke, F. (2014). "Threefolded Motion Perception during Immersive Walkthroughs," in Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology, Edinburgh, Scotland, November 11–13, 2014, 177–185.

Flansbjer, U.-B., Blom, J., and Brogårdh, C. (2012). The Reproducibility of Berg Balance Scale and the Single-Leg Stance in Chronic Stroke and the Relationship between the Two Tests. *PM&R* 4 (3), 165–170. doi:10.1016/j.pmrj.2011.11.004

Goodwin, G. M., McCloskey, D. I., and Matthews, P. B. C. (1972). Proprioceptive Illusions Induced by Muscle Vibration: Contribution by Muscle Spindles to Perception? *Science* 175 (4028), 1382–1384. doi:10.1126/science.175.4028.1382

Hagimori, D., Yoshimoto, S., Sakata, N., and Kiyokawa, K. (2019). "Tendon Vibration Increases Vision-Induced Kinesthetic Illusions in a Virtual Environment," in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, March 23–27, 2019 (IEEE), 952–953.

Hothorn, L. A. (2016). The Two-Step Approach-A Significant ANOVA F-Test before Dunnett's Comparisons against a Control-Is Not Recommended. *Commun. Stat. - Theor. Methods* 45 (11), 3332–3343. doi:10.1080/03610926.2014.902225

Je, S., Lim, H., Moon, K., Teng, S. Y., Brooks, J., Lopes, P., et al. (2021). "Elevate: A Walkable Pin-Array for Large Shape-Changing Terrains," in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, May 8–13, 2021, 1–11.

Kajimoto, H. (2013). "Illusion of Motion Induced by Tendon Electrical Stimulation," in 2013 World Haptics Conference (WHC), Daejeon, Republic of Korea, April 14–17, 2013 (IEEE), 555–558.

Laakso, I., and Hirata, A. (2013). Computational Analysis Shows Why Transcranial Alternating Current Stimulation Induces Retinal Phosphenes. *J. Neural Eng.* 10 (4), 046009. doi:10.1088/1741-2560/10/4/046009

Moore, J. C. (1984). The Golgi Tendon Organ: A Review and Update. *Am. J. Occup. Ther.* 38 (4), 227–236. doi:10.5014/ajot.38.4.227

Rogers, J. H. (1980). Romberg and His Test. *J. Laryngol. Otol.* 94 (12), 1401–1404. doi:10.1017/s002221510009023x

Saltzman, C. (1997). Orthopaedic Foot and Ankle Society Ad Hoc Committee Report, January 1996. *Foot Ankle Int.* 18, 310–311.

Sasagawa, S., Ushiyama, J., Masani, K., Kouzaki, M., and Kanehisa, H. (2009). Balance Control under Different Passive Contributions of the Ankle Extensors: Quiet Standing on Inclined Surfaces. *Exp. Brain Res.* 196 (4), 537–544. doi:10.1007/s00221-009-1876-4

Schmidt, D., Kovacs, R., Mehta, V., Umapathi, U., Köhler, S., Cheng, L. P., et al. (2015). "Level-ups: Motorized Stilts that Simulate Stair Steps in Virtual Reality," in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, April 18–23, 2015, 2157–2160.

Seki, H., Yona, M., Naito, Y., and Muro, M. (2011). Effects of Vibration Applied on Forearm and Upper Limb Muscles on Human Voluntary Multi-Wrist and Elbow Movements. *Jpn. J. Phys. Fit. Sports Med.* 60 (2), 229–237. doi:10.7600/jspfsm.60.229

Shigaki, L., Rabello, L. M., Camargo, M. Z., Santos, V. B. D. C., Gil, A. W. D. O., Oliveira, M. R. D., et al. (2013). Comparative Analysis of One-Foot Balance in Rhythmic Gymnastics Athletes. *Rev. Bras Med. Esporte* 19, 104–107. doi:10.1590/s1517-86922013000200006

Soames, R. W., and Atha, J. (1982). The Spectral Characteristics of Postural Sway Behaviour. *Europ. J. Appl. Physiol.* 49 (2), 169–177. doi:10.1007/bf02334065

Steel, R. G. D. (1959). A Multiple Comparison Rank Sum Test: Treatments Versus Control. *Biometrics* 15 (4), 560–572. doi:10.2307/2527654

Takahashi, A., Tanabe, K., and Kajimoto, H. (2016). "Relationship between Force Sensation and Stimulation Parameters in Tendon Electrical Stimulation," in International AsiaHaptics Conference, Chiba, Japan, November 29–December 1, 2016 (Singapore: Springer), 233–238.

Takahashi, N., Amemiya, T., Narumi, T., Kuzuoka, H., Hirose, M., and Aoyama, K. (2021). "Perception of Illusory Body Tilt Induced by Electrical Tendon Stimulation," in International Conference on Human-Computer Interaction, July 24–29, 2021 (Cham: Springer). [Online], 357–368. doi:10.1007/978-3-030-78321-1_27

Tamaki, E., Chan, T., and Iwasaki, K. (2016). "UnlimitedHand: Input and Output Hand Gestures with Less Calibration Time," in Proceedings of the 29th Annual Symposium on User Interface Software and Technology, 163–165.

Tanaka, S., Ushiyama, K., Takahashi, A., and Kajimoto, H. (2020). "Movement-Free Virtual Reality Interface Using Kinesthetic Illusion Induced by Tendon Vibration," in International Conference on Human Haptic Sensing and Touch Enabled Computer Applications, Leiden, Netherlands, September 6–9, 2020 (Cham: Springer), 316–324. doi:10.1007/978-3-030-58147-3_35

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: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Virtual Reality

ORIGINAL RESEARCH
published: 27 April 2022
doi: 10.3389/frvir.2022.860872

Changing Finger Movement Perception: Influence of Active Haptics on Visual Dominance

Loën Boban[1,2]*, David Pittet[1], Bruno Herbelin[2] and Ronan Boulic[1]

[1]School of Computer and Communication Sciences, Immersive Interaction Group (IIG), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, [2]School of Life Sciences, Laboratory of Cognitive Neuroscience (LNCO), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

The perception of one's own body is a complex mechanism that can be disturbed by conflicting sensory information and lead to illusory (mis-) perceptions. Prominent models of multisensory integration propose that sensory streams are integrated according to their reliability by approximating Bayesian inference. As such, when considering self-attribution of seen motor actions, previous works argue in favor of visual dominance over other sensations, and internal cues. In the present work, we use virtual reality and a haptic glove to investigate the influence of an active haptic feedback on one's visual and agency judgments over a performed finger action under experimentally manipulated visual and haptic feedbacks. Data overall confirm that vision dominates for agency judgment in conditions of multisensory conflict. Interestingly, we also show that participants' visual judgment over their finger action is sensitive to multisensory conflicts (vision, proprioception, motor afferent signals, and haptic perception), thus bringing an important nuance to the widely accepted view on a general visual dominance.

Keywords: multisensory integration, multisensory conflicts, virtual reality, haptics, sense of agency, sense of ownership, visual perception, motor perception

OPEN ACCESS

Edited by:
Stephen Palmisano, University of Wollongong, Australia

Reviewed by:
Juno Kim, University of New South Wales, Australia
Daniel Rene Mestre, Aix-Marseille Université, France

***Correspondence:**
Loën Boban
loen.boban@epfl.ch

Specialty section:
This article was submitted to Virtual Reality and Human Behaviour, a section of the journal Frontiers in Virtual Reality

Received: *23 January 2022*
Accepted: *11 March 2022*
Published: *27 April 2022*

Citation:
Boban L, Pittet D, Herbelin B and Boulic R (2022) Changing Finger Movement Perception: Influence of Active Haptics on Visual Dominance. Front. Virtual Real. 3:860872. doi: 10.3389/frvir.2022.860872

1 INTRODUCTION

Bodily perception is a complex mechanism that can be disturbed by multimodal conflicts. Remarkably, the felt position of one's hand can be altered by inducing ownership over a rubber hand counterpart (Botvinick and Cohen, 1998). In the rubber hand illusion (RHI), multimodal conflicts are introduced between visual, tactile, and proprioceptive signals, so that the congruence of vision and touch can be considered to be more reliable than proprioception alone, thus giving rise to the ownership illusion of the rubber hand and to a visual dominance for locating the limb. This interpretation of the neural mechanism of the RHI is supported by prominent understanding on multisensory integration, according to which the brain integrates multiple sensory streams weighted by their reliability and approximating Bayesian inference (Knill and Pouget, 2004). A Bayesian brain model can indeed explain the proprioceptive drift to come from the computation of the felt position of the hand using a weighted sum of available sensory signals (Holmes and Spence, 2005). Extending this principle, if a sensory modality is added and is reliable enough, it can influence the multisensory integration, and thus affect one's feeling of ownership of one's limbs as well as its felt location. For instance, it can be the case when adding haptic feedback in a Virtual Reality (VR) context in which the real movements of one's hand are not perfectly matching with the provided visual feedback (i.e., visuo-proprioceptive conflict). In this situation, careful manipulation of the haptic information can influence the multisensory integration by making one configuration more reliable (e.g., congruent visuo-haptic feedback), thus influencing the acceptance of the proprioceptive distortion (Lee et al., 2015).

In such Bayesian framework contexts, when vision is supposed to be reliable, the integration of conflicting visuo-motor information may lead to a visual dominance over afferent motor signals and proprioception. This phenomenon was indeed exemplified by Salomon et al. (2016) by visually displaying a virtual hand that did not move according to the subject's actual movement. They could observe that, when the movement of two fingers was swapped (e.g., index finger moving instead of the middle one), participants could wrongly self-attribute the seen motor action. Although striking, this experiment does not allow concluding on a generalized visual dominance for agency judgments because the haptic feedback was not manipulated so its influence could not be evaluated. The multisensory integration in Salomon et al. (2016) indeed involved internal cues (proprioceptive and afferent motor information), visual feedback (observed finger movement, in conflict with the internal cues), but also a passive haptic feedback from touching the table on which the hand was resting. To elaborate on this work and evaluate how haptic feedback influences multisensory integration and, in particular, if it can alter visual dominance for motor perception, we propose to study more deeply the multisensory integration leading to the misattribution of virtual fingers' movements in an experimental paradigm involving systematic haptic feedback manipulation.

This paper is organized as follows. The next section presents the related work about multisensory integration in cognitive neuroscience and multisensory conflict. **Section 3** describes the experiment, including the setup, the task and conditions, the protocol, and our hypotheses. The statistical analysis is detailed in **Section 4**, while results are discussed in **Section 5**, followed by the conclusion in **Section 6**.

2 RELATED WORK

Multisensory integration designates the brain's process involved in converting a given set of sensory inputs into a unified (integrated) perception. A prominent model about this process is based on Bayesian inference, proposing that incoming sensory signals are integrated according to the inverse of their variance, i.e., their reliability, thus giving more weight to more precise modalities (Knill and Pouget, 2004; Friston, 2012). Illustrating this phenomenon, the rubber hand illusion involves presenting participants with a rubber hand placed nearby their hidden real hand (Botvinick and Cohen, 1998). The illusion of ownership for the rubber hand is then induced by brushing synchronously the subject's real hand and its rubber counterpart. The illusion then rises from a multisensory conflict in which the visuo-tactile congruency is considered as more reliable than proprioception. Noteworthy, in this particular context of illusory ownership toward a virtual bodily counterpart, a reciprocal link between multisensory integration, and illusory ownership was found: the ownership illusion emerges from multisensory integration, while in return it influences upcoming multisensory processing. For example, Maselli et al. (2016) showed that the ownership illusion acts as a "unity assumption" and, congruently with bayesian causal inference models, enlarges the temporal window for integrating visuo-tactile stimulus.

In the particular context of the rubber hand illusion (RHI), the ownership illusion induces a proprioceptive drift, meaning that one's perception of one's hand location shifts toward the rubber hand location. But proprioceptive drifts had been observed earlier in a similar paradigm, without tactile stimulation, using a prism to manipulate one's static hand position (Hay et al., 1965). This phenomenon was later confirmed, and more deeply investigated, by Holmes and Spence (2005) who used a mirror to modulate the visual position of the participants' right hand. They showed that participants estimate their hand position to be closer to the visual information when asked to move their fingers than when only presented with a passive visual hand reflection. Taken together, these findings could lead to the idea that vision necessarily dominates over proprioception. It is indeed often considered that, in the context of VR, vision is more reliable than other sensory cues, leading to a "visual dominance," or "visual capture," that overrides other factors of movement perception (Burns et al., 2005). But other factors and combinations of conditions are to be considered, as they can lead to an increase or a decrease of the reliability associated to a specific modality. As counterexamples, it is known that the spatial direction of the visual distortion strongly impacts the visuo-proprioceptive integration for the judgment of hand location (van Beers et al., 2002), or that the haptic perception can dominate vision when one is judging about shapes if two radically different configurations are presented to vision and touch (Heller, 1983).

More dramatically, beside the manipulation of one's proprioceptive judgment, the ownership illusion can also enable manipulations of one's sense of authorship toward a movement, namely one's sense of agency (Blanke and Metzinger, 2009). Illustrating this possibility, and laying the foundation for the present research, Salomon et al. (2016) manipulated their participant's sense of agency using an anatomical distortion of their hand's movements. Participants were asked to rest their right hand on a table and to lift either their index or middle finger. Their real hand was hidden while they were presented with a virtual representation of their hand. On trials with distortion, the consistency principle of agency (Jeunet et al., 2018) was challenged by visually moving the virtual hand's index finger if the participant moved the middle finger, and conversely. Under these circumstances, Salomon et al. showed not only that participants would experience a lower sense of agency, but that they could wrongly self-attribute the motor action of the virtual finger they saw moving. This motor illusion is a typical example of visual capture in a case of conflicting integration, as the visual stimuli here dominates over tactile (passive tactile feedback when bringing back the

FIGURE 1 | Global view of the experimental setup.

moved finger on the table), proprioceptive and motor afferent signals. These results however do not allow us to make conclusions on the influence of haptic feedback in the resolution of multisensory conflicts. We thus propose to investigate whether adding an active haptic feedback (a reaction force) could influence the multisensory conflict leading to agency judgment and motor attribution, thereby providing an important nuance to the generally accepted view on visual dominance, and opening new research perspectives on agency manipulations.

3 MATERIALS AND METHODS

Our protocol replicated Salomon et al. (2016) and extended it with *congruent* and *incongruent* active haptic feedback in order to test whether or not such active haptic feedback can dominate vision in a motor perception context.

3.1 Setup and Virtual Scene

Participants sat on a chair in front of a table, and were equipped with an HTC Vive Pro Eye Head Mounted Display (HMD). They were wearing a Dexmo glove from Dextra robotics (https://www.dextarobotics.com/) on their right hand, on which an HTC Vive tracker was fixed (**Figure 1**).

The Dexmo haptic glove can prevent a finger from moving, or apply a force of varying intensity on the user's fingertips. We used the finger-tracking data it provides to animate the virtual hand [androgynous right hand from Schwind et al. (2018)].

The participants' right wrist was kept static using a strap attached to the table. The right hand was resting on a rectangular surface (**Figure 2**), while the left one was holding a HTC vive controller to answer questions.

Participants saw a virtual environment through a virtual reality head mounted display (HMD) from a first person perspective (**Figure 3**). The virtual scene was composed of a virtual table placed congruently with the real table, on which a rectangular prop was placed. On top of this prop were attached two virtual obstacles (**Figure 2**) that the virtual index and middle fingers would hit when being lifted. The eye tracking system was used to ensure that participants were continuously looking at a red fixation cross located between the index and middle fingers (**Figure 3**).

3.2 Task and Experimental Conditions

During the experiment, participants were asked to lift either their index or middle finger until it touched the corresponding virtual obstacle, and then to bring their finger back on the table. When a participant lifted the index or middle finger, the virtual hand would simultaneously lift the same finger (*congruent* visual) or the other finger (*incongruent* visual). When the moving virtual finger touched the obstacle, haptic feedback was applied either on the lifted finger (*congruent* haptic) or on the other finger tip (*incongruent* haptic). That is, when the haptic feedback is *congruent*, the glove applies a force on the moving finger to prevent it from passing through the virtual obstacle. Conversely, when the haptic feedback is *incongruent*, the moving finger does not receive any haptic feedback from the

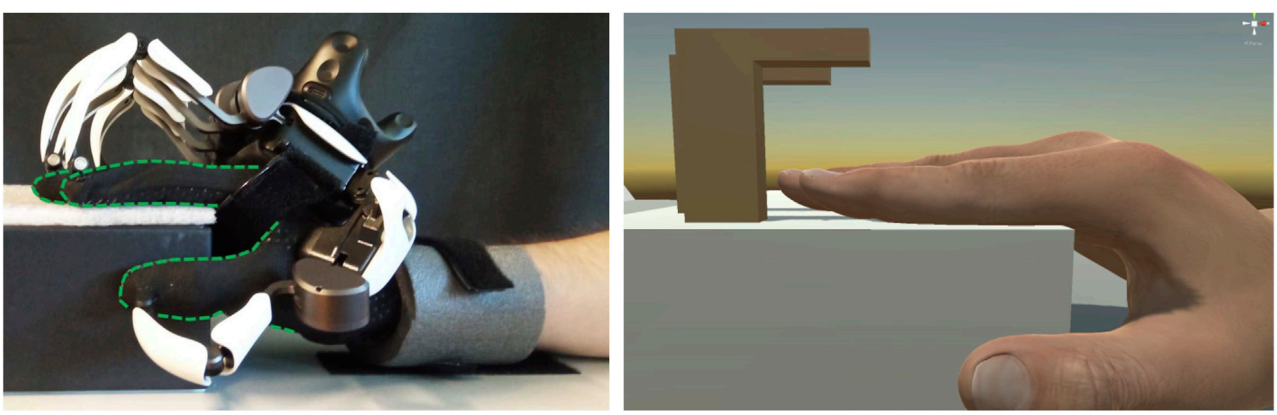
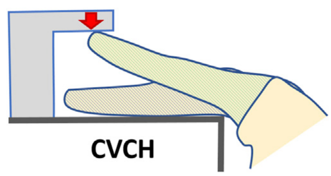
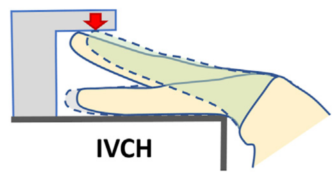
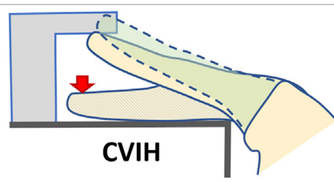
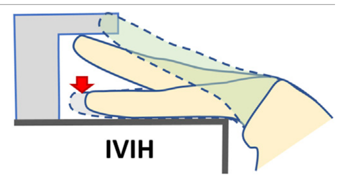


FIGURE 2 | Participant's real hand (outlined in green) and corresponding virtual hand.

FIGURE 3 | The virtual scene viewed from the first person perspective; during the task, participants had to look at the red cross located on the surface between the index and middle fingers.

TABLE 1 | Experimental conditions illustrated in the context of the subject moving their index finger (highlighted in green with plain line if displayed or dotted line otherwise).

	Congruent visual	Incongruent visual
Congruent haptic		
Incongruent haptic		

The displayed hand and fingers are in yellow or in beige (when consistent with proprioception). The real pose of the middle finger is indicated in grey with dotted lines for the incongruent *visual condition.*

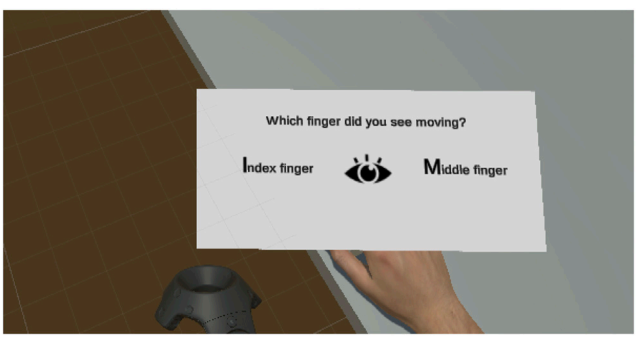


FIGURE 4 | Visual perception question panel.

glove, but a force is applied on the resting finger's tip, pushing it against the platform. Our experiment follows a 2 × 2 factorial design, with two factors (haptic and visual feedbacks) manipulated with two levels (*congruent* or *incongruent* with the movement actually performed by the participant). **Table 1** sums-up the experimental conditions.

3.3 Experimental Protocol

3.3.1 Trials and Blocks

Each trial consisted of one task (i.e., finger lifting) followed by a question. During the movement, participants had to look at the fixation cross located between their index, and middle fingers (**Figure 3**).

When participants brought back their moved finger on the rectangular platform, a question appeared in front of the virtual hand (**Figure 4**). It was randomly selected between *visual* perception ("Which finger did you see moving?") and motor perception ("Which finger did you move?"), and answers were presented in random order. Participants

TABLE 2 | Agency and ownership questionnaire.

Agency items	• *I felt as if I was causing the movement I saw* • *The virtual hand moved just like I wanted it to, as if it was obeying my will* • *I felt as if I was controlling the movements of the virtual hand* • *Whenever I moved my finger I expected the virtual finger to move in the same way*
Ownership items	• *I felt as if I was looking at my own hand* • *It seemed as if I were sensing the movement of my finger in the location where the virtual finger moved* • *I felt as if the virtual hand was part of my body* • *I felt as if the virtual hand was my hand*

answered the question using the Vive controller held in their left hand.

A full experimental block consisted in 30 trials of each condition, thus in 120 trials in total, presented in pseudo-random order. For each trial, participants could freely choose to move either their index or middle finger; they were however asked to keep the balance as much as possible. To help participants, how many times they moved each finger was displayed when half of a block was completed.

3.3.2 Procedure

When a participant arrived, the experiment was first explained to them, they were shown the setup, signed a consent form and filled in a demographic questionnaire. Then, they sat on the chair in front of the table and put on the HMD. First, a stereoscopic vision test was conducted to ensure they would correctly see the fingers movements during the experiment. Then the eye-tracking system was calibrated. Next, the participant removed the HMD and was helped to put on the glove and to place their hand correctly for the calibration to be performed.

The *calibration phase* went as follows. First, the height of the physical prop was adjusted to fit the participant's hand size. This way, the angle between the index/middle fingers plane and the palm was ensured to be sufficiently below 180° to allow participants to lift their fingers while wearing the glove. Indeed the maximal extension of the hand tolerated by the haptic glove corresponds to a flat hand (180° between the palm and the index/middle fingers plan). Finally, the virtual hand pose was adjusted to fit the participant's fingers pose on the rectangular prop, as well as the position to achieve for triggering the haptic feedback.

Following the calibration, participants put back the HMD and completed a two-minutes *acclimatization phase* to get used to the virtual hand. During this phase, participants were asked to keep their thumb still and the other fingers fully extended, but could freely rotate their wrist and lift their fingers from the rectangular prop. A *training phase* followed, consisting in 48 trials. The training block was composed of 12 trials of each condition, presented in random order.

After the training, each participant went though the three *main blocks*, thus performing 360 trials in total. They could take a break between each block if they wanted to. Once completed, subjects had to perform the *Agency and Ownership blocks*: four additional blocks of 48 trials each. Each block corresponded to an experimental condition and the order of blocks was randomized. During this phase, the visual and motor perception questions were removed. After each block, as in the study by Salomon et al., participants were asked to fill a questionnaire on agency and ownership toward the virtual hand (**Table 2**). Subjects had to answer the questions on a seven-points Likert scale ranging from 1 ("Totally disagree") to 7 ("Totally agree").

Except during the training blocks, white noise was used to isolate participants from the external world. Moreover, if between two blocks the participant's hand had moved from its initial position, for example if the participant took a pause, the calibration phase was repeated.

3.4 Participants

Following a sample size estimation based on Salomon et al. (2016), 20 healthy right-handed volunteers were recruited for this study (twelve females), aged from 18 to 35 ($m = 21.65$, $sd = 4.36$). All participants gave informed consent and the study was approved by our local ethics committee. The study and methods were carried out in accordance with the guidelines of the declaration of Helsinki.

3.5 Hypotheses

First, we expect the replication of previous results from Salomon et al. (2016) (H1). We thus have four sub-hypothesis (see **Table 1** illustrating the conditions acronyms): (H1a) Visual judgment accuracy is not affected by visual congruency (CVCH not significantly different from IVCH), (H1b) Motor judgment accuracy is reduced when the visual feedback is *incongruent* with the motor actions (CVCH > IVCH), (H1c) agency is lower under *incongruent* visual feedback (CVCH > IVCH), and (H1d) ownership is reduced when the visual feedback is *incongruent* (CVCH > IVCH).

Second, our (H2) hypothesis focuses on the influence of an active haptic feedback on one's perception. Concerning visual perception, as Salomon et al. showed vision to be considered as highly reliable in a very similar set-up, we anticipate no effect of the active haptic feedback on visual judgment accuracy. Thus, we hypothesize that (H2a) visual judgment accuracy is not significantly affected by visual nor by haptic feedback congruency.

When it comes to motor perception judgment accuracy, we believe that vision will be predominant in our context as no visual cues indicates the visual feedback to be erroneous (Power, 1980). We thus hypothesize that (H2b) motor judgment accuracy is higher if the visual modality is *congruent* than if it is *incongruent*, no matter the associated haptic feedbacks (CV...H > IV... H).

Given that agency is typically induced when one can control the virtual bodily counterpart, it should be experienced by participant in our setup when the visual feedback is *congruent* with the movement performed. In this condition, and as our haptic feedback manipulations do not disturb the matching

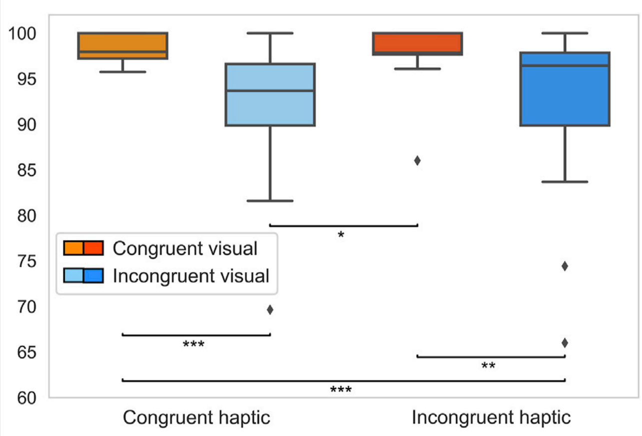


FIGURE 5 | Motor perception judgment accuracy across conditions (in percent).

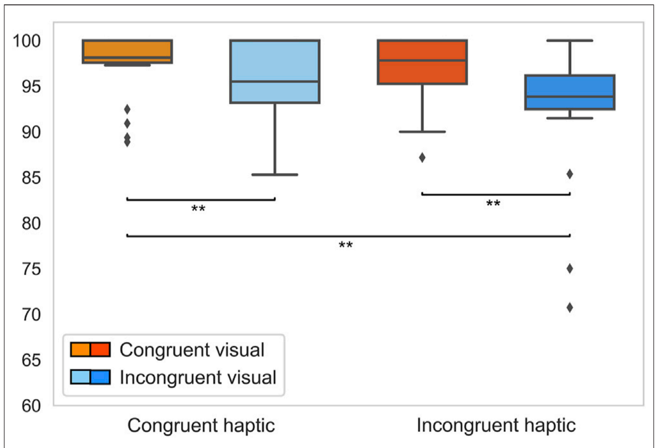


FIGURE 6 | Visual perception judgment accuracy across conditions (in percent).

between the performed and the seen movement, we do not expect haptic information to disturb the sense of agency. Conversely, in conditions when the visual feedback is *incongruent* with the performed movement, participants should not experience a strong sense of agency. Similarly as before, as the haptic modality does not influence the performed/seen movement consistency, we do not expect *congruent* haptic feedback to alter the sense of agency. As a sum-up, under both visual feedback conditions, we do not expect the haptic modality to disturb the feeling of agency (H2c).

Finally, concerning ownership, Bovet et al. (2018) previously showed the propensity of visual-haptic conflict to reduce ownership. We thus expect that, under *congruent* visual feedback, *incongruent* haptic feedback should lower ownership levels. Conversely, under *incongruent* visual feedback we do not expect any significant effect of the haptic modality as, in this condition, the scores already reported by Salomon et al. indicate low levels of ownership toward the virtual hand (H2d).

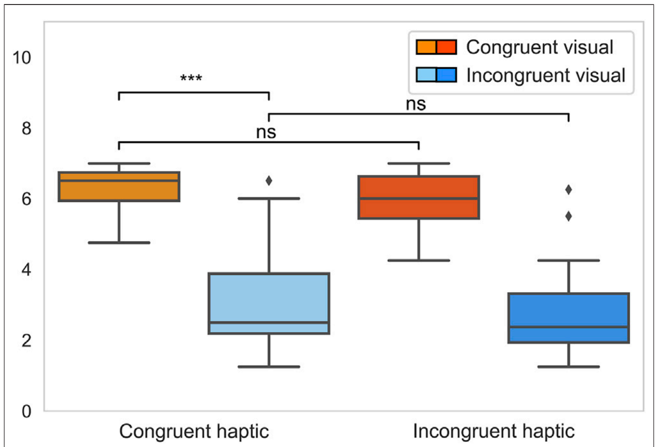


FIGURE 7 | Agency scores across conditions, computed for each participant using the mean of answers to the corresponding questions in the last four blocks of the experiment.

4 STATISTICAL ANALYSIS

Visual and motor judgment accuracies were computed for each participant as the percentage of correct answers to the corresponding perception questions. As for both measurements the samples included several outliers and were not approximating a normal distribution (Shapiro-Wilk test), the statistical analysis was conducted using non-parametric tests.

Concerning agency and ownership scores, they were computed separately for each participant using the mean of answers to the corresponding questions in the last four blocks of the experiment. For both agency and ownership data, the samples were not outliers-free. As there was no reason to exclude the outliers from the dataset, and as they influenced the results of the associated t-tests, we conducted the statistical analysis using non-parametric tests.

In case of multiple comparisons, p-values were corrected with the conservative Bonferonni method and reported using the following notation: p_{corr}.

4.1 Motor Perception Judgment Accuracy

A one-sided Wilcoxon test was applied on the pair (CVCH; IVCH) to test order, revealing a significantly ($p < 10^{-3}$) higher motor perception accuracy under *congruent* visual feedback (**Figure 5**), thus validating (H1b).

To test whether motor judgment accuracy is influenced by congruency (H2b), one-sided Wilcoxon tests were applied on pairs comparing *congruent* and *incongruent* visual feedback. The motor perception accuracy was significantly higher for the CVCH condition than for the IVCH one ($p < 10^{-3}$ and $p_{corr} < 10^{-3}$) and the IVIH one ($p \simeq 0.003$ and $p_{corr} \simeq 0.012$). Similarly, the CVIH condition led to a significantly higher motor judgment accuracy than the IVIH ($p \simeq 0.002$ and $p_{corr} \simeq 0.01$) and the IVCH one ($p < 10^{-3}$ and $p_{corr} < 10^{-3}$). Thus, those results validate (H2b) (**Figure 5**).

4.2 Visual Perception Judgment Accuracy

A two-sided Wilcoxon test was applied to compare the CVCH and IVCH samples, revealing a significant effect of visual

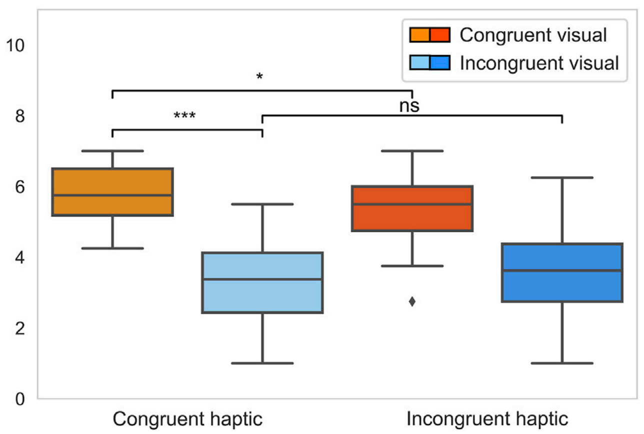


FIGURE 8 | Ownership scores across conditions, computed for each participant using the mean of answers to the corresponding questions in the last four blocks of the experiment.

feedback on visual perception accuracy ($p \simeq 0.02$), thus rejecting (H1a) (**Figure 6**). The associated one-sided Wilcoxon test showed visual judgment accuracy to be decreased when the visual modality is *incongruent* ($p < 10^{-2}$).

A Friedman test applied to the four samples rejected the similarity of distributions ($p < 10^{-3}$), thus rejecting our (H2a) hypothesis according to which visual perception accuracy would not be affected by visual nor haptic congruency.

A post-hoc analysis was conducted using one-sided Wilcoxon tests on each possible pair of conditions (**Figure 6**). The visual judgment accuracy of the IVIH condition was found to be significantly lower than for the CVCH and CVIH ones ($p < 10^{-3}$ and $p_{corr} < 10^{-2}$ for both tests).

4.3 Agency

A one-sided Wilcoxon test was applied on *congruent* haptic feedback samples, revealing a significantly ($p < 10^{-3}$) higher agency under *congruent* than *incongruent* visual feedback (**Figure 7**), thus validating (H1c).

To test for (H2c), according to which agency would not be affected by the haptic feedback congruency, two-sided Wilcoxon tests were applied on the pairs (CVCH; CVIH), and (IVCH; IVIH). None was significant ($p > 0.05$), thus failing to reject this hypothesis (**Figure 7**).

4.4 Ownership

A one-sided Wilcoxon test was applied on *congruent* haptic feedback samples, revealing a significantly ($p < 10^{-3}$) higher ownership under *congruent* than *incongruent* visual feedback (**Figure 8**), thus replicating previous findings and validating (H1d).

To test for (H2d), according to which ownership would be affected by the haptic feedback congruency only under congruent visual information, a one-sided Wilcoxon test was applied on the pair (CVCH; CVIH) while a two-sided one was applied on the pair (IVCH; IVIH). Under *congruent* visual information, the *congruent* haptic feedback led to significantly ($p \simeq 0.03$) higher ownership scores than the *incongruent* one (**Figure 8**). Under *incongruent* visual feedback, no effect of haptic feedback congruency was found ($p \simeq 0.29$).

5 DISCUSSION

First, we reproduced and extended Salomon et al. (2016) results on the influence of visual feedback on judgment of motor authorship. More precisely, motor perception is affected by the congruency between the performed and the observed movement: when haptic feedback is *congruent* with the action performed, *congruent* visual information leads to a higher motor perception accuracy than *incongruent* visual information (CVCH > IVCH). It thus indicates that, as previously shown in Salomon et al. (2016), vision can overrule action information (proprioceptive, tactile, and motor afferent signals) when one is judging about the movement one performed. Said differently, one's awareness of one's movement and proprioceptive consequences (motor afferent signals, proprioception, and tactile incoming signals) can be overruled by conflicting visual information.

As an extension of this result, we show that in our setup, no matter the haptic feedback congruency, *congruent* visual feedback always leads to less motor judgment errors than an *incongruent* one. Of particular interest, if visual or haptic information is *incongruent* (conditions CVIH and IVCH), then *congruent* visual feedback with *incongruent* haptic stimuli leads to a better motor perception accuracy ($m \simeq 98\%$) than *incongruent* visual feedback coupled with *congruent* haptic stimuli ($m \simeq 92\%$), thus showing that haptic information does not challenge vision in the motor judgment context. Put together, when it comes to motor judgment, these findings argue in favor of the visual dominance over the other modalities being integrated to produce one's perception.

Concerning visual perception accuracy, our findings are different from (Salomon et al., 2016) as when the haptic feedback is *congruent* with the performed movement, visual judgment accuracy is significantly higher under *congruent* than *incongruent* visual feedback (CVCH > IVCH). As for their experimental manipulation, when the visual stimulus is *incongruent*, vision contradicts proprioception, motor afferent signals, and haptic perception. As shown by our data, this conflict influences visual perception, meaning that the consistency of this group of three modalities is sufficient to cast doubt on which finger one saw moving, thus resulting in a drop in visual perception accuracy. One possible explanation for the difference between our results and previous work from Salomon et al. (2016) is the difference in terms of haptic feedback. Indeed, the haptic feedback is stronger in our setup as it not only comes from passive touch of the table but also from the contact with the virtual obstacle, thus strengthening non-visual stimuli in the multisensory integration process. Consistently with (H1a) rejection, when comparing the four experimental conditions, we found visual perception accuracy to be higher when the visual feedback is *congruent* than when both visual and haptic modalities are *incongruent* (CVCH, CVIH > IVIH). These findings indicate that when both visual and haptic feedbacks are *incongruent* with the movement performed, the haptic feedback does not reinforce visual information enough to make vision as reliable as if it was *congruent* with the movement performed. Thus, incongruencies between the performed movement and the visual-haptic feedback can impact visual perception accuracy by blurring one's visual judgment, e.g., casting doubt on which finger one saw moving.

Now considering perception judgment accuracy under *congruent* haptic feedback, Salomon et al. (2016) made the distinction between two accuracy patterns by manipulating the visual feedback congruency. They experimentally changed the position of the virtual hand to be either in the same orientation as the physical subject's hand (0° rotation), in a plausible orientation (90° rotation), or in an impossible posture (180°or 270° rotation). They observe that with a plausible hand posture, the accuracy pattern presents a small decrease in visual perception accuracy (but not significant) and a large decrease in motor perception accuracy when visual feedback is switched from *congruent* to *incongruent*. Conversely, when the virtual hand is in an impossible posture, both visual and motor perception accuracy largely decrease when visual feedback is *incongruent*. In our experiment, the virtual hand is presented with a 0° rotation, and our results replicate the corresponding pattern from Salomon et al. (2016). Indeed, we observe a slight decrease in visual judgment accuracy when the visual feedback switches from *congruent* to *incongruent*. This decrease is significant in our setup ($m \simeq 97.2\%$ to $m \simeq 95.3\%$) but not in Salomon et al. (2016) ($m \simeq 94.8\%$ to $m \simeq 93.8\%$), which may be explained by a more pronounced haptic feedback in our setup. However, as we also observe a large drop in motor perception accuracy when the visual feedback becomes *incongruent* ($m \simeq 98.4\%$ to $m \simeq 92\%$), our results fit the pattern described by Salomon et al., thus confirming their general finding. As improvements in motor perception can in turn bear potential benefits for motor performance, future work could explore the impact of such active haptic feedback on motor performance to infer its potential benefit for motor training (Ramírez-Fernández et al., 2015; Kreimeier et al., 2019; Odermatt et al., 2021).

Finally, considering the agency and ownership scores, we replicated previous findings showing that, under *congruent* haptic feedback, both scores are higher when participants are provided with *congruent* visual information than *incongruent*. This is in line with agency and ownership being typically induced when one is presented with a virtual bodily counterpart they can control and that is placed in an anatomically plausible position. Moreover, because the haptic feedback factor does not impair one's control of the virtual hand as it does not disturb the matching between the performed and the seen movement, we did not observe any significant difference in agency scores when comparing conditions differing only by the haptic feedback condition. In addition, as hypothesized, when provided with *congruent* visual information, the participant's sense of ownership was significantly lower in the *incongruent* haptic feedback condition than in the *congruent* one. This result is in line with previous work from Bovet et al. (2018) showing the propensity of a visual-haptic conflict to reduce ownership. Conversely, when the visual information was *incongruent*, haptic feedback had no significant effect on ownership scores, and in particular, it did not bring back the ownership illusion.

6 CONCLUSION

In cases when a multisensory conflict arises, such as in a VR simulation, vision is considered as dominant over other sensory cues. Typically, when considering self-attribution of seen motor actions, previous work tends to confirm this tendency for fingers movements (Salomon et al., 2016). In the present work, through the use of VR and of an haptic glove, we provide an important nuance to this widely accepted view on visual dominance by investigating the influence of an active haptic feedback on one's visual and agency judgments about a seen movement.

Overall, we confirm vision dominates multisensory conflicts when it comes to agency judgments of motor action. But conversely, we show that visual dominance can be challenged by the association of proprioceptive and motor afferent signals with active haptic perception, as shown in our results with a lower accuracy for the judgment of which finger one saw moving. Finally, our experiment shows an effect of a punctual and transient haptic stimulation on the sense of ownership for a virtual hand. When the active haptic feedback is incongruent with the performed movement, participants experience a loss of ownership for the hand, even if visuals were congruent.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the CER-VD (Commission cantonale d'éthique de la recherche sur l'être humain) N°2018-01601. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: LB, BH, and RB. Performed the experiments: DP. Analyzed the data: LB. Contributed reagents/materials/analysis tools: DP and LB. Wrote the manuscript: LB, BH, and RB.

FUNDING

This work has been supported by the SNFS project "Immersive Embodied Interactions" grant 200020-178790.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2022.860872/full#supplementary-material

REFERENCES

Blanke, O., and Metzinger, T. (2009). Full-Body Illusions and Minimal Phenomenal Selfhood. *Trends Cognitive Sciences* 13, 7–13. doi:10.1016/j.tics.2008.10.003

Botvinick, M., and Cohen, J. (1998). Rubber Hands 'Feel' Touch that Eyes See. *Nature* 391, 756. doi:10.1038/35784

Bovet, S., Debarba, H. G., Herbelin, B., Molla, E., and Boulic, R. (2018). The Critical Role of Self-Contact for Embodiment in Virtual Reality. *IEEE Trans. Vis. Comput. Graphics* 24, 1428–1436. doi:10.1109/tvcg.2018.2794658

Burns, E., Whitton, M. C., Razzaque, S., McCallus, M. R., Panter, A. T., and Brooks, F. P. (2005). "The Hand Is Slower Than the Eye: A Quantitative Exploration of Visual Dominance over Proprioception," in IEEE Proceedings. VR 2005. Virtual Reality, 2005, Bonn, Germany, 12-16 March 2005 (Bonn, Germany: IEEE), 3–10. doi:10.1109/VR.2005.1492747

Friston, K. (2012). The History of the Future of the Bayesian Brain. *NeuroImage* 62, 1230–1233. doi:10.1016/j.neuroimage.2011.10.004

Hay, J. C., Pick, H. L., and Ikeda, K. (1965). Visual Capture Produced by Prism Spectacles. *Psychon. Sci.* 2, 215–216. doi:10.3758/BF03343413

Heller, M. A. (1983). Haptic Dominance in Form Perception with Blurred Vision. *Perception* 12, 607–613. doi:10.1068/p120607

Holmes, N. P., and Spence, C. (2005). Visual Bias of Unseen Hand Position with a Mirror: Spatial and Temporal Factors. *Exp. Brain Res.* 166, 489–497. doi:10.1007/s00221-005-2389-4

Jeunet, C., Albert, L., Argelaguet, F., and Lecuyer, A. (2018). "Do You Feel in Control?": Towards Novel Approaches to Characterise, Manipulate and Measure the Sense of Agency in Virtual Environments. *IEEE Trans. Vis. Comput. Graphics* 24, 1486–1495. doi:10.1109/tvcg.2018.2794598

Knill, D. C., and Pouget, A. (2004). The Bayesian Brain: The Role of Uncertainty in Neural Coding and Computation. *Trends Neurosciences* 27, 712–719. doi:10.1016/j.tins.2004.10.007

Kreimeier, J., Hammer, S., Friedmann, D., Karg, P., Bühner, C., Bankel, L., et al. (2019). "Evaluation of Different Types of Haptic Feedback Influencing the Task-Based Presence and Performance in Virtual Reality," in Proceedings of the 12th acm international conference on pervasive technologies related to assistive environments, Rhodes Greece, June 5–7, 2019, 289–298. doi:10.1145/3316782.3321536

Lee, Y., Jang, I., and Lee, D. (2015). "Enlarging Just Noticeable Differences of Visual-Proprioceptive Conflict in Vr Using Haptic Feedback," in 2015 IEEE World Haptics Conference (WHC) (IEEE), Northwestern University, Evanston IL, June 22–26, 2015, 19–24. doi:10.1109/whc.2015.7177685

Maselli, A., Kilteni, K., López-Moliner, J., and Slater, M. (2016). The Sense of Body Ownership Relaxes Temporal Constraints for Multisensory Integration. *Sci. Rep.* 6, 30628. doi:10.1038/srep30628

Odermatt, I. A., Buetler, K. A., Wenk, N., Özen, Ö., Penalver-Andres, J., Nef, T., et al. (2021). Congruency of Information rather Than Body Ownership Enhances Motor Performance in Highly Embodied Virtual Reality. *Front. Neurosci.* 15, 678909. doi:10.3389/fnins.2021.678909

Power, R. P. (1980). The Dominance of Touch by Vision: Sometimes Incomplete. *Perception* 9, 457–466. doi:10.1068/p090457

Ramírez-Fernández, C., Morán, A. L., and García-Canseco, E. (2015). "Haptic Feedback in Motor Hand Virtual Therapy Increases Precision and Generates Less Mental Workload," in 2015 9th international conference on pervasive computing technologies for healthcare (PervasiveHealth), Istanbul, Turkey, 20-23 May 2015 (IEEE), 280–286.

Salomon, R., Fernandez, N. B., van Elk, M., Vachicouras, N., Sabatier, F., Tychinskaya, A., et al. (2016). Changing Motor Perception by Sensorimotor Conflicts and Body Ownership. *Sci. Rep.* 6, 25847. doi:10.1038/srep25847

Schwind, V., Lin, L., Di Luca, M., Jörg, S., and Hillis, J. (2018). "Touch with Foreign Hands: The Effect of Virtual Hand Appearance on Visual-Haptic Integration," in Proceedings of the 15th ACM Symposium on Applied Perception, Vancouver British Columbia Canada, August 10–11, 2018, 1–8.

van Beers, R. J., Wolpert, D. M., and Haggard, P. (2002). When Feeling Is More Important Than Seeing in Sensorimotor Adaptation. *Curr. Biol.* 12, 834–837. doi:10.1016/S0960-9822(02)00836-9

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.

frontiers | Frontiers in Virtual Reality

TYPE Original Research
PUBLISHED 02 December 2022
DOI 10.3389/frvir.2022.1056782

OPEN ACCESS

EDITED BY
Eike Langbehn,
Hamburg University of Applied Sciences, Germany

REVIEWED BY
Sandra Malpica,
University of Zaragoza, Spain
Andy Thomas Woods,
University of London, United Kingdom

*CORRESPONDENCE
Hai-Ning Liang,
haining.liang@xjtlu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Technologies for VR, a section of the journal Frontiers in Virtual Reality

RECEIVED 29 September 2022
ACCEPTED 22 November 2022
PUBLISHED 02 December 2022

CITATION
Wu Z, Shi R, Li Z, Jiang M, Li Y, Yu L and Liang H-N (2022), Examining cross-modal correspondence between ambient color and taste perception in virtual reality.
Front. Virtual Real. 3:1056782.
doi: 10.3389/frvir.2022.1056782

Examining cross-modal correspondence between ambient color and taste perception in virtual reality

Zhen Wu[1†], Rongkai Shi[2†], Ziming Li[2], Mengqi Jiang[3†], Yue Li[2], Lingyun Yu[2] and Hai-Ning Liang[2*]

[1]Division of Integrative Systems and Design, Hong Kong University of Science and Technology, Hong Kong, Hong Kong SAR, China, [2]Department of Computing, Xi'an Jiaotong-Liverpool University, Suzhou, China, [3]School of System Design and Intelligent Manufacturing, Southern University of Science and Technology, Shenzhen, China

This research explores the cross-modal correspondence effect of ambient color on people's taste perception in virtual reality (VR). To this end, we designed and conducted two experiments to investigate whether and how taste-congruent ambient colors in VR influence taste perception measured by four aspects: 1) taste ratings of a neutral drink; 2) taste association with virtual environments; 3) associated scenarios when immersed in these virtual environments; and 4) participants' liking of these environments. In Experiment 1, participants adjusted the ambient light with different cross-modal-related colors in the immersive environments and reported their scaling of the Virtual Reality Sickness Questionnaire (VRSQ). Comfortable light intensity for each ambient color was obtained and color recognition problems were observed. In Experiment 2, participants tasted black tea (as the neutral drink), after being exposed to eight different virtual environments with different ambient colors. Results showed that the pink ambient color significantly increased the sweetness ratings. Differences in the color-taste association and environment liking were also observed in the ambient color conditions. Our results provide new insights into the cross-modal correspondence effect on ambient color and taste perception not found in prior work in VR scenarios.

KEYWORDS

virtual reality, human-food interaction, cross-modal correspondence, visual perception, user study, tasting

1 Introduction

Cross-modal correspondence (CMC) is the interaction among various senses, where people map sensations in one sensory modality onto other modalities in a consistent manner (Spence, 2011). Taste is one of the most common sensations that people experience daily. Basic taste perceptions include sweetness, sourness, bitterness, saltiness, and umami[1] (Keast and Breslin, 2003). Prior studies have explored how taste perception can be formed with other senses *via* CMC (Knöferle and Spence, 2012; Kerruish, 2019; Halabi and Saleh, 2021). Existing research has provided evidence that color is an intrinsic and extrinsic visual stimulus influencing taste. Intrinsic colors are those features of the food itself, including the color of the food (Maga, 1974; Johnson and Clydesdale, 1982; Clydesdale. et al., 1992; Charles et al., 2010; Spence and Piqueras-Fiszman, 2016). Extrinsic colors are those external to the food, including the influence of color itself (Koch and Koch, 2003; Wan et al., 2014), the color of the container (like the cup or the bowl) (Harrar et al., 2011; Piqueras-Fiszman and Spence, 2012; Spence and Velasco, 2018; Sugimori and Kawasaki, 2022), and the color of the environment (Oberfeld et al., 2009).

Recently, there has been an increasing trend of using relatively newer interactive technologies to enhance or modify taste perception (Gallace et al., 2012; Moser and Tscheligi, 2013; Nishizawa et al., 2016; Arnold, 2017; Kerruish, 2019; Halabi and Saleh, 2021). For example, Narumi et al. (2011) proposed MetaCookie+, a pseudo-gustatory system that uses an augmented reality head-mounted display (AR HMD) to overlay both visual and olfactory information on real cookies to observe taste alterations. Another work by Nishizawa et al. (2016) introduced a projective-AR system to modify the flavor and taste of chips by changing their hue properties. Besides AR environments, researchers found immersive virtual reality (VR) environments can also influence people's food or beverage taste experience. VR is regarded as a new medium to provide visual stimulation and is gradually becoming a research tool for the discipline of perception (Kuliga et al., 2015; Lipson-Smith et al., 2020). Previous works show that this medium has several benefits. The users can have higher engagement and taste perception consistency in an immersive virtual environment compared to a physical environment (Carlson et al., 2011; Ledoux et al., 2013; Bangcuyo et al., 2015; Stelick et al., 2018). VR also presents a cost-effective way to provide various stimuli (Lipson-Smith et al., 2020) and thus can be used to conduct empirical studies on psychology or perception.

Researchers have utilized this medium to study the effects of color on taste perception, especially the beverage's (intrinsic) color on its tasting. For example, research from Wang et al. (2020) modified the color of the coffee in VR and Huang et al. (2019) studied the influence of tea color simulation in VR. In addition to changing the color of the food, Chen et al. (2020) studied the synthesis of multiple visual elements, with the environmental color as one of the variables. Although these works explored the feasibility of conducting taste perception studies related to color in virtual environments, they either focused on manipulating the intrinsic color of the foodstuff or dealing with complex VR scenes with various visual elements or items than colors only. Compared with the modification of the intrinsic colors or textures, the focus on environmental color may be more suitable as VR environments emphasize the visual immersion of users. However, from Chen et al. (2020)'s experiment, we can not decide whether the change in taste perception was because of the color or other visual stimuli. A simple and specific type of extrinsic visual stimulus from the virtual environment is worthy of investigation because we can then isolate the effect of specific visual aspects of such environments and leverage the findings to develop VR environments tailored to more specific purposes.

In this work, we investigate the influence of ambient light color variations in virtual environments. This type of stimulus has been studied in *physical settings*, presenting a significant influence on taste perception (Oberfeld et al., 2009; Spence et al., 2014), but it has not been explored in *VR environments*. Ambient light color is ubiquitous in such environments and fits the characteristic of immersion as an extrinsic indirect visual cue. However, limited research has investigated how could this parameter in immersive virtual environments influences taste perception. To fill this research gap in our understanding of how ambient light affects taste perceptions in virtual environments, we run two experiments that help answer two research questions (RQs).

RQ1: What will the suitable environment parameter be when we apply the cross-modal-related colors to the VR environments? is the main research question explored in Experiment 1. Previous work revealed two potential problems that color studies in VR environments may suffer: **1)** participants may experience visually induced motion sickness (VIMS) due to the unpleasant colors (Bonato et al., 2004; So and Yuen, 2007; Gusev et al., 2018; Grassini et al., 2021); and **2)** there may be a recognition deviation between the color of the ambient light presented through the VR HMD and the color perceived in the physical setting. Based on these, we conducted Experiment 1 to evaluate the impact of these issues and set up the ambient light conditions to be suitable for color-taste studies. We formulated the testing virtual environments with 11 ambient colors, including black, green, grey, orange, pink, purple, red, yellow, white, blue, and brown. Participants adjusted the light intensity of these environments to make them comfortable, then reported their perceived VIMS and recognized color when exposed to the environment. Our findings show that brown, white, and grey ambient lights were considered

1 Umami is a category of taste in food (besides sweet, sour, salt, and bitter), corresponding to the flavor of glutamates, especially monosodium glutamate.

to be more comfortable. However, these three colors were not well recognized and thus were excluded in Experiment 2.

RQ2: How will taste perception be influenced by different ambient colors? is the research question for Experiment 2, where we further studied the influence of the ambient color setups *via* Experiment 1 on perceived taste in VR environments. For this research question, we investigated the following three aspects of taste perception: **1)** participants' perceived taste of the tastant in terms of sweetness, sourness, bitterness, and saltiness; **2)** participants' taste association under the stimuli together with the associated scenarios that participants imagined; and **3)** participants' preference of the tested virtual environments where they tasted the tastant. According to Wan et al. (2014), we excluded umami in our study because it did not show a strong association with any color, shape, or texture in real-life settings. Moreover, the term umami may be difficult to understand due to its inconsistent perceptions across different cultural backgrounds (Gotow et al., 2021). We selected black tea as our neutral tastant based on a pilot study. Our findings reveal that ambient color significantly influenced the sweet taste of the tastant. Pink and purple showed a significant association with the sweet taste rather than the bitter taste while black and green were associated with bitterness more frequently compared with sweetness. In terms of the associated scenarios coupled with the color-taste association, most of the participants associated the sweet taste with food-related items, but the bitter taste with specific scenes. Finally, orange ambient light was preferred over red.

Our work presents two main contributions:

- An experiment protocol that uses ambient color for taste perception studies in virtual environments.
- Evidence that ambient color is a significant factor for perceived sweetness and bitterness sensations in VR.

2 Related work

2.1 Taste perception in virtual environments

Taste experience can be introduced to different types of virtual environments, ranging from curved computer-based screens to AR/VR HMDs (Wang et al., 2021). There is an increasing trend of using virtual environments in human-food interaction research because they allow researchers to easily add environmental features or properties and minimize certain negative effects of real settings (Bangcuyo et al., 2015; Delarue et al., 2019). For instance, compared to traditional booths, the hedonic data collected in a virtual coffee house can be more identifiable and reliable as a predictor for future coffee evaluation (Bangcuyo et al., 2015). Delarue et al. (2019) have shown that a multi-sensory immersive room could minimize the differences in product testing results produced at different periods in a lab setting. Moreover, introducing the taste experience to VR also benefits the immersive experience. Eating and drinking unify both virtual and physical sensory cues (Spence, 2016). Thus, they can increase users' presence in the virtual environments, i.e., their sense of being there (Slater and Wilbur, 1997). Wang et al. (2021) also suggested that introducing physical interaction mechanisms that the users are used to in the physical world into the immersive virtual environments can enhance the psychological feeling of presence. Furthermore, Gallace et al. (2012) stated that increasing the number of senses stimulated in VR can increase users' enjoyment, memorability, and presence. Since taste is one of the major sensory modalities familiar to people in real-life settings, it is beneficial to have it in VR. Eating and food-orientating activities may also become integral to the virtual experience in social contexts as VR becomes more available and allow people to meet on platforms like VRChat[2] for socialization and Horizon Workrooms[3] for business meetings.

On the other hand, taste-related results in virtual environments may be different from those obtained in physical environments. For example, the virtual environment may significantly influence users' hedonic responses to foodstuffs. Beef can be rated higher in terms of liking when consumed in a VR restaurant compared to a traditional sensory laboratory condition (Crofton et al., 2021). Stelick et al. (2018) also found that VR had a significant influence on the taste perception of blue cheese, especially the level of pungency. Ledoux et al. (2013) showed that food craving produced by VR was marginally greater than a neutral cue in a physical setting.

2.2 Cross-modal effect between color and taste

Cross-modal correspondence (CMC) is about the interactions among different senses. People combine information captured from various senses to form a more comprehensive view of the external world or activity they are doing (e.g., when people see and smell food in front of them). Interaction between these systems is an essential part of this process (Stein and Meredith, 1993). In particular, based on CMC, people can map features or sensations in one sensory modality onto features of other modalities in a consistent manner (Spence, 2011). For example, the taste of sweetness can correspond to a round shape (visual) or high pitch (audio). Surprisingly, CMC is not confined to one group (cultural or otherwise) but is often universally shared (Knöferle and Spence, 2012). Taste is one of

2 https://hello.vrchat.com/.

3 https://www.oculus.com/workrooms/.

the most common sensations that people experience daily. Basic taste perceptions include sweetness, sourness, bitterness, saltiness, and savor (umami) (Keast and Breslin, 2003), which can often be formed with other senses, such as smell or sound (Knöferle and Spence, 2012; Kerruish, 2019; Halabi and Saleh, 2021).

Color is an essential aspect of CMC between vision and taste. Color-taste correspondence is of great interest to the food industry, especially in helping design food product packaging because of its direct influence on consumers' expectations (Cardello, 1994; Koch and Koch, 2003). To a large extent, it serves as a sensory property of the product before its consumption (Spence and Velasco, 2018). Color is a focused factor that significantly affects taste perception (Piqueras-Fiszman and Spence, 2014). For example, from the studies of color itself, Koch and Koch (2003) have stated that red is connected to sweetness while yellow is connected to sour, citrus, and fruity tastes. In addition, Wan et al. (2014) have shown that some strong CMCs are similar in different cultures, such as mapping black with bitterness and pink with sweetness.

Factors influencing multi-sensory flavor perception can be divided into either *intrinsic* to the food (e.g., texture, aroma) or *extrinsic* to it (e.g., related to the packaging, receptacle, or external environment) (Wang et al., 2019), both of which have been studied by previous research on color-taste CMC. When color represents the original feature of the foodstuff, it is considered an intrinsic visual stimulus feature (Spence, 2019b). For example, Johnson and Clydesdale (1982) found that applying red color to water could increase the sweet taste. Similarly, Maga (1974) showed that green statistically increased sweet taste threshold sensitivity while yellow decreased taste sensitivity. Different from intrinsic features, extrinsic factors are outside of the food itself. For example, Piqueras-Fiszman and Spence (2012) have shown that orange and dark-cream cups could enhance the chocolate flavor of a drink. Harrar et al. (2011) investigated popcorn in four different colored bowls and found that the sweet popcorn, in addition to being sweet, was perceived as saltier when consumed in a colored bowl than in a white bowl.

Specifically, we noticed that the color of ambient light is a type of extrinsic factor that can influence taste. Spence and Carvalho (2020) stressed the importance of the environment where the beverages are consumed. The ambient color of an environment could significantly affect the beverage's flavor and taste (Oberfeld et al., 2009). Oberfeld et al. (2009) showed that wine was perceived to be spicier in blue or green ambient lighting than in red or white. Spence et al. (2014) explored people's taste perception of red wine in a room that changes ambient color and music and found a significant difference in taste and beverage liking. Their results provide empirical support for the claim about the influence of ambient color on flavor from Oberfeld et al. (2009).

In general, there are three possible mechanisms underpinning the CMCs (Spence, 2011; Parise and Spence, 2013): 1) *structural correspondence*, described from the neural science side, is the byproduct of the innate cognitive system; 2) *statistical correspondence*, describes a learned process associating the established experience and the environment; and 3) *semantic correspondence*, stresses that the same terms can be used to describe different stimuli linguistically. For color-taste CMC, there are mainly two types of specific explanations for their correspondence. The first type is the association with the source object(s), meaning that the color-taste pairs result from associating specific colors with certain gustatory cues (Wang et al., 2020). Moreover, Spence and Levitan (2021) found that some color-taste associations are not limited to a particular source object. The second explanation is users' emotional mediation. For example, the pink-sweet association arises because both stimuli are linked with happiness independently and are connected by people (Spence and Levitan, 2021). It is important to note that some colors with positive emotional valance in abstract scenarios may lead to negative emotions when matching with food. For instance, blue is a favorite color to many people, but blue-colored food can trigger negative feelings (Spence, 2019a).

2.3 Color-taste study in VR

The color-related research in the physical world has inspired some researchers to apply them in VR. Wang et al. (2020) designed an experiment manipulating the color of coffee in VR so that their participants saw either a dark brown or light brown liquid as they consumed the real coffee. Similarly, Huang et al. (2019) designed three experiments to study the influence of actual tea color simulation in VR on users' taste ratings of Chinese red and green tea. Moreover, Chen et al. (2020)'s study went beyond colors. They created a sweet, bitter, and neutral scene integrating shape, color, and visual textures, and assessed their influence on perceived sweetness and environment liking. Results showed that the sweet-congruent environment significantly increased the perceived sweetness of the beverage they tested. However, these studies either focus on manipulating the intrinsic color of the foodstuff in VR or dealing with complex VR scenes.

As discussed in the previous section, the color of the ambient light is an effective stimulus studied in the physical environment. This inspired us to consider if similar effects can be reproduced and leveraged in VR environments. The virtual environments may cost-effectively control or simulate the ambient color rooms while providing fully colored immersive environments. Firstly, the color of the ambient light is a simple parameter that can be controlled in VR and be compared with previous works done in the physical setting. Although not related to taste, researchers have already suggested using VR for ambient color studies. For example, Oberfeld and Hecht (2011) studied the influence of environmental color on the perceived height and width of interior rooms virtually. von Castell et al. (2018) also studied the effects of the hue, saturation, and luminance of color on the perceived height of interior spaces *via* a VR HMD. Lipson-Smith et al. (2020) did a study focusing on the mood and preference towards the virtual environments varying in colors. In the meantime, applying ambient light as an environmental parameter to VR leverages the immersion provided by this medium well. In VR development tools, ambient light hue is an important parameter in creating the visual appearance of the scene and its objects. Changing ambient light hue can effectively alter the color scheme of all the 3D models (objects) in the VR scene (Unity, 2019). Based on these, we chose ambient light varying in colors as our stimuli to study its influence on taste perception in VR. While ambient light is essential to and a foundation building block to all VR environments, its effect on taste perception has not been explored in detail. Our research aims to fill this important but underexplored area.

2.4 Potential issues for VR color-taste studies

In a VR environment, color may induce Visually Induced Motion Sickness (VIMS) and make the application unusable (Gusev et al., 2016). Bonato et al. (2004) simulated the optokinetic drum (spinning sphere) with specific color patterns in a virtual environment. They suggested that the chromaticity might increase the inconsistency between visual and vestibular inputs. However, So and Yuen (2007) found the change in the colors would not influence VIMS significantly in VR. Besides, colored environments may lead to a higher VIMS compared to monochrome environments (Gusev et al., 2018). Based on the literature, it is not clear whether the ambient color in VR would cause VIMS and, if so, to what extent the VIMS would be. To avoid discomfort and bias in our color-taste experiment, we conducted the first experiment to adjust the ambient colors in the virtual environments to ensure they were not harmful to the experience. Color recognition can be another potential issue for VR color-taste studies. It is important to consider whether people could depict the ambient color hue in a VR environment as they perceive it in the real world (Lipson-Smith et al., 2020). We also want to filter out this issue before running the color-taste experiment.

Selecting a suitable tastant is also important for a color-taste study. Some of the existing studies in CMC between taste and other stimuli used word-based descriptions or images of the food. Crisinel and Spence (2009) used Implicit Association Test and applied food images as visual stimuli to investigate the association between basic taste and pitch (auditory parameter). Similarly, Wan et al. (2014) used graphics with words referring to taste as the visual stimuli in their online study. The participants dragged the image to the word they thought it was associated with. On the contrary, Crisinel and Spence (2010) did a repeated experiment using real tastants. However, they got similar taste perception results as those evoked by the food names. As reported by Crisinel and Spence (2009) in their work, although it was convenient to use the names of food or drinks in the study, it entailed several disadvantages. First, between-participant differences might exist, such as the same word "beer" may associate with various distinctive tastes. Second, excluding real tastants prevented them from studying the direct interaction between the sensory modalities. As such, in our study, we did not use words and instead used a real tastant because this is more aligned with the use of VR, as a platform that brings users to experience virtual environments but allows them to still be closely associated with the physical world.

Previous studies in taste and virtual environments have utilized a variety of tastant choices. We collected some representative ones and listed also in Table 1. Among these, we found that a beverage is the most frequently used tastant. In addition, drinking is relatively more convenient and safer than eating. Thus, we used beverages as our taste stimuli. We anticipated that a beverage with a very strong taste would lead towards one specific taste and bias the results. Due to this, a neutral tastant that would not affect participants' taste perception is required. However, limited research has discussed the rationale behind choosing the tastants and whether their tastes are neutral. We decided to run a pilot study to choose a neutral tastant before our formal color-taste experiment.

2.5 Summary

Table 1 summarizes the related works mentioned above. This study has three main differences compared to prior studies:

1) Unlike prior work about taste perception in immersive virtual scenes (Section 2.1), our focus was CMC-related stimuli instead of realistic scene simulation.
2) In terms of research about visual-taste CMC in the physical environment (Section 2.2), our work is different from the experiments from Koch and Koch (2003) and Wan et al. (2014) because our work included a real tastant. In addition, our work differs from studies dealing with ambient light (e.g. Oberfeld et al., 2009; Spence et al., 2014). We introduce VR as a platform to determine its level of visual-taste CMC. To do this, several changes and considerations are required but not found in prior research. The use of VR as a medium may bring interdisciplinary insights to applications in VR, which has grown rapidly in the last few years and is positioned to shift significantly how we interact with others, whether for

TABLE 1 A summary of related work about taste and the immersive environments with the used stimuli. Tastants are highlighted in bold and tested stimuli were marked in italic.

References	Description	Main difference between this work and our research
Taste perception in immersive virtual scenes (see also Section 2.1)		
Crofton et al. (2021)	Studying the impact of different virtual environments on people's hedonic ratings of beef and chocolate. Scenarios were *traditional sensory booths, a VR restaurant, a VR Irish countryside, and a VR busy city*	(1) Their focus was specific scene simulation instead of CMC-related variables. (2) Their emphasis was on users' hedonic experiences instead of taste
Stelick et al. (2018)	Adapting VR to sensory evaluation *via* a pilot test to collect participants' feelings on blue cheese in virtual contexts using *Custom-recorded 360-degree videos of a sensory booth, a park bench, and a cow barn in a VR HMD.*	Same as above
Bangcuyo et al. (2015)	Studying the use of immersive techniques in consumers' hedonic teste of (whole bean) coffee in *virtual coffeehouses with or without contextual information*	Same as above
Delarue et al. (2019)	Studying the experience of alcohol-free beer in *multi-sensory immersive rooms simulating a nightclub and a beach*	Same as above
Visual-taste cross-modal correspondence in the physical environment (see also Section 2.2)		
Koch and Koch. (2003)	Examining users' preconception of taste, using *10 colors and 8 tastes words in the form of scale-based questions on paper*	(1) We included a real tastant. (2) We did the color experiment in virtual environments rather than through a computer screen or paper-based questionnaires
Wan et al. (2014)	Investigating the crossmodal-correspondences between visual features (including colors, shapes, textures) and basic taste words, using *Text descriptors and images of the visual stimuli, presented against a gray background*	Same as above
Sugimori and Kawasaki. (2022)	Investigating the effect of background color on taste using chocolate and green tea with varying bitterness levels. Tested stimuli were *black or pink wrapping paper* for chocolate and *clear blue cup or clear cup* for green tea	We did the experiment in a virtual environment seeing through a VR HMD and focused on ambient light
Harrar et al. (2011)	Investigating whether the color of the bowl would affect the taste of the popcorn it contains. Comparing *Four different colored bowls: white, green, red, and blue*	Same as above
Oberfeld et al. (2009)	Investigating the effect of the ambient room color on the flavor of wine, while maintaining the color of the beverage. Comparing four types of *ambient lighting vary in colors*: blue, green, red, white	(1) Although we both focused on ambient colors, we introduced VR and took into account aspects not explored before. (2) The tested parameters of this work were focused on wine (e,g, fruitiness, spiciness). Our work is more general to explore basic tastes
Spence et al. (2014)	Investigating the combination of ambient light and music on the wine drinking experience. Tested stimuli were: *(1) white, red, green lighting with sour music, and red lighting with sweet music. (2) white, green, red lighting with sweet music, and green lighting with sour music*	Same as above and we focused on ambient light only and not specific to wine
Visual-taste cross-modal effect in immersive virtual environments (see also Section 2.3)		
Narumi et al. (2011)	Investigating the effectiveness of augmented reality flavor gustatory display based on edible cues and cross-modal interaction (using cookies as the tastant) by *overlaying visual and olfactory information* onto a real cookie with special AR marker patterns	We modified the extrinsic color (an environmental factor) as opposed to the intrinsic color of the foodstuff

(Continued on following page)

TABLE 1 (*Continued*) A summary of related work about taste and the immersive environments with the used stimuli. Tastants are highlighted in bold and tested stimuli were marked in italic.

References	Description	Main difference between this work and our research
Huang et al. (2019)	Assessing the influence of *virtual color of Chinese green and red tea in a VR HMD* on the taste of Chinese tea	Same as above
Chen et al. (2020)	Investigating the impact of visual-taste congruency on perceived sweetness and product liking in immersive VR (using syrup diluted with water as the tastant). Stimuli were *rooms with sweet-congruent, bitter-congruent, and neutral visual cues in a VR HMD.*	Same as above
Wang et al. (2020)	Exploring whether making coffee look milkier in a VR environment can alter its perceived flavor and liking. Stimuli were *dark brown or light brown coffee in a VR HMD.*	Our choice of stimuli is more generic and focuses on environmental colors instead of including other visual elements together
Our experiment	Examining cross-modal correspondence between *ambient colors* and taste perception in virtual reality, using black tea as the tastant, considered a neutral tastant by participants and including 8 ambient light colors in a VR environment	

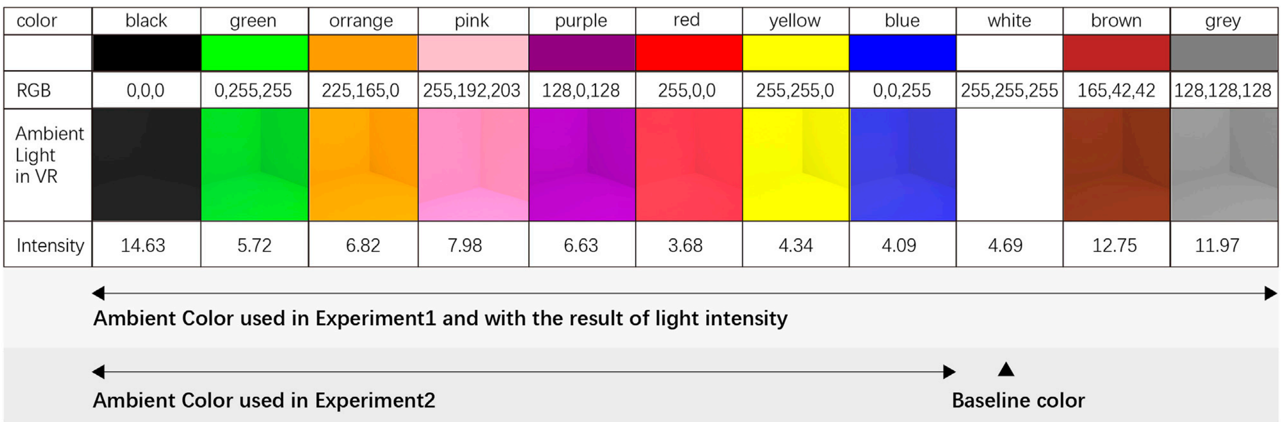


FIGURE 1
Ambient color stimuli that were utilized in this study, with their RGB values, ambient light effects in VR, and the suitable light intensities obtained from Experiment 1. The first row with an arrow represents the colors tested in Experiment 1. The second row indicates the colors in Experiment 2, with white as a baseline color.

social occasions (e.g., in platforms such as VRChat) or for production purposes (e.g., within Horizon Workrooms) that aim to emulate many aspects of the physical world.

3) Unlike previous work on taste-related CMC in VR (Section 2.3), our selection of ambient colors as the visual stimuli is novel and has not yet been explored systematically before. Ambient colors are foundational to VR applications due to their emphasis on visual immersive and focus on extrinsic modification. Our understanding of how ambient colors affect people's perception of taste will enable us to develop VR applications that can be more tailored to specific situations.

3 Experiment 1: Adjustment of the ambient color in the virtual environments (RQ1)

In this experiment, we aim to explore and determine the suitable settings of different ambient color environments, which is the premise for their usage in the follow-up taste perception study. As mentioned in Section 2.4, color in VR environments may cause visually induced motion sickness (VIMS) (Bonato et al., 2004; So and Yuen, 2007; Gusev et al., 2018; Grassini et al., 2021). Directly adapting the colors from former color-taste studies in physical settings to the ambient color in virtual

FIGURE 2
Experiment 1 setup. A participant is adjusting the light intensity of pink ambient color in the virtual room displayed *via* an VR HMD, which is connected to a laptop.

environments with the default light setting may cause discomfort to users. Adjustments to the related lighting parameters for these colors with high visual saturation are necessary. In addition to the hue (i.e., RGB values), there are several parameters influencing the visual presentation and rendering of the environment: ambient mode, compression, environment reflections, light intensity, and bounces in Unity engine (Unity, 2019). We set our main adjustment parameter as **light intensity** since it directly influences the brightness of the ambient light in the VR scene. Also, as mentioned in Section 2.4, it is uncertain whether color recognition in VR deviates from the physical world, which might lead to ambiguous or biased results in color-taste studies. Thus, we also tested the recognition accuracy of the ambient color in VR in Experiment 1.

3.1 Stimuli

We utilized the colors described in Wan et al. (2014)'s taste-color synesthesia, including black, green, grey, orange, pink, purple, red, yellow, white, blue, and brown. These colors were used as the hue of the ambient light in the VR environments. The RGB values of these colors are also listed in Figure 1. Specifically, we created 11 hexagonal virtual rooms to demonstrate these ambient colors (see Figure 1). In each room, we placed a spotlight at the top middle of the room to make sure it illuminated the room (i.e., the ambient color covered the entire interior of the room), and assigned the RGB value of the color. The spot angle of the light was 129.8°. We set the range for controlling how far the light is emitted from the center of the object from 0 to 20. The default light intensity for each room is 10. We left the rooms empty to prevent any confounding factors.

3.2 Participants and apparatus

Twelve healthy participants (6 females, 6 males; mean age = 21.2 ± 1.25 years) from a Sino-Chinese university located in a mid-size city in China were recruited for the first experiment. The participants were not paid and joined the experiment voluntarily. They did not have color blindness and did not report any issues regarding color recognition. Seven participants had prior experiences with VR HMD.

We used an Oculus Rift S VR HMD to display the virtual environments. The HMD has a resolution of 1,280 × 1,440 px per eye, a refreshing rate of 80 Hz, and a 115° field of view. Participants could move their heads in 6DOF. It was connected to an Intel Core i7-6700HQ processor laptop with an NVIDIA GeForce GTX 1070 GPU card. Participants could interact with the virtual environment *via* an Oculus hand-held controller. The virtual environments were created using the Unity3D platform (version 2019.1.14f1) and the related scripts were programmed in C#. The experiment was conducted in an empty and quiet room where the participants were equipped with the VR HMD and controllers. Figure 2 shows the setup.

3.3 Experimental design and task

A within-subjects design was used in this experiment. Participants would enter all the virtual rooms to complete the given tasks. When first entering a room, the initial intensity of the tested ambient light was set at 10. Participants were required to adjust the light intensity up and down starting from this value until they found the lighting condition that was the most comfortable. For experimental purposes, we set the adjustment range of the light intensity from 0 (dimmest light) to 20 (brightest light for our 3D assets). Participants could press Button A or B to make small adjustments to the light intensity (increased or decreased by 0.5 per press) and push the thumbstick to make big adjustments (see Figure 2). Once they settled down the light intensity for the current room, they stopped interacting with the controller for about 10 s. The finalized light intensity was then recorded by the program. After that, they needed to answer the question "what ambient color do you think is the current environment?" and complete a Virtual Reality Sickness Questionnaire (VRSQ) (Kim et al., 2018) in oral form. The experimenter asked these questions from VRSQ one by one and manually recorded the answers. Then the participants moved to the next room. The order of the room for each participant was counterbalanced using a Latin-Square design. Note that in our experimental design, participants' answers toward color recognition might not have a relationship with their adjustment of light intensity. Although each participant had encountered several levels of intensity, they were exposed to the same intensity condition at first and were

10.3389/frvir.2022.1056782

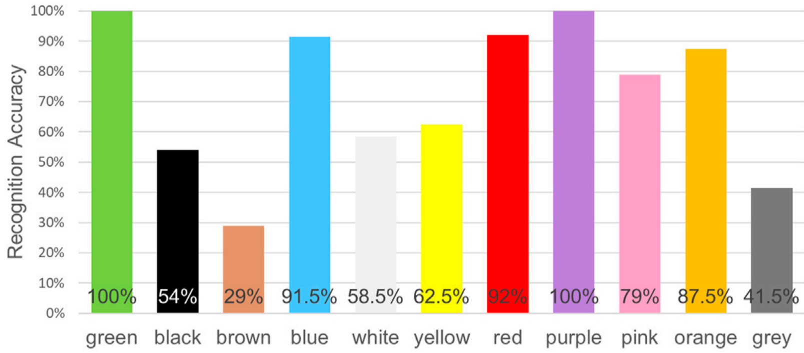


FIGURE 3
Recognition accuracy (in percentage) for each ambient color.

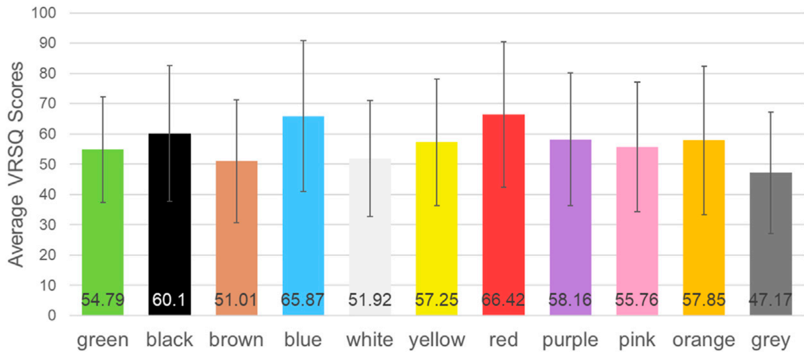


FIGURE 4
Average VRSQ scores for each ambient color, the error bars represent the standard deviations.

informed that they would make the adjustment for one specific color each time.

3.4 Procedure

The whole experiment lasted approximately 30 min for each participant. Before the formal experiment began, participants were asked to fill in a questionnaire for demographic background. We asked about their prior VR experience and if they had difficulty recognizing colors. Then participants were introduced to the research aims, tasks, controls of the experiment, and the questions that would be asked. Then they wore the VR HMD and started the experiment. After they adjusted the light intensity for all the ambient colors, they would take off the HMD and have a break. After the break, we asked participants to adjust the light intensity for the ambient colors in reverse order. This was intended to eliminate the residue effects from the previous colored room.

3.5 Results and discussion

The average light intensities for each ambient color are 14.63 for black, 5.72 for green, 11.97 for grey, 6.82 for orange, 7.98 for pink, 6.63 for purple, 3.68 for red, 4.34 for yellow, 4.69 for white, 4.09 for blue, and 12.75 for brown. As mentioned before, the values are between 0 (the dimmest condition) to 20 (the brightest condition).

In total, we collected 24 color recognition results (12 participants × 2 rounds). Figure 3 summarizes the recognition accuracy for each ambient color. The recognition accuracy for the ambient color of green, blue, red, purple, pink, and orange was high. However, the ambient color of brown and grey was low (29% and 41.5%, respectively). Though not low in

FIGURE 5
Tastant choices and the container of the tastant.

recognition accuracy, white was frequently mixed with light pink (4 times). For grey and brown, there was a grey-green, grey-blue confusion, and a brown-pink confusion. Figure 4 is the summary of the average VRSQ score for each ambient color following the calculating method provided by Kim et al. (2018). A higher score indicates that the simulator sickness caused by the ambient color environment is more severe. Participants regarded brown, grey, and white as more comfortable than other colors in the VR environment after self-adjusting the light intensity. However, the difference was not significant.

Colors with low recognition accuracy were excluded from the follow-up experiment. We excluded white as it was mixed with pink while keeping pink since pink could be recognized successfully in the virtual environment. Grey and brown were excluded for the same reason. Although the color recognition accuracy of black was not high, it is an important color showing a high correlation with the bitter taste in the physical environment. So, it was useful to keep it and see its effectiveness in the virtual environment.

The results of recognition accuracy show that gaps may exist between the ambient colors that a developer wants the audience to perceive and the actual perceived color by this audience, especially when creating virtual environments. Developers are suggested to pay attention to color recognition accuracy, especially when color is an important factor in the virtual environment. Another contributing result was the comfortable intensity for the 11 ambient colors in the virtual environment with their VRSQ value, working as a guideline for developers who want to utilize these colors in their designs.

4 Experiment 2: CMC between ambient color and taste in virtual environments (RQ2)

In this experiment, we investigated how ambient colors would influence users' perception of taste in immersive virtual environments, including sweet, sour, bitter, and salty tastes, corresponding to **RQ2**. As mentioned in Section 2.4, we needed a real tastant for the color-taste experiment rather than using words describing different tastes. In addition, the tastant should be as neutral as possible in its perceived taste so that it would not affect participants' taste perception. Thus, we run a pilot study to find a neutral tastant before the formal experiment.

4.1 Pilot study

We invited the same 12 participants in Experiment 1 to complete this pilot study. The participants filled in a demographic questionnaire before the formal experiment. In the questionnaire, we inquired about their attention to taste in the form of scaling questions from 0-low to 6-high, "to how much extent do you pay attention to taste during your daily dieting?". Their average ratings on taste attention were: 5.08 for sweetness, 4.69 for sourness, 3.38 for bitterness, and 5.06 for saltiness. We prepared four common beverages, including apple juice (brand: Huiyuan Juice), black tea (brand: Nongfu spring), coconut water (brand: Goodfarmer), and yogurt soda (brand: Energetic Forest), all of which could be bought at a local grocery shop. Our main consideration was to provide several possible choices of drinks that do not have a strong taste and are also familiar to users. After some trials before the pilot study with a number of possibilities, the four chosen ones represent a wide range. The pilot study then helped us narrow done to one.

We opted to use mini paper cups (100 ml) to ensure that it was convenient to drink the beverage while wearing an HMD in the main experiment later. The cups were covered with lids to prevent participants from recognizing the beverage by seeing the content or by smelling the odor (see Figure 5). Participants were asked to drink each of the four beverages, holding it in their mouth for at least 3 s before swallowing it. Then they needed to rate the taste of the beverage in terms of sweetness, sourness,

 10.3389/frvir.2022.1056782

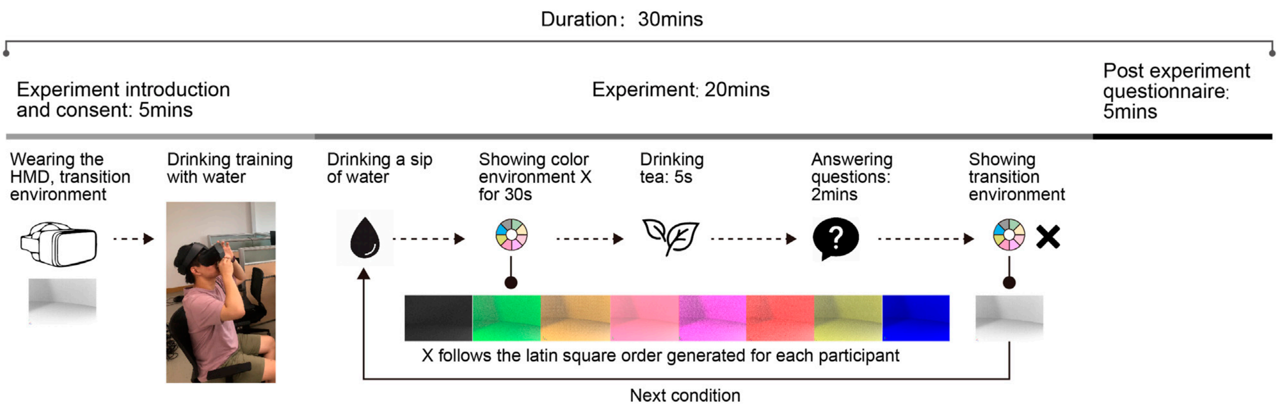


FIGURE 6
Overview of experiment 2.

bitterness, saltiness, and neutrality using a 7-scale questionnaire. Water was given to participants between two beverages to remove any residual, left-over taste. The sequence of the delivered beverages was counterbalanced. Note that, water was not used as the tastant because it was used as the "palate cleanser" to clear participants' taste from the previous drink, and as such using it as the tastant is not appropriate (Huang et al., 2019; Chen et al., 2020; Wang et al., 2020).

Black tea was rated the most neutral among the four beverages. The remaining tastants have strong sweet tastes based on the ratings. Therefore, we used black tea as the taste stimulus in our experiment.

4.2 Stimuli

We removed the colors showing low recognition accuracy from our Experiment 1 (as described in Section 3.5). Thus, the color stimuli in the main study included black, green, orange, pink, purple, red, yellow, and blue. Although white was excluded, we set it as a starting, ending, and transition color for the experiment to avoid the residue effects from the previous colored room. A similar approach can be found in Spence et al. (2014). Figure 1 shows a preview of each virtual environment with the RGB parameter and intensity level of each ambient color. Notice that we did not include the 3D model of the beverage in the virtual environment because we wanted to focus on the influence of the immersive environment only. We used black tea as the tastant based on the results of the pilot study. Like the pilot study, the tastant was served in 100 ml paper cups covered with lids.

4.3 Participants and apparatus

We recruited another 16 participants (7 females, 9 males; 22.2 ± 1.66 years, ranging from 21 to 27 years). Thirteen participants had used VR HMD before the experiment and all of them had normal or corrected-to-normal vision and had no history of color blindness. Regarding their drinking behaviors, thirteen participants reported drinking beverages several times a week and twelve participants pay much attention to the taste of the beverage during their daily drinking activities. We used the same VR apparatus as in Experiment 1.

4.4 Design and task

Figure 6 provides an overview of Experiment 2. We applied a within-subjects design in this experiment with ambient color as the independent variable. As mentioned, we used a white room as a starting, ending, and transition room between the two rooms. The order of ambient color conditions was counterbalanced with a Latin-Square design. In each condition, the participants would be exposed to the virtual room for 30 s, starting from when they entered the scene and ending when they heard a notification to drink the tastant. They were asked to have a sip from the 100 ml tastant in their mouths for 3 s before swallowing it. After they drank the tastant, they were then asked to report their feelings according to the following questions:

- **Q2-1.** Can you rate the taste of the drink in the just-finished scene in terms of sweet, sour, bitter, and salty? Ratings are from 1 (very weak) to 7 (very strong).
- **Q2-2.** What type of taste among sweetness, sourness, bitterness, and saltiness do you think matches the VR environment the most? You can also choose none if you cannot come up with an association.
- **Q2-3.** Do you have some scenarios in your mind when coming up with this association when answering the second question? You should first answer yes or no, and then provide a specific answer if you say yes.

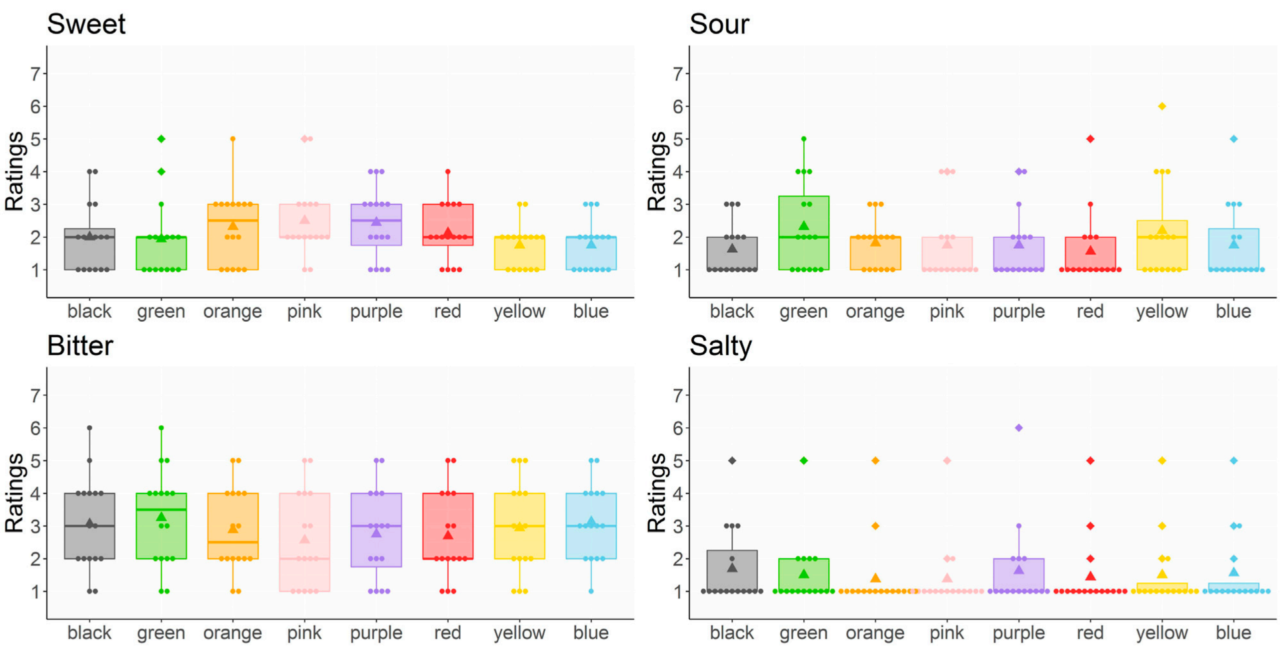


FIGURE 7
The descriptive statistics and the distribution of the results of taste rating (**Q2-1**). The dots, triangles, lines and diamonds represent participants' ratings, average ratings, median and outliers, respectively.

- **Q2-4.** How much do you like the just-finished scene when drinking the beverage? Ratings are from 1 (dislike it very much) to 7 (like it very much).

Before going to the next ambient color room, they were exposed to a white room and delivered a mini cup of pure water for palate cleansing. After the participants experienced all eight conditions, they were asked to take the VR HMD off, and answer a post-experiment questionnaire to collect their demographic information.

4.5 Procedure

The participants were first briefed about the task and procedure of the experiment. They were encouraged to raise any questions if something was unclear. They were also informed that they could stop at any point during the experiment if they felt uncomfortable. The participants then put the VR HMD on and were asked not to take it off before the end of the experiment. Thus, they could not see and did not receive any information about the drinks during the whole experiment. They were first exposed to a white room and were given a cup of pure water to accommodate the 'blinded' drinking action and cleansing palate. After that, they were given the mini cup containing the tastant and kept it in their hand. Next, the experimenter switched the scene to the first room to start the formal experiment. After the participants were immersed in the ambient color virtual environment for 30 s, the background voice began to remind the participant of starting drinking. After the drinking is finished, participants were asked for answering the above questions orally. Before going to the next ambient color room, they were exposed to a white room and the cups in their hands were replaced with new ones containing pure water. They then drank the pure water for palate cleansing. After that, they entered the next ambient color room with a new mini cup of tastant delivered, replacing the cup containing pure water. Noticing that, when immersed in the virtual environment without seeing the content in the physical world, most of the participants were unaware that they were drinking the same tastants throughout the study.

After the participants experienced all eight conditions, they were asked to take the VR HMD off, drink the taste stimuli once more, and then report their feelings. Finally, we gave participants a post-experiment questionnaire to collect their demographic information, past VR experience, and beverage consumption habits.

4.6 Results

SPSS (version 26) was used for data analysis. The results were transformed with Aligned Rank Transform (Wobbrock et al., 2011). We performed one-way Repeated Measure-(RM-) ANOVA with the transformed data for **Q1-1** and **Q1-4**. If the

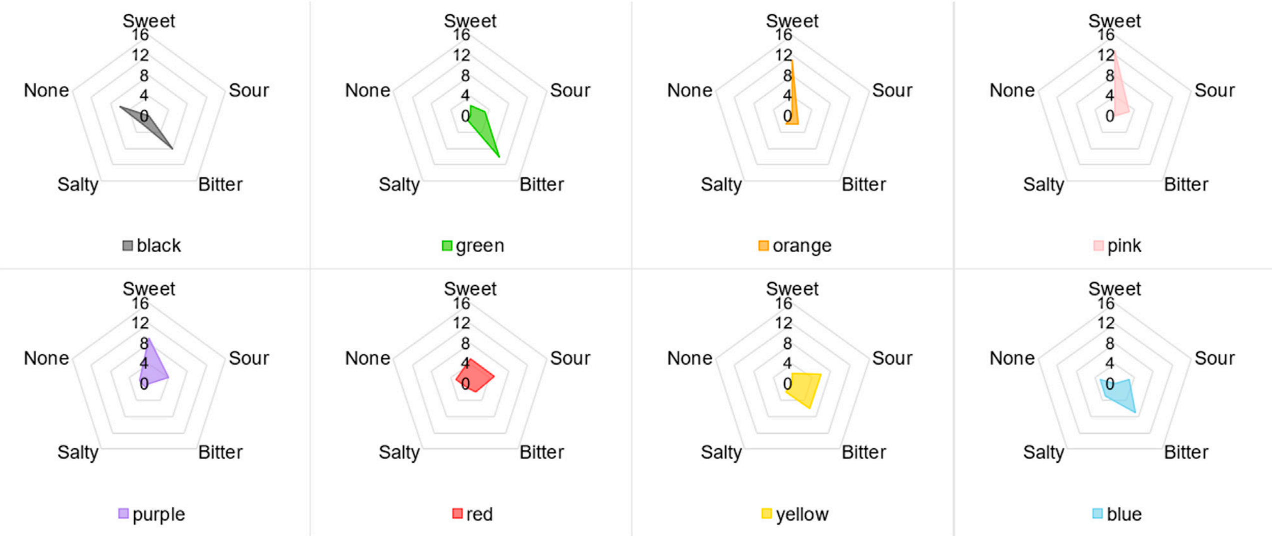


FIGURE 8
The frequencies and pairwise comparisons of the answers to **Q2-2** in Experiment 2. The colors in the figure correspond to the ambient colors in the virtual envrionmes.

TABLE 2 The frequencies and pairwise comparisons of the answers to Q2-2 in Experiment 2. The values in bold indicate the most frequently selected taste. Each subscript letter denotes a subset of color categories whose proportions do not differ significantly from each other at the 0.05 level reading across a row.

	Sweet	Sour	Bitter	Salty	None
black	0_c	$0_{a,c}$	$\mathbf{8_{a,b}}$	$2_{a,b,c}$	6_b
blue	0_b	$3_{a,b}$	$\mathbf{7_a}$	3_a	3_a
green	2_b	$3_{a,b}$	$\mathbf{10_a}$	$1_{a,b}$	$0_{a,b}$
orange	$\mathbf{11_a}$	1_a	2_a	2_a	$0_{a,b}$
pink	$\mathbf{13_b}$	$3_{a,b}$	0_a	$0_{a,b}$	$0_{a,b}$
purple	$\mathbf{9_b}$	$4_{a,b}$	0_a	$1_{a,b}$	$2_{a,b}$
red	$\mathbf{5_a}$	$\mathbf{5_a}$	2_a	1_a	3_a
yellow	2_a	$\mathbf{6_a}$	$\mathbf{6_a}$	2_a	0_a

transformed results violated the assumption of sphericity, we report the results with Greenhouse-Geisser or Huynh-Feldt corrections according to the case ($\epsilon < 0.75$ or > 0.75, respectively). We used Fisher's exact test to examine whether there is a significant color-taste association (i.e., **Q1-2**). Post-hoc pairwise comparisons were run with Bonferroni corrections if the above tests revealed significant differences or associations.

4.6.1 Q2-1. Perceived taste of the drink

Q2-1 refers to the influence on real taste. Figure 7 summarizes the descriptive statistics and the distribution of the ratings of **Q2-1**. As can be seen from the figure, the VR room with pink ambient color (AC) was on average rated the highest for sweet taste ($M = 2.500$, SD = 1.119) among eight *AC* conditions. For sour taste, the VR room with green *AC* reached the highest average rating ($M = 2.312$, SD = 1.310). Similarly, bitter taste in this *AC* environment was also rated the highest ($M = 3.250$, SD = 1.436). Participants did not report a strong feeling of salty taste. The ratings on saltiness were all low, but the VR room with black *AC* reached the highest average score of salty ($M = 1.688$, SD = 1.158).

RM-ANOVA tests revealed a significant main effect of *AC* on the ratings of sweet taste ($F(7, 105) = 2.296$, $p = 0.032, \eta_p^2 = 0.133$), but not on sour taste ($F(7, 105) = 1.547$, $p = 0.159, \eta_p^2 = 0.094$), bitter taste ($F(7, 105) = 1.106$, $p = 0.365, \eta_p^2 = 0.069$), or salty taste ($F(3.716, 55.736) = 0.957$, $p = 0.466, \eta_p^2 = 0.060$). We further performed a post-hoc analysis with the ratings of sweet taste. However, the pairwise comparison did not show significant differences in the ratings of sweet taste in the different *AC* conditions.

4.6.2 Q2-2. Color-taste associations

Q2-2 focuses on scenario-taste matching. Figure 8 shows the frequencies of taste matching for different ambient color environments. Results of Fisher's exact test show that there was a significant association between color and taste ($\chi^2 = 87.230$, $p < 0.001$). The frequencies and pairwise comparisons are summarized in Table 2. Black and green showed a significant association with bitter taste compared with sweet. Pink and purple presented a significant association with sweetness rather than bitterness. No significant difference was observed in the other colors. According to the frequencies, sour was

10.3389/frvir.2022.1056782

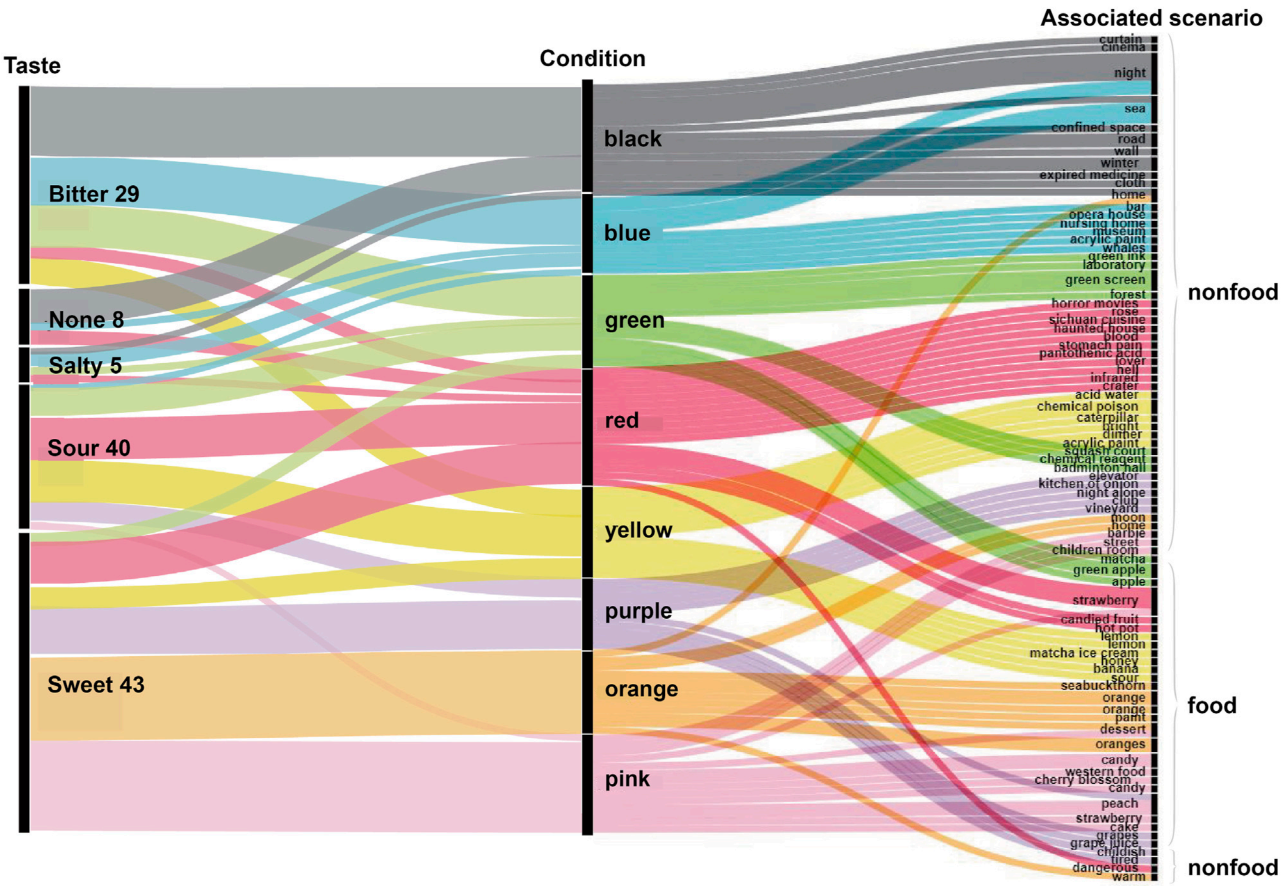


FIGURE 9
General associated objects and scenes under different taste choice (**Q2-3**). Participants gave the answers after matching the ambient color with the taste. Answers of associated scenario were categorized into food and non-food and listed in the right side. Lines were colored for indicating what ambient colors the answers were based on. Notice that, not all the conditions have associated scenarios.

associated more frequently with yellow, while salty was associated more frequently with blue.

4.6.3 Q2-3. Association of VR environment to real scenarios

Figure 9 shows the participants' reported real scenarios associated with the ambient light virtual environments, after matching them with the taste. Their answers were listed on the right side, with lines in colors indicating the association was under which ambient color virtual environment. Overall, there were three types of answers: scenes, food, and feelings. There were both answers related to foodstuff and non-foodstuff. Most participants associated sweet taste with food-related items. For example, after being exposed to the pink ambient light in the virtual environment, some participants came up with "candy" in their minds. Other common answers were lemon-yellow ambient light, apple-green ambient light, and oranges-orange ambient light. On the contrary, participants came up with specific scenes other than food (e.g., sea and night) after matching the ambient light with bitterness. Regarding the success of coming up with scenarios in their minds, associated scenarios were frequently described after being matched with sweetness (43, 21, 29, and 5 for the number of answers in sweet, sour, bitter, and salty taste correspondingly). Few participants described scenarios after matching the ambient color environment with the salty taste.

4.6.4 Q2-4. Degree of liking of ambient colors

Figure 10 provides an overview of the participants' answers to **Q1-4**. It can be seen from the figure that participants preferred orange and pink environments. Black, green, and red ambient environments were disliked by some participants. There was a significant main effect on the ratings in their preferred VR room (F (7, 105) = 2.463, p = 0.022). Post-hoc pairwise comparisons showed that ratings of orange (M = 4.313, SD = 1.158) was significantly higher than red (M = 3.000, SD = 1.173; p = 0.024, adjusted).

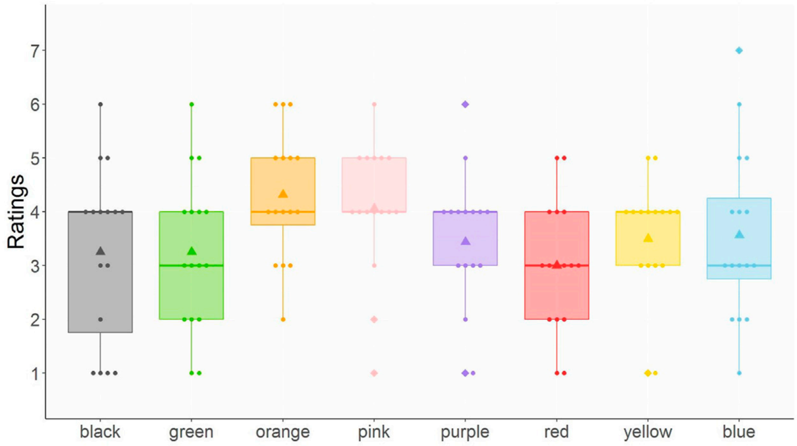


FIGURE 10
The descriptive statistics and the distribution of the results of color environment liking (**Q2-4**). The dots, triangles, lines and diamonds represent participants' ratings, average ratings, median and outliers, respectively. Asterisks mark the comparison that differed significantly ($p < 0.05$) in the post-hoc pairwise comparison.

5 Discussion

5.1 Comparison with related work

In this work, we focused on exploring how environmental colors would affect people's taste perception in VR than verifying a psychological phenomenon in VR. We used VR as a platform for providing new perceptions and a higher level of immersion to study the cross-modal correspondence (CMC) experience.

As listed in the Related Work (Section 2), our research can be positioned differently from prior work in physical and VR settings. Our results in taste association present an interesting comparison with the results given by studies conducted using a desktop computer screen as a medium for showing colors (Wan et al., 2014). Pink was significantly associated with sweetness in both VR and the real world. However, different from Wan et al. (2014)'s results from experiments in the real world where green was found to be more associated with sourness and black with a bitter taste, our results in VR showed that green can also be associated with a bitter taste. This indicates that the same color studied *via* different mediums may bring the same or different results in taste perception.

Unlike prior work regarding taste-related CMC in virtual environments, our selection of ambient colors as visual stimuli is novel and has not been explored systematically before. Choosing a parameter that is not related to the environment (e.g., color change of the foodstuff, change of the container shape) may lose the benefit of the immersion provided by VR headsets. In addition, our focus is on the basic aspect of VR than the complex ones. The color of ambient light suits these requirements. Our work is different from Chen et al. (2020) which used VR environments involving a diversity of visual elements, surface textures, and types of rooms. Moreover, while the experiment by Wang et al. (2020) and Huang et al. (2019) focused on the changing the color of the drink in VR (i.e., an intrinsic property of the drinks), our study dealt with the ambient color of the VR environment (i.e., an extrinsic property of the drinks). In our experiment, the sweetness of the beverage seems to be increased by the pink environment, and the orange VR environment was preferable. These findings can be compared with Wang et al. (2020)'s results showing that the intrinsic beverage color did not influence the perceived sweetness or liking of the beverage.

5.2 Conceptual model and emotional valence under the ambient colors

Participants' answer to the associated scene in the VR environment shows that most participants' color-taste associations were based on specific scenarios. This mainly includes foodstuff and non-food-related scenarios. According to the explanation by Spence (2019b) (see Section 2.2), people can often point to specific *source object(s)* that embody both color and taste, which is also in line with the associated scenarios regarding foodstuff in our experiment. The ambient color may trigger participants' recall of real-life experiences with food, and then the association with taste would come up based on these experiences. The variety of answers indicates the association may be personal and subjective, which may also be a topic for further investigation.

Some answers about scenarios other than foodstuff could be explained by *emotional mediation* (Spence, 2011) (see Section 2.2) of the relation between taste and the words regarding emotion in a particular culture. Categorized by taste association, some answers such as "chemical poison", "nursing house", "horror movie" appeared under bitter association while "vineyard", "club", "children's room" appeared under sweet association. When categorized by color, there were some negative words like "prison" for the black virtual environment, "horror" for the red virtual environment, and "depressed" for the green virtual environment. This kind of negative association observed with certain ambient colors may be in line with some recent studies that have explored negative emotions in VR (Lavoie et al., 2021; Magdin et al., 2021). These studies showed that virtual environments involving a higher level of absorption may increase negative emotional responses. Also, Elliot et al. (2007) suggested that attention must be paid to how color can act as a subtle environmental cue that has important influences on human behavior. Regarding color study in virtual environments, research work has shown that the context in a virtual environment where the ambient color was assigned can also influence emotion and valance (Lipson-Smith et al., 2020). This leads to a future question to be discussed since, in this experiment, we did not include 3D models in the scenes. Based on this, we suggest that future researchers or designers applying ambient color in VR systems can pay special attention to the emotional association the color can trigger.

5.3 Future applications

Our experiment results could contribute to the design of human-food applications, presenting the possibility of how the dieting experience can be altered by using specific environmental colors. Researchers have looked into the use of diet activities as positive psychological interventions to increase everyday happiness and well-being (Fischler, 2011; Cook et al., 2020). Moreover, existing works have explored what and how advanced technology can support healthy dieting using wearable devices and mobile apps (Dong and Biswas, 2013; Pan et al., 2019). Our work shows the possibility of using VR to change people's taste perception, especially the sensation of sweet taste. Given the health issues caused by high levels of sugary drink intake (Essman et al., 2021; Leung et al., 2021), our findings could have applications to support healthier drinking habits. We found that modifying the ambient color in a VR environment can enhance users' sensation of the sweet taste of the beverage. This approach provides a novel beverage-drinking experience and a cost-effective approach to changing users' diet behavior (by providing users' perception of sweetness with drinks that are lower in actual sugar content). Future research could look into this potential application *via* longitudinal experiments.

6 Limitations and future work

The first limitation of this research is that our sample population is limited to university students within a similar age group and is relatively small. To examine the topic further, a larger sample size including different age groups may be helpful taking our experiment framework as a reference. Also, in our experiments, participants did the taste ratings after receiving the stimuli at a consistent time interval. Further work is needed to check whether the same results will be obtained and whether other insights will be found when we extend or shorten the exposure time of the ambient color. Moreover, the aroma of the tea can be crucial for influencing the taste and could represent a bias for studying taste perception. However, our focus was the influence of ambient light on basic tastes, and we did not want to introduce the aroma of tea as another variable. We tried to reduce the influence of tea aroma on taste perception by covering it with lids for participants to hold in their hands to reduce the influence. Despite these limitations, we believe that our work contributes to the initial direction and provides a solid foundation on top of which further research could be conducted to better understand CMC, especially the ones evoked by ambient color in virtual environments.

Thinking more broadly, the virtual environment can be rendered in various ways and the VR HMD we used in this experiment is just one of them. In other types of virtual environments, including CAVE, the screen-based interface can also be taken into account to see if there is a difference in taste perception or not. Moreover, the only beverage we used was a neutral drink. In the future, it could be interesting to use several types of beverages and explore whether for example using beverages with a strong taste can have a greater or lesser effect.

7 Conclusion

Our research provides a solid example of using a Virtual Reality (VR) environment as an experimental platform to study the cross-modal correspondence between ambient color and taste perception. Results from the first experiment suggest the need for paying attention to the visually induced motion sickness and color recognition of people when doing color-related experiments in VR environments. Results from the second experiment demonstrated that the pink ambient color could change people's perception of the sweet taste of a neutral drink. The black and green ambient colors are associated with bitterness, while pink and purple are linked with sweetness. Moreover, the orange ambient color is preferable for drinking beverages in VR. Our research fills a gap in cross-modal correspondence targeting ambient color in VR, providing helpful and practical insights that can be applied to human-food interaction and VR applications that involve socialization and healthy habit promotion. Its results can also form the basis for further research.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by University Research Committee. The patients/participants provided their written informed consent to participate in this study.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication. ZW and MJ conducted this work when they were affiliated with Xi'an Jiaotong-Liverpool University.

Funding

This work is funded in part by Xi'an Jiaotong-Liverpool University (XJTLU) Key Special Fund (KSF-A-03) and XJTLU Research Development Fund (RDF-17-01-54), and the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJB520038).

Acknowledgments

The authors want to thank the participants for volunteering their time to join the experiments. We also thank the reviewers for the insightful comments and suggestions that helped improve our paper.

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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Arnold, P. (2017). You better eat to survive! exploring edible interactions in a virtual reality game. *Proc. 2017 CHI Conf. Ext. Abstr. Hum. Factors Comput. Syst.* '17, 206–209. doi:10.1145/3027063.3048408

Bangcuyo, R. G., Smith, K. J., Zumach, J. L., Pierce, A. M., Guttman, G. A., and Simons, C. T. (2015). The use of immersive technologies to improve consumer testing: The role of ecological validity, context and engagement in evaluating coffee. *Food Qual. Prefer.* 41, 84–95. doi:10.1016/j.foodqual.2014.11.017

Bonato, A., Bubka, F., and Alfieri, L. (2004). Display color affects motion sickness symptoms in an optokinetic drum. *Aviat. Space Environ. Med.* 75, 306–311.

Cardello, A. V. (1994). *Consumer expectations and their role in food acceptance.* Boston, MA: Springer US, 253–297. doi:10.1007/978-1-4615-2171-6_10

Carlson, P., Kirpes, C., Pavlik, R. A., Vance, J. M., Yin, L., Scott-Cooper, T., et al. (2011). "Comparison of single-wall versus multi-wall immersive environments to support a virtual shopping experience," in ASME 2011 World Conference on Innovative Virtual Reality, Milan, Italy, June 27–29, 2011 (World Conference on Innovative Virtual Reality), 287–291. doi:10.1115/WINVR2011-5582

Charles, S., Carmel, A., L., Maya, U. S., and Massimiliano, Z. (2010). Does food color influence taste and flavor perception in humans? *Chemosens. Percept.* 3, 68–84. doi:10.1007/s12078-010-9067-z

Chen, Y., Huang, A. X., Faber, I., Makransky, G., and Perez-Cueto, F. J. A. (2020). Assessing the influence of visual-taste congruency on perceived sweetness and product liking in immersive vr. *Foods* 9, 465. doi:10.3390/foods9040465

Clydesdale, F. M., Gover, R., and Fugardi, C. (1992). The effect of color on thirst quenching, sweetness, acceptability and flavor intensity in fruit punch flavored beverages. *J. Food Qual.* 15, 19–38. doi:10.1111/j.1745-4557.1992.tb00973.x

Cook, M., Kuntsche, S., Labhart, F., and Kuntsche, E. (2020). Do different drinks make you feel different emotions? Examination of young adolescents' beverage-specific alcohol expectancies using the alcohol expectancy task. *Addict. Behav.* 106, 106375. doi:10.1016/j.addbeh.2020.106375

Crisinel, A.-S., and Spence, C. (2010). As bitter as a trombone: Synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Atten. Percept. Psychophys.* 72, 1994–2002. doi:10.3758/APP.72.7.1994

Crisinel, A.-S., and Spence, C. (2009). Implicit association between basic tastes and pitch. *Neurosci. Lett.* 464, 39–42. doi:10.1016/j.neulet.2009.08.016

Crofton, E., Murray, N., and Botinestean, C. (2021). Exploring the effects of immersive virtual reality environments on sensory perception of beef steaks and chocolate. *Foods* 10, 1154. doi:10.3390/foods10061154

Delarue, J., Brasset, A.-C., Jarrot, F., and Abiven, F. (2019). Taking control of product testing context thanks to a multi-sensory immersive room. a case study on alcohol-free beer. *Food Qual. Prefer.* 75, 78–86. doi:10.1016/j.foodqual.2019.02.012

Dong, B., and Biswas, S. (2013). "Wearable diet monitoring through breathing signal analysis," in 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 03-07 July 2013, 1186–1189. doi:10.1109/EMBC.2013.6609718

Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., and Meinhardt, J. (2007). Color and psychological functioning: The effect of red on performance attainment. *J. Exp. Psychol. General* 136, 154–168. doi:10.1037/0096-3445.136.1.154

Essman, M., Stoltze, F. M., Carpentier, F. D., Swart, E. C., and Taillie, L. S. (2021). Examining the news media reaction to a national sugary beverage tax in South Africa: A quantitative content analysis. *BMC public health* 21, 454–514. doi:10.1186/s12889-021-10460-1

Fischler, C. (2011). Commensality, society and culture. *Soc. Sci. Inf.* 50, 528–548. doi:10.1177/0539018411413963

Gallace, A., Ngo, M. K., Sulaitis, J., and Spence, C. (2012). "Multisensory presence in virtual reality: Possibilities & limitations," in *Multiple sensorial media advances and applications: New developments in MulSeMedia* (Pennsylvania, United States: IGI Global), 1–38. doi:10.4018/978-1-60960-821-7.ch001

Gotow, N., Skrandies, W., Kobayashi, T., and Kobayakawa, T. (2021). Traditional Japanese confection overseas: Cultural difference and retronasal aroma affect flavor preference and umami perception. *Food Qual. Prefer.* 92, 104204. doi:10.1016/j.foodqual.2021.104204

Grassini, S., Laumann, K., de Martin Topranin, V., and Thorp, S. (2021). Evaluating the effect of multi-sensory stimulations on simulator sickness and sense of presence during hmd-mediated vr experience. *Ergonomics* 64, 1532–1542. doi:10.1080/00140139.2021.1941279

Gusev, D. A., Eschbach, R., Westin, T., and Yong, J. (2018). Motion-sickness-related aspects of inclusion of color-deficient observers in virtual reality. *Int. J. Child Health Hum. Dev.* 11, 177–181.

Gusev, D. A., Whittinghill, D. M., and Yong, J. (2016). "A simulator to study the effects of color and color blindness on motion sickness in virtual reality using head-mounted displays," in *Mobile and wireless technologies 2016*. Editors K. J. Kim, N. Wattanapongsakorn, and N. Joukov (Singapore: Springer Singapore), 197–204.

Halabi, O., and Saleh, M. (2021). Augmented reality flavor: Cross-modal mapping across gustation, olfaction, and vision. *Multimed. Tools Appl.* 1, 36423–36441. doi:10.1007/s11042-021-11321-0

Harrar, V., Piqueras-Fiszman, B., and Spence, C. (2011). There's more to taste in a coloured bowl. *Perception* 40, 880–882. doi:10.1068/p7040

Huang, F., Huang, J., and Wan, X. (2019). Influence of virtual color on taste: Multisensory integration between virtual and real worlds. *Comput. Hum. Behav.* 95, 168–174. doi:10.1016/j.chb.2019.01.027

Johnson, J., and Clydesdale, F. M. (1982). Perceived sweetness and redness in colored sucrose solutions. *J. Food Sci.* 47, 747–752. doi:10.1111/j.1365-2621.1982.tb12706.x

Keast, R. S., and Breslin, P. A. (2003). An overview of binary taste–taste interactions. *Food Qual. Prefer.* 14, 111–124. doi:10.1016/S0950-3293(02)00110-6

Kerruish, E. (2019). Arranging sensations: Smell and taste in augmented and virtual reality. *Senses Soc.* 14, 31–45. doi:10.1080/17458927.2018.1556952

Kim, H. K., Park, J., Choi, Y., and Choe, M. (2018). Virtual reality sickness questionnaire (vrsq): Motion sickness measurement index in a virtual reality environment. *Appl. Ergon.* 69, 66–73. doi:10.1016/j.apergo.2017.12.016

Knöferle, K., and Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychon. Bull. Rev.* 1, 992–1006. doi:10.3758/s13423-012-0321-z

Koch, C., and Koch, E. C. (2003). Preconceptions of taste based on color. *J. Psychol.* 137, 233–242. doi:10.1080/00223980309600611

Kuliga, S., Thrash, T., Dalton, R., and Hölscher, C. (2015). Virtual reality as an empirical research tool — Exploring user experience in a real building and a corresponding virtual model. *Comput. Environ. Urban Syst.* 54, 363–375. doi:10.1016/j.compenvurbsys.2015.09.006

Lavoie, R., Main, K., King, C., and King, D. (2021). Virtual experience, real consequences: The potential negative emotional consequences of virtual reality gameplay. *Virtual Real.* 25, 69–81. doi:10.1007/s10055-020-00440-y

Ledoux, T., Nguyen, A. S., Bakos-Block, C., and Bordnick, P. (2013). Using virtual reality to study food cravings. *Appetite* 71, 396–402. doi:10.1016/j.appet.2013.09.006

Leung, C. Y., Abe, S. K., Sawada, N., Ishihara, J., Takachi, R., Yamaji, T., et al. (2021). Sugary drink consumption and risk of kidney and bladder cancer in Japanese adults. *Sci. Rep.* 11, 21701–21709. doi:10.1038/s41598-021-01103-x

Lipson-Smith, R., Bernhardt, J., Zamuner, E., Churilov, L., Busietta, N., and Moratti, D. (2020). Exploring colour in context using virtual reality: Does a room change how you feel? *Virtual Real.* 25, 631–645. doi:10.1007/s10055-020-00479-x

Maga, J. A. (1974). Influence of color on taste threshold. *Chem. Senses* 1, 115–119. doi:10.1093/chemse/1.1.115

Magdin, M., Balogh, Z., Reichel, J., Francisti, J., Koprda, Š., and György, M. (2021). Automatic detection and classification of emotional states in virtual reality and standard environments (lcd): Comparing valence and arousal of induced emotions. *Virtual Real.* 1, 1029–1041. doi:10.1007/s10055-021-00506-5

Moser, C., and Tscheligi, M. (2013). "Playful taste interaction," in IDC '13: Proceedings of the 12th International Conference on Interaction Design and Children, New York, NY, 24 June 2013 (New York: Association for Computing Machinery), 340–343. doi:10.1145/2485760.2485828

Narumi, T., Nishizaka, S., Kajinami, T., Tanikawa, T., and Hirose, M. (2011). "Augmented reality flavors: Gustatory display based on edible marker and cross-modal interaction," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 07 May 2011 (Association for Computing Machinery), 93–102. CHI '11. doi:10.1145/1978942.1978957

Nishizawa, M., Jiang, W., and Okajima, K. (2016). "Projective-ar system for customizing the appearance and taste of food," in Proceedings of the 2016 Workshop on Multimodal Virtual and Augmented Reality, New York, NY, USA, 16 November 2016 (Association for Computing Machinery). MVAR '16. doi:10.1145/3001959.3001966

Oberfeld, D., Hecht, H., Allendorf, U., and Wickelmaier, F. (2009). Ambient lighting modifies the flavor of wine. *J. Sens. Stud.* 24, 797–832. doi:10.1111/j.1745-459X.2009.00239.x

Oberfeld, D., and Hecht, H. (2011). Fashion versus perception: The impact of surface lightness on the perceived dimensions of interior space. *Hum. Factors* 53, 284–298. doi:10.1177/0018720811407331

Pan, S., Ren, X., Brombacher, A., and Vos, S. (2019). "Designing technology to encourage healthy eating at work," in Proceedings of the 9th International Conference on Digital Public Health, New York, NY, USA, 20 November 2019 (Association for Computing Machinery), 131. DPH2019. doi:10.1145/3357729.3357759

Parise, C., and Spence, C. (2013). "Audiovisual cross-modal correspondences in the general population," in *The Oxford handbook of synesthesia* (Oxford, United Kingdom: Oxford University Press).

Piqueras-Fiszman, B., and Spence, C. (2014). Colour, pleasantness, and consumption behaviour within a meal. *Appetite* 75, 165–172. doi:10.1016/j.appet.2014.01.004

Piqueras-Fiszman, B., and Spence, C. (2012). The influence of the color of the cup on consumers' perception of a hot beverage. *J. Sens. Stud.* 27, 324–331. doi:10.1111/j.1745-459X.2012.00397.x

Slater, M., and Wilbur, S. (1997). A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence. (Camb).* 6, 603–616. doi:10.1162/pres.1997.6.6.603

So, R. H., and Yuen, S. (2007). "Comparing symptoms of visually induced motion sickness among viewers of four similar virtual environments with different color," in *Virtual reality* (Berlin, Germany: Springer Berlin Heidelberg), 386–391. doi:10.1007/978-3-540-73335-5_42

Spence, C., and Carvalho, F. M. (2020). The coffee drinking experience: Product extrinsic (atmospheric) influences on taste and choice. *Food Qual. Prefer.* 80, 103802. doi:10.1016/j.foodqual.2019.103802

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Atten. Percept. Psychophys.* 73, 971–995. doi:10.3758/s13414-010-0073-7

Spence, C., and Levitan, C. A. (2021). Explaining crossmodal correspondences between colours and tastes. *i-Perception* 12, 204166952110182. doi:10.1177/20416695211018223

Spence, C. (2019a). Multisensory experiential wine marketing. *Food Qual. Prefer.* 71, 106–116. doi:10.1016/j.foodqual.2018.06.010

Spence, C. (2019b). On the relationship(s) between color and taste/flavor. *Exp. Psychol.* 66, 99–111. doi:10.1027/1618-3169/a000439

Spence, C. (2016). Oral referral: On the mislocalization of odours to the mouth. *Food Qual. Prefer.* 50, 117–128. doi:10.1016/j.foodqual.2016.02.006

Spence, C., and Piqueras-Fiszman, B. (2016). "Food color and its impact on taste/flavor perception," in *Multisensory flavor perception* (Amsterdam, Netherlands: Elsevier), 107–132.

Spence, C., Velasco, C., and Knoeferle, K. (2014). A large sample study on the influence of the multisensory environment on the wine drinking experience. *Flavour* 3, 8–12. doi:10.1186/2044-7248-3-8

Spence, C., and Velasco, C. (2018). On the multiple effects of packaging colour on consumer behaviour and product experience in the 'food and beverage' and 'home and personal care' categories. *Food Qual. Prefer.* 68, 226–237. doi:10.1016/j.foodqual.2018.03.008

Stein, B. E., and Meredith, M. A. (1993). *The merging of the senses*. Massachusetts, United States: MIT Press.

Stelick, A., Penano, A. G., Riak, A. C., and Dando, R. (2018). Dynamic context sensory testing–a proof of concept study bringing virtual reality to the sensory booth. *J. Food Sci.* 83, 2047–2051. doi:10.1111/1750-3841.14275

Sugimori, E., and Kawasaki, Y. (2022). Cross-modal correspondence between visual information and taste perception of bitter foods and drinks. *Food Qual. Prefer.* 98, 104539. doi:10.1016/j.foodqual.2022.104539

Unity (2019). Global illumination. Available at: https://docs.unity.cn/2019.1/Documentation/Manual/GlobalIllumination.html (Accessed November 24, 2022).

von Castell, C., Hecht, H., and Oberfeld, D. (2018). Which attribute of ceiling color influences perceived room height? *Hum. Factors* 60, 1228–1240. doi:10.1177/0018720818789524

Wan, X., Woods, A. T., van den Bosch, J. J., McKenzie, K. J., Velasco, C., and Spence, C. (2014). Cross-cultural differences in crossmodal correspondences between basic tastes and visual features. *Front. Psychol.* 5, 1365. doi:10.3389/fpsyg.2014.01365

Wang, Q. J., Barbosa Escobar, F., Alves Da Mota, P., and Velasco, C. (2021). Getting started with virtual reality for sensory and consumer science: Current practices and future perspectives. *Food Res. Int.* 145, 110410. doi:10.1016/j.foodres.2021.110410

Wang, Q. J., Meyer, R., Waters, S., and Zendle, D. (2020). A dash of virtual milk: Altering product color in virtual reality influences flavor perception of cold-brew coffee. *Front. Psychol.* 11, 595788. doi:10.3389/fpsyg.2020.595788

Wang, Q. J., Mielby, L. A., Junge, J. Y., Bertelsen, A. S., Kidmose, U., Spence, C., et al. (2019). The role of intrinsic and extrinsic sensory factors in sweetness perception of food and beverages: A review. *Foods* 8, 211. doi:10.3390/foods8060211

Wobbrock, J. O., Findlater, L., Gergle, D., and Higgins, J. J. (2011). "The aligned rank transform for nonparametric factorial analyses using only anova procedures," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA, 07 May 2011 (Association for Computing Machinery), 143–146. CHI '11. doi:10.1145/1978942.1978963

frontiers | Frontiers in Virtual Reality

TYPE Original Research
PUBLISHED 02 October 2023
DOI 10.3389/frvir.2023.1261093

Check for updates

OPEN ACCESS

EDITED BY
Hai-Ning Liang,
Xi'an Jiaotong-Liverpool University, China

REVIEWED BY
Diego Vilela Monteiro,
ESIEA University, France
Stefan Marks,
Auckland University of Technology, New Zealand

*CORRESPONDENCE
Belmir Jose De Jesus Junior,
belmir.jesus@inrs.ca
Tiago H. Falk,
tiago.falk@inrs.ca

RECEIVED 18 July 2023
ACCEPTED 15 September 2023
PUBLISHED 02 October 2023

CITATION
De Jesus Junior BJ, Perreault L, Lopes MKS, Roberge M-C, Oliveira AA and Falk TH (2023), Using multisensory virtual reality nature immersion as a therapeutic modality for improving HRV and cognitive functions in post-traumatic stress disorder: a pilot-study. *Front. Virtual Real.* 4:1261093. doi: 10.3389/frvir.2023.1261093

Using multisensory virtual reality nature immersion as a therapeutic modality for improving HRV and cognitive functions in post-traumatic stress disorder: a pilot-study

Belmir Jose De Jesus Junior[1]*, Léa Perreault[1,2], Marilia K. S. Lopes[1], Marie-Claude Roberge[2], Alcyr A. Oliveira[3] and Tiago H. Falk[1]*

[1]Institut National de la Recherche Scientifique, University of Quebec, Montreal, QC, Canada, [2]Traumas Côte-Nord, Sept-Iles, QC, Canada, [3]Federal University of Health Sciences of Porto Alegre, Porto Alegre, RS, Brazil

Introduction: Immersive virtual reality (VR) applications are burgeoning within healthcare as they promote high levels of engagement. Notwithstanding, existing solutions only stimulate two of our five senses (audio and visual), thus may not be optimal in the sense of promoting immersion and of "being present". In this paper, we explore the benefits of an immersive multisensory experience as a therapeutic modality for participants suffering from post-traumatic stress disorder (PTSD).

Methods: In addition to 360-degree videos and corresponding natural sounds, nature smells are also presented by means of a portable ION 2 scent diffusion device attached to an Oculus Quest 2 VR head-mounted display. A 3-week 12-sessions protocol was applied to a sample of 20 participants diagnosed with PTSD.

Results and discussion: We report the outcomes seen from a battery of qualitative metrics, including cognitive functioning tests, psychological symptoms, severity of PTSD, and several self-reported questionnaires and heart rate variability (HRV) metrics. Results are compared not only between pre-and post intervention, but also after a 3-month follow-up period. Results suggest a decrease in the severity of PTSD, as well as improvements in processing speed and sustained attention post-intervention, but also sustained decrease in the severity of PTSD and in dissociative tendencies at the 3-month follow-up. Overall, participants rated the experience as highly immersive and produced very mild to no symptoms of cybersickness, thus corroborating the feasibility and usefulness of the proposed multisensory immersive VR tool for reducing PTSD symptoms.

KEYWORDS

virtual reality, relaxation training, olfactory, multisensory, mental health

1 Introduction

Post-traumatic stress disorder (PTSD) is a mental health condition that results from experiencing or witnessing a traumatic life event. People suffering from PTSD may experience psychological symptoms such as intrusive thoughts, nightmares, flashbacks, hypervigilance, and severe anxiety. Cognitive functions might also be impaired, leading to attention and memory deficits. Furthermore, PTSD can lead to diminished physical health, debilitated social and daily functioning, decline in work performance and increase in demand for healthcare and mental health resources. Therefore, PTSD impacts not only the individual, but also their family, their community, and society in general (Shalev et al., 2017).

Even though extensive research has been conducted to reduce PTSD symptoms, its complexity and individual variations make the elaboration of treatments challenging. Exposure therapy, for example, has been regarded as a useful method to treat PTSD (Rauch et al., 2012). However, it still entails high dropout rates and insufficient accessibility to many patients. Some clinicians opt for other treatment methods, as exposing the patient to their trauma can be a distressing experience (Bomyea and Lang, 2012). Less invasive therapies also exist, such as relaxation training or eye movement desensitization and reprocessing (EMDR). However, the treatments have been found to be less efficacious than exposure therapy (Taylor et al., 2003). Therefore, the need for new therapeutic modalities that increase patient's motivation and are clinically effective in diminishing PTSD symptoms is more present than ever. In response, modern technology has earned scientific interest in the past few years for its unique advantages and promising results.

Technological advances in computer graphics hardware, communications, and immersive media software have brought an ascension in computer-simulated environments, commonly referred to as virtual reality (VR). Today, applications of VR go beyond entertainment, and many treatment pathways are emerging in healthcare (Halbig et al., 2022). Within mental health research, VR has been used to treat phobias, addictions, attention deficit and hyperactivity disorder, autism spectrum disorder, and post-traumatic stress disorder (Freeman et al., 2017; Kothgassner et al., 2019; Emmelkamp and Meyerbröker, 2021). VR allows different stimuli to be invoked in immersive virtual worlds, which would often be hard or impossible to accomplish in a usual real-life treatment scenario (Rizzo and Shilling, 2017). A recent systematic review has shown that treatment delivered in an intense format, for example, with multiple or continuous sessions over 1 week, is effective and reduces the amount of time required to obtain its benefits, which in turn diminishes dropout rates (Ragsdale et al., 2020). The use of VR in the clinic and at home could allow for such intense sessions to be conducted and could be a good adjuvant to regular therapy, thus increasing the frequency of services for patients in need.

The use of VR immersive content to promote relaxation has also been investigated recently, not only among people suffering from mental health problems, but also among healthy individuals. For relaxation, most of the applications consist of exposing the subject to a relaxing environment, usually featuring nature content, or even using the content to induce a meditative state (Noronha and Campos, 2021; Lopes et al., 2022; Riches et al., 2023). In addition, nature immersion delivered by head-mounted displays (HMDs) has been self-reported as more relaxing than when they are delivered on a PC screen (Knaust et al., 2022). Recently, Mistry et al. (2020) reported that promoting relaxation through VR for PTSD patients increased positive affect. Although more positive affect was obtained from meditating with VR than in the non-VR condition, the authors emphasize the need for further exploration of VR to promote greater relaxation states. VR applications are known to induce a strong sense of presence in the user while immersed in the virtual environment, making them a useful tool for medical interventions, as they are more engaging and can increase participant's motivation. One main limitation of existing VR systems is that they only stimulate two of our five senses and, thus, may be considered sub-optimal for immersion. Introducing smells into the VR content has been revealed to improve the sense of realism, presence, immersion and emotion in users (De Jesus Jr et al., 2022). Additionally, Lopes et al. (2022) showed the potential of multisensorial VR applications to promote greater relaxation states in healthy individuals. Furthermore, Aiken and Berry (2015) stated the possibility of translating the benefits of olfaction therapy to patients suffering from PTSD symptoms with the utilization of VR exposure therapy.

In this paper, we explore the use of multisensory nature immersion, where VR audio-visual nature scenes are combined with nature smells, as a new therapeutic modality for PTSD patients. In particular, we explore the benefits of multisensory VR on psychological, cognitive, and physiological symptoms of PTSD.

TABLE 1 Demographics of study participants. CEGEPs are exclusive to the Québec provincial education system and correspond to a vocational college.

Variable	Categories	*n*	%
Sex	Female	7	35
	Male	13	65
Marital Status	Single	4	20
	Common-law	5	25
	Married	7	35
	Divorced	2	10
	Other	2	10
Parental Status	No children	7	35
	At least 1 child	13	65
Educational Level	Elementary	1	5
	High school	4	20
	Professional Diploma	3	15
	CEGEP	3	15
	Bachelor's degree	7	35
	Masters/Ph.D. degree	1	5
	Other	1	5
Professional Status	Full/Part time job	6	30
	Invalid	12	60
	Retired	2	10

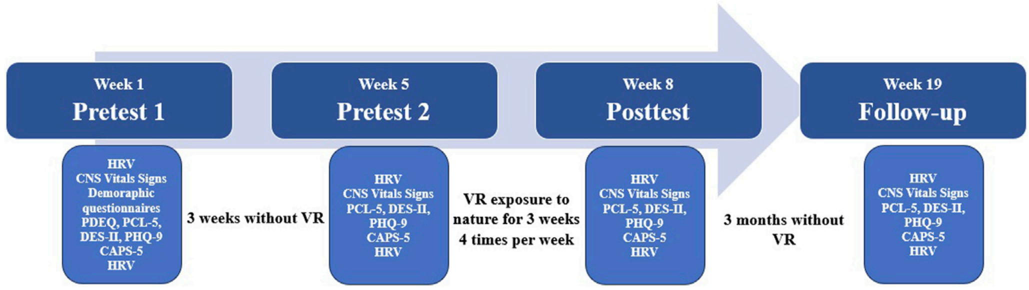


FIGURE 1
Protocol of VR nature immersion for each participant.

2 Materials and methods

2.1 Participants

Data were collected from July 2022 to May 2023 at a clinic specializing in PTSD assessment and treatment located in Sept-Îles, QC, Canada. The Ethics Committee of Research with Human Beings (CER-22-669) from INRS (University of Québec) approved the study. The trial was registered under the ISRCTN registry identifier: ISRCTN68760993. Twenty-four subjects with a diagnosis of PTSD recognized by their family doctor were initially recruited. Twenty completed the protocol, as four participants withdrew from the study due to personal reasons. The remaining participants had an average age of 42.2 ± 11.35 years old. Most of the participants were male (65%), married (35%), with at least one child (65%), and holding a bachelor's degree or higher education (40%). Sixty percent of the participants were not working due to disability. Table 1 displays a summary of the demographic characteristics of the participants. All participants were native French speakers or had an acceptable understanding of French. The nature of traumatic events in the sample included transportation accident (3), work accident (2), physical (5) or sexual aggression (5), war (2), witness of the death of another (1) and having or witnessing an injury or a sickness (2). Fifteen percent of participants declared only one traumatic event, whereas 25% identified two, 20% identified three, 5% identified four and 35% identified five or more. Eighty percent of participants followed psychotherapy during the intervention, and all participants had access to psychological support on their demand if they experienced mental distress due to the modality.

2.2 Procedure

Participants had to be over 18 years old, understand French, and have a PTSD diagnosis recognized by their doctors. They could not have a severe addiction, uncontrolled epilepsy or trauma linked with nature to participate. They provided their written and informed consent to take part in the experiment. Figure 1 depicts the protocol followed. At week 1 of the protocol, they were exposed to a short VR environment with olfactory stimuli to assess the presence of symptoms of so-called cybersickness (i.e., motion sickness caused by immersion in VR content) or any other impediment that could cause the exclusion of the participant from the experiment. Participants who could continue the protocol then proceeded to the first pretest phase, which was composed of a 30–45 min neurocognitive assessment (CNS Vital Signs, United States), a demographic questionnaire, four questionnaires measuring psychological symptoms and subjective distress, and a semi-structured interview administered by a trained psychologist to determine the symptoms and severity of PTSD (see Section 2.5). At the beginning and at the end of pretest 1, participants' heart rate variability was assessed to measure the activation level of the autonomic nervous system with a one-minute heart rate variability test using the EmWave Pro Plus software (HeartMath, United States) (see Section 2.6).

Next, a three-week period free of VR immersion was prescribed for every participant (those who followed psychotherapy continued as usual) to monitor changes in variables unrelated to the experiment. Then, 1 day before starting immersion in the multisensory VR environments, each subject completed a second pretest that included all the measures from the first pretest, except for the demographic questionnaire and the PDEQ questionnaire. The protocol itself was comprised of 12 sessions, distributed across 3 weeks, four times per week. Each session lasted 15 min and was comprised of three different virtual reality environments, each with a duration of 5 minutes (more details below). The order in which the participants viewed the environments was randomized per participant and per session to avoid possible biases. Once the participants finished all 12 sessions, they completed a posttest identical to the second pretest, as well as a three-month follow-up to evaluate the long-term effects of the protocol. At the end of each VR exposition session, participants also reported their sense of presence, immersion, and cybersickness using standard questionnaires (see Section 2.7).

2.3 Hardware

A VR headset (Oculus Quest 2, Meta, United States) with 1832 × 1920 resolution, up to 120 Hz display refresh rate, and 90° field-of-view was used in our studies. The HMD was instrumented with electroencephalography (EEG), electrocardiography (ECG), facial electromyography (EMG), and electro-oculography (EOG) sensors following advice from Cassani et al. (2018, 2020). More specifically, the VR headset was instrumented with 16 ExG sensors connected to a wireless bioamplifier (OpenBCI Cyton/Daisy) operating at a sample rate of 125 Hz. EEG data were obtained from 12 channels where sensors were embedded in the foam and

FIGURE 2
Captures from environment one (first row), two (second row), and three (third row). Environment one: Cabin Near Misty Lake, by Eric Fassbender. Licensed by Atmosphaeres. Environment two and three: Reproduced with permission from Martin Demassieux.

straps of the headset (i.e., Fp1, Fpz, Fp2, F3, F4, FCz, C3, C4, O1, O2, P3, and P4), four EOG electrodes were strategically placed on the foam of the headset (two pairs of horizontal and vertical electrodes to track eye activity), heart rate was obtained through PPG via a sensor that was placed on the left upper side of the visual foam, and two electrodes were placed on mastoids as a reference. In this work, we will focus only on experiment outcomes and leave the analyses of these biosignals for future exploration. Lastly, the olfactory stimulation was provided by an OVR ION2 scent diffuser device (OVR Technologies, United States) attached to the instrumented headset. The ION 2 device was calibrated to disperse up to nine different nature scents, out of which five were mainly used (flowers, earth dirt, forest, ocean breeze, and grass).

2.4 Virtual environments

As mentioned above, each of the 12 sessions was comprised of three nature scenes chosen to promote relaxation. The first scene brought the participant to the shore of an in-mountain lake (Figure 2—first row). This scenario acted as a baseline for the nature immersion, as no other stimulus (guided meditation, relaxing music or breathing exercise) was added to the existing natural sounds and odours associated with the scenario. The second scenario depicted a local beach in Sept-Îles (QC, Canada), in which the participant was placed near the water. In this scenario, a cardiac coherence exercise was added to observe its effects on the level of relaxation in the natural environment. At every moment, the participant could synchronize their breathing with a sphere moving up and down at 5 seconds intervals (Figure 2—middle row). The third scenario also depicted a local rocky seaside location in Sept-Îles, surrounded by forest. The participant was guided by an audio-guided meditation female voice instructing them to focus on their breathing and relaxation (Figure 2—third row).

2.5 Psychological and cognitive measurements

2.5.1 PDEQ

The French version of the Peritraumatic Dissociative Experiences Questionnaire (PDEQ) (Marmar et al., 2004) was used to assess if a dissociative episode occurred during and/or after the traumatic event that led to PTSD. Participants rate 10-items on a scale of 1 (1 = not at all true) to 5 (5 = extremely true) the presence of symptoms such as depersonalization, derealization, and amnesia that might have occurred during and/or after the event. A total score above 15 indicates significant dissociation. This instrument has shown satisfactory convergent validity [Pearson's correlation (*r*) of 0.39–0.54] with other measures of PTSD, including PTSD diagnosis supported by CAPS-5 interview (Bomyea and Lang, 2012).

2.5.2 PCL-5

The French version of the Post-traumatic Stress Disorder Checklist (PCL-5) (Weathers et al., 1993); validated in French by

Ashbaugh et al. (2016) was used to assess PTSD symptoms severity as perceived by the participants. The instrument consists of a 20-items questionnaire associated with each PTSD symptom as they appear in their respective clusters (intrusions, avoidance, changes in cognition and mood, and changes in reactivity). Patients rate each item on a scale from 0 (Not at all) to 4 (Extremely). The subjective endorsement of the necessary DSM-5 criterion for PTSD indicates a provisional diagnosis of PTSD, which usually sums up to a total score of 33 or higher. PCL-5 has shown average to strong internal consistency (Cohen's alpha (α) of 0.94) and test-retest reliability (α = 0.89), as well as moderate to strong validity of construct and criterion, as shown by its correlation with the CAPS-5 (r = 0.90) and other measures of PTSD. It was also used in other studies to measure the subjective perception of changes in PTSD symptoms pre- and post-experiment (Forkus et al., 2022; Ashbaugh et al., 2016).

2.5.3 CAPS-5

The French version of the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5) (Weathers et al., 2018); validated in French by Rivest-Beauregard et al. (2022) was used to objectively assess PTSD symptoms severity based on the DSM-5 criteria. It is a semi-structured interview in which the clinician evaluates the presence, frequency, and intensity of each symptom of PTSD on a scale from 0 (0 = absent) to 4 (4 = extreme/incapacitating). A rating of 2 (moderate) represents the threshold to qualify a symptom as clinically significant. Following this scale, a mean score of the first 20 questions (criteria B, C, D and E) indicates the presence and severity of PTSD (0 to 1 = Absent to mild PTSD; 1 to 2 = Moderate PTSD; 2 to 3 = Severe PTSD; 3 to 4 = Extreme PTSD). The CAPS-5 has been validated and used in many studies involving PTSD patients with great internal consistency (α = 0.90), test-retest reliability (Cohen's Kappa = 100) over a 1-month period and convergent validity with the PCL-5 (r = 0.30) (Rivest-Beauregard et al., 2022).

2.5.4 DES-II

The French version of the Dissociative Experiences Scale (DES-II) (translated by Saintonge, 1999) was used to assess dissociation, amnesia and absorption symptoms. This self-assessment is composed of 28-items depicting day-to-day situations of dissociative symptoms. Patients rate on a scale of 0–100 the percentage of time that they spend in each dissociative state described. The total score is obtained by calculating the mean of all 28 items, and a score above 30 indicates significant dissociation in day-to-day life. DES-II was found to have high test-retest reliability ($0.79 < r < 0.84$) and internal consistency (α = 0.95) (Carlson and Putnam, 1993).

2.5.5 PHQ-9

The French version of the Patient Health Questionnaire (PHQ-9) (Kroenke et al., 2001) was used to assess symptoms of depression. The measure consists of a self-assessment based on a 9-items questionnaire associated with cognitive, affective, and somatic symptoms of depression. Patients must rate the frequency to which they presented each symptom in the last 2 weeks on a scale from 0 (Never) to 3 (Almost every day). The cut-offs are of 4 and under for no depression, 5 to 9 for minor depression, 10 to 14 for moderate depression, 15 to 19 for moderate to severe depression and 20 and above for severe depression. PHQ-9 presents great construct validity for depression and anxiety, has been found sensible to changes pre- and post-treatment and has a good internal consistency (α = 0.87) (Beard et al., 2016).

2.5.6 CNS Vital Signs cognitive test

The CNS Vitals Signs was used to assess cognitive functions. It is a computerized battery containing seven tests (verbal and visual memory, finger tapping, symbol digit coding, Stroop Test, shifting attention and continuous performance). The test picks from a bank of words, numbers, and drawings for each testing session to avoid memorization, and is sensible to invalid responses. Standard scores of verbal and visual memory, complex attention, processing speed, working memory, and sustained attention were analyzed according to the test's normative categories (Above average = > 109; Average = 90 to 109; Low average = 80 to 89; Low = 70 to 79; Very low = < 70). CNS Vitals Signs has shown good test-retest reliability (r = 0.65–0.88) (Gualtieri and Johnson, 2006) and has been shown to be unbiased for repeated measures following the second testing session.

2.6 Physiological indicators

Literature has shown that PTSD patients have increased heart rate (Pole, 2007) and decreased heart rate variability (Schneider and Schwerdtfeger, 2020) when compared to controls. Therefore, the EmWave Ear sensor from HeartMath was used to assess heart rate variability before and after each pretest, posttest, and follow-up for a duration of 1 minute. Lo et al. (2017) showed that the Emwave Pro sensor, attached to the earlobe, is not obtrusive to participants and provides data with comparable validity to other heart rate variability devices. Metrics such as heart rate (HR), inter-beat interval (IBI), maximal heart rate reserve (MHRR), standard deviation of NN intervals (SDNN), and root mean square of successive differences (RMSSD) were extracted from the segments.

2.7 VR experience measurements

After each VR session, the french-Canadian versions of the Immersive Tendencies Questionnaire (QPI), the Presence Questionnaire (QÉP) (Witmer and Singer, 1998) and Simulator Sickness Questionnaire (QC) (Kennedy et al., 1993) were used. All questionnaires were validated in French by the Laboratory of Cyberpsychology of the University of Quebec in Outaouais (UQO) with Cronbach alpha's of 0.86 for the QC (Bouchard et al., 2011), 0.78 for the QPI, and 0.84 for the QÉP (Robillard et al., 2002). The QC contains 16 items covering different symptoms of cybersickness rated on a scale from 0 (Not at all) to 3 (Severely). QPI contains 18 items rated on a scale from 1 (Never or Not at all) to 7 (Often or A lot). QÉP contains 24 items in total, but only the basic scale of 19 items was used and rated on a scale from 1 (Not at all, Very Artificial, Not reactive) to 7 (Completely, Very reactive, Completely natural). The total score of all three questionnaires is equal to the sum of the rating of each item, with the inversion of items 14, 17 and 18 in the QÉP.

TABLE 2 Psychological tests correlations.

	Pretest 1				Pretest 2		
	PDEQ	CAPS-5	PCL-5	PHQ-9	CAPS-5	PCL-5	PHQ-9
CAPS-5	0.488*						
PCL-5	0.502*	0.921***			0.844***		
PHQ-9	0.291	0.791***	0.727***		0.617**	0.691***	
DES-II	0.421	0.278	0.247	0.234	0.265	0.418	0.317
	Posttest				Follow-up		
	PDEQ	CAPS-5	PCL-5	PHQ-9	CAPS-5	PCL-5	PHQ-9
PCL-5		0.881***			0.882***		
PHQ-9		0.724***	0.741***		0.67**	0.765***	
DES-II		0.397	0.237	0.319	0.423	0.46*	0.554*

* p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

Therefore, a higher score represents the higher intensity of the concept. QPI was only distributed at the first session, as it is considered a trait rather than a state.

2.8 Data analysis

The protocol used is a single-centre longitudinal pretest-posttest study that aims to examine whether virtual nature immersion has an impact on HRV, cognitive functions and affective symptoms in participants with PTSD. Therefore, a non-parametric option to the one-way repeated measures analyses of variance (ANOVA) was conducted (Cleophas et al., 2016). Friedman's test was chosen for the analysis since the Shapiro-Wilk test's null hypothesis for normality was rejected for some groups of measurements. Non-parametric *post hoc* analyses were performed by the Nemenyi *post hoc* test for the significant results of Friedman's test to verify which pairwise groups had a significant difference. If participants had a missing value for any of the tests, they were excluded from that specific analysis.

3 Results and discussions

3.1 Psychological and cognitive measurements

Table 2 outlines the Pearson correlation coefficients measured among the psychological measurements taken. As expected, scores of PDEQ, CAPS-5, PHQ-9, and PCL-5 were positively and significantly correlated with each other at pretest 1. From pretest 1 to pretest 2, no significant change was found in psychological symptoms. In general, PCL-5 was significantly and strongly correlated with CAPS-5 in all four measurements, since both tests are used to measure PTSD symptom severity. Additionally, there was a significant correlation between PHQ-9 scores and PTSD severity measurements, implying that the severity of PTSD and depression symptoms are correlated. Therefore, participants with higher scores of depression and dissociative tendencies presented higher severity of PSTD on the CAPS-5 interview. DES-II scores were not significantly correlated with other psychological variables.

Figure 3 depicts the distribution of the psychological test variables collected throughout the experiment. Overall, a declining trend in the median of all psychological variables from pretests to posttest measurements can be observed. Furthermore, this decrease is still observable for CAPS-5 and DES-II during the three-month follow-up measurements, suggesting the experiment's lasting efficacy at least 3 months after its termination on PTSD severity and dissociative tendencies. Even though the declining trend does not continue for PCL-5 and PHQ-9 in the follow-up measurements, their lower quartile levels are decreased in comparison to the pretest 2 lower quartiles. The fact that a decrease in CAPS-5 scores is observable at follow-up but not in PCL-5 scores may be linked to the self-reported nature of the second instrument. It would seem that subjectively, participants did not perceive their decrease in PTSD severity as lasting, but a more objective, exterior measure of the disorder's severity did. This finding is coherent with Lee et al. (2022) results, which also observed lower self-reported improvement in PTSD symptoms at posttest compared to clinician-administered CAPS-5 interview.

Figure 4 displays the CAPS-5 and PCL-5 for each participant for pretest 2, posttest, and follow-up. On both pretests, a total of 16 participants presented PCL-5 above 30 and 15 participants exhibited a CAPS-5 above two, which indicates clinical significance of the disorder at pretest. At follow-up, 15 participants had CAPS-5 scores lower than in pretest 2 (represented by a blue dot in Figure 4). From the remaining 5 subjects that did not decrease in CAPS-5 score at follow-up, three presented a decline from pretest 2 to posttest. In the case of PCL-5, 13 participants resulted in a lower measurement in the follow-up compared to pretest 2. Four participants declined from pretest 2 to posttest but increased from posttest to follow-up. A possible explanation for this is that punctual events happening in the 3 months gap between follow-up and posttest can trigger the rise of symptoms from a subjective point of view. Nonetheless, the VR protocol decreased PSTD severity from a subjective as well as an objective point of view, from pretests to posttest, for most participants. This suggests that this therapeutic modality might

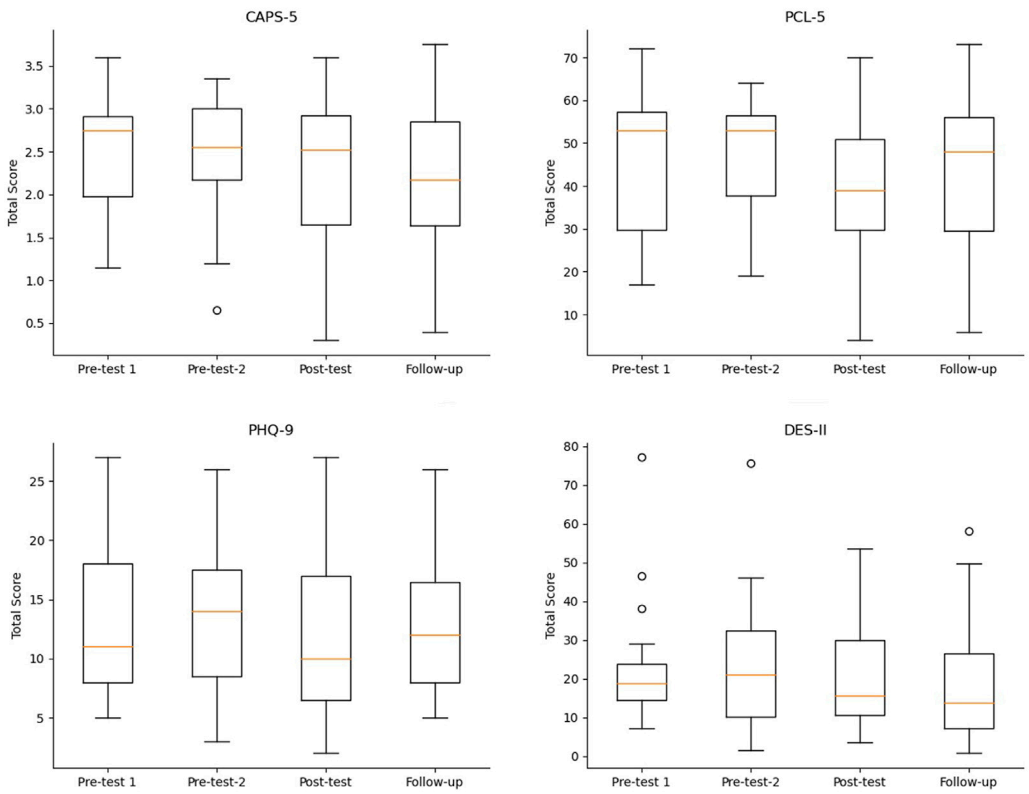


FIGURE 3
Boxplot showing the distributions from the psychological measurements.

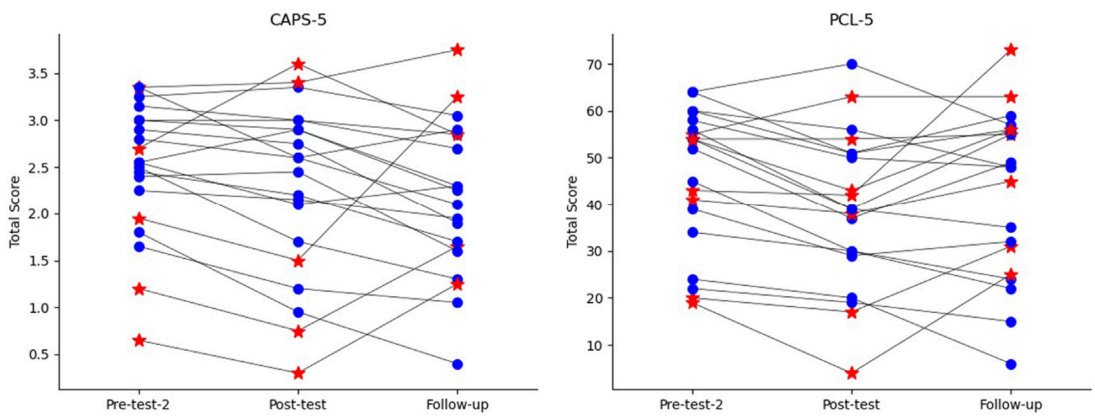


FIGURE 4
CAPS-5 and PCL-5 plots per subject.

be clinically more useful to reduce high peaks of PTSD symptoms in a short period of time, for example, when patients face temporary stressors.

Figure 5 displays the progression in standardized cognitive scores from pretest 1 to follow-up. It is possible to notice a rising trend between pretest 2 and posttest for processing speed and sustained attention variables, but the trend does not maintain on the follow-up measurements. Furthermore, from the curves of working memory and complex attention, a decrease in the standard deviation for the posttest analysis compared to the pretests can be observed. Therefore, a tendency to regroup around higher scores can be observed throughout the sample as the protocol progresses, even though no significant improvement occurred on these cognitive variables. Overall, a slight improvement in the cognitive faculties evaluated by the CNS Vitals Signs neurocognitive assessment was observed between pretest 1 and pretest 2, possibly as a result of the practice effect. However, Littleton et al. (2015) also observed this phenomenon in their

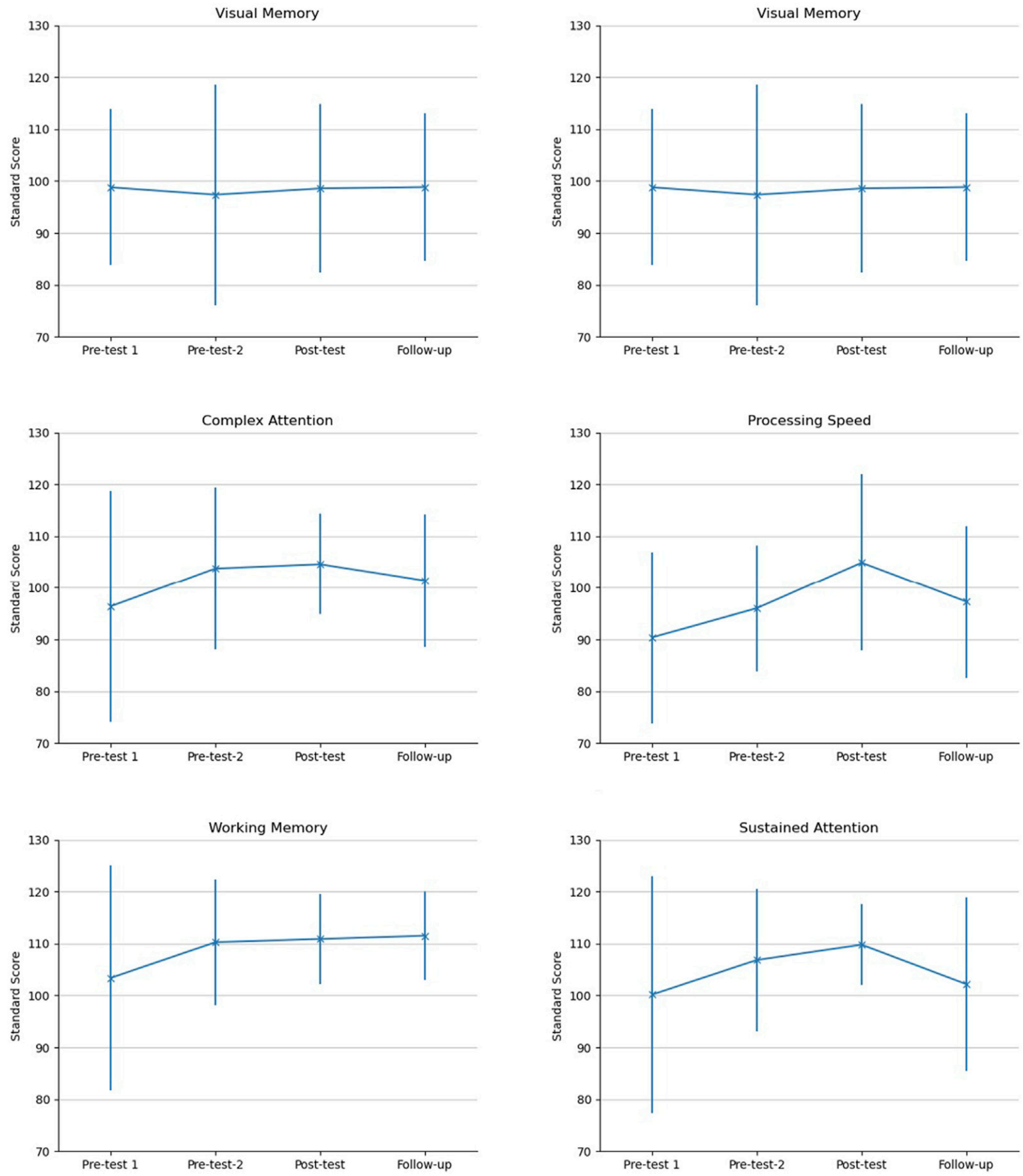


FIGURE 5
Mean and standard deviation of standard scores of cognitive variables.

validation study of the CNS Vitals Signs, and no additional improvement was observed in participants without training following the second testing session. These findings increase confidence in the results indicating improvement in cognitive function due to the VR immersion from pretest 2 to posttest.

Table 3 demonstrates the mean, standard deviation, and the result of the Friedman statistical test for the psychological variables and cognitive variables. All the psychological variables resulted in a statistical difference among the groups of measurements, while only two variables showed differences among the cognitive factors (i.e., processing speed, sustained attention). The following scores exhibit differences according to the *post hoc* analysis: PCL-5 (pretest 1 and posttest, $p < 0.05$ and pretest 2 and posttest, $p < 0.01$); CAPS-5 (pretest 1 and follow-up, $p < 0.05$ and pretest 2 and follow-up, $p < 0.05$); PHQ-9 (pretest 1 and posttest, $p < 0.05$); DES-II (pretest 1 and follow-up, $p < 0.05$); CNS Vitals Signs processing speed (pretest 1 and posttest, $p < 0.01$) and CNS Vitals Signs sustained attention (pretest 1 and posttest, $p < 0.05$).

Upon further analysis of the difference between groups of participants according to cut-off scores of psychological measures, no significant difference in cognitive variables was found linked to higher or lower scores of PHQ-9, CAPS-5, DES-II, or to the presence of psychotherapy during the intervention. Non-depressed people (PHQ-9 score below 10) tended to start at pretest 1 with higher overall cognitive scores on all variables, but improvement between pretest 1 and posttest stayed similar and exclusive to processing speed and sustained attention for both depressed and non-depressed participants. Improvement of cognitive functions following a relaxation immersion in VR is yet to be observed elsewhere in the scientific literature, which is why these findings deserve further consideration with a greater sample. Furthermore, it would seem

TABLE 3 Psychological and Cognitive tests mean (M) and standard deviation (SD).

Variable	Pretest 1		Pretest 2		Posttest		Follow-up		Friedman test	*p*-value
	M	SD	M	SD	M	SD	M	SD		
PDEQ	29.75	10.19								
PCL-5	46.85	16.06	45.9	14.64	39.10	16.16	42.7	17.43	13.35	0.004**
CAPS-5	2.49	0.73	2.47	0.70	2.27	0.91	2.16	0.81	11.31	0.01*
PHQ-9	13.05	5.92	13.15	5.59	11.74	6.37	13.00	5.66	9.0	0.03*
DES-II	22.57	15.82	24.38	17.42	19.82	13.19	19.48	15.57	8.7	0.034*
Verbal Memory	96.45	15.95	93.8	13.19	94.25	15.51	93.45	16.27	1.81	0.612
Visual Memory	97.35	15.91	96.85	20.83	97.7	16.23	98.7	13.82	0.3	0.961
Complex Attention	96.35	22.32	103.77	15.66	104.59	9.7	101.41	12.86	5.17	0.16
Processing Speed	89.8	16.35	95.05	12.6	103.95	17.13	96.55	14.67	14.89	0.002**
Working Memory	103.38	21.68	110.25	12.11	110.88	8.72	111.5	8.55	7.11	0.068
Sustained Attention	100.17	22.83	106.83	13.72	109.78	7.8	102.17	16.7	10.07	0.018*

* p-value < 0.05, ** p-value < 0.01, *** p-value < 0.001.

that cognitive improvement took place regardless of other conditions. Multiple studies have found that cognitive functioning is impaired in people with PTSD (Qureshi et al., 2011). If VR relaxation in nature proves to be efficient in this population to improve this particular problem, and thus even in the presence of comorbidity of depression, dissociative tendencies or a severe case of PSTD, it would be an addition to the clinical arsenal of therapeutic modalities for PTSD symptoms that impair daily functioning.

3.2 Physiological indicators

Figure 6 displays the HR and RMSSD distributions of the measures taken prior to and posterior to the tests. Table 4 complements the information displaying the mean and standard variation of the variables collected from HR and HRV. The mean values and standard deviation of heart rate are lower in the posttest and follow-up than during pretest two, showing a reverse tendency as expected for patients with PTSD that have higher heart rate than controls (Pole, 2007). In addition, it is possible to notice that heart rate data below the middle quartile measured before the subjects have done the tests gets lower towards the posttest when compared to the pretests. PTSD patients also show decreased HRV when compared to controls (Schneider and Schwerdtfeger, 2020). However, HVR scores above the middle quartile for the RMSSD variable increase towards the posttest in comparison to the after-test HVR scores on pretests 1 and 2 as displayed in Figure 6. Additional analysis of HR and HRV will be further explored from the biosignals collected from the participants during each section of immersion.

3.3 VR experience measurements

Figure 7 depicts the average and standard deviation results of the sense of presence and cybersickness for each session of the treatment. The results showed that the participants graded their sense of presence highly and almost no symptoms of cybersickness. These results and the rate of participants that finished the entire intervention reinforce the utilization of VR as a high motivator and safe tool.

3.4 Summary of findings

This study examined the feasibility and clinical outcomes of a multisensory VR nature immersion in participants suffering from PTSD symptoms. Preliminary results suggest that the procedure may be efficacious in promoting a statistically significant decline in patients' self-reported PTSD symptoms (PCL-5) and clinician-examined severity of PTSD (CAPS-5). Additionally, depressive symptoms declined from pretest to posttest, as well as dissociative symptoms. Decline in PTSD severity assessed by CAPS-5 interview and dissociative tendencies were maintained at the 3-month follow-up. Considering the analyzed cognitive variables, only two presented a statistically significant improvement in processing speed and sustained attention showing the benefits of the procedure on the cognitive aspects of the participants. Although we could not observe a statistically significant improvement, the variables of working memory and complex attention also present an increased average up to the posttest analysis. In addition, visual and verbal memory results stayed consistent across the period of VR immersion. These results show that improvement in cognition is not attributable to learning from repeated testing. These results display a reverse tendency between PTSD severity and processing speed as well as sustained attention, suggesting that a decrease in PTSD symptoms could have improved these cognitive functions in our participants. This interesting trend should be further explored with a greater sample size. Moreover, according to the questionnaires evaluating the experience, most of the participants graded the experience as highly immersive and free of considerable symptoms of cybersickness. All the outcomes give us an early indication that the treatment is feasible and safe. Notwithstanding the not-very-conclusive results from heart rate analysis, we will be analyzing the EEG and PPG signals

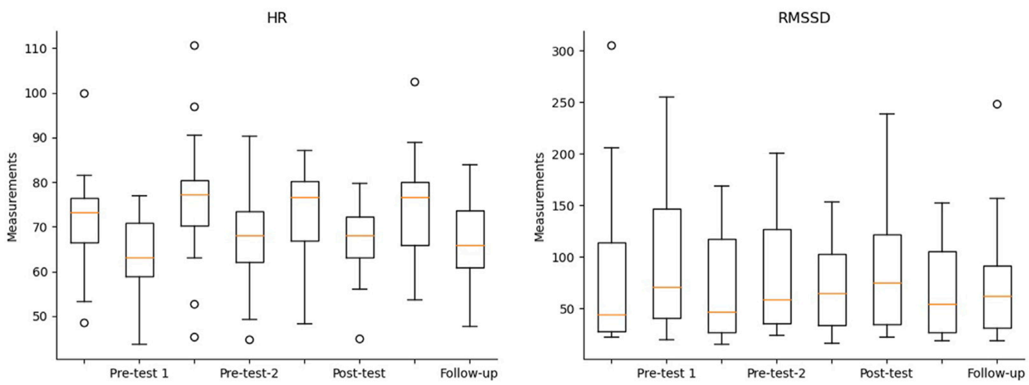


FIGURE 6
Boxplot showing the distributions of HR and HRV measurements.

TABLE 4 HR and HRV mean ± standard deviation.

Variable	Pretest 1		Pretest 2		Posttest		Follow-up	
	Before	After	Before	After	Before	After	Before	After
HR	71.45 ± 11.55	63.87 ± 8.23	76.49 ± 14.53	67.76 ± 10.77	73.15 ± 9.79	66.83 ± 7.98	74.56 ± 12.01	67.09 ± 9.76
IBI	875.68 ± 154.38	968.76 ± 142.71	825.59 ± 181.55	921.69 ± 164.61	846.88 ± 132.65	923.73 ± 130.57	835.5 ± 143.9	923.99 ± 140.3
MHRR	19.7 ± 11.20	17.86 ± 7.99	18.85 ± 9.97	18.38 ± 8.6	19.34 ± 9.58	18.06 ± 9.92	17.83 ± 8.70	17.48 ± 10.62
SDNN	89.23 ± 57.53	99.86 ± 48.06	79.97 ± 50.37	91.65 ± 49.5	84.34 ± 41.0	93.99 ± 43.97	77.72 ± 49.79	85.46 ± 46.46
RMSSD	86.34 ± 78.68	95.07 ± 65.7	74.26 ± 52.91	83.37 ± 56.96	73.09 ± 43.41	84.18 ± 55.43	66.89 ± 44.67	77.27 ± 58.06

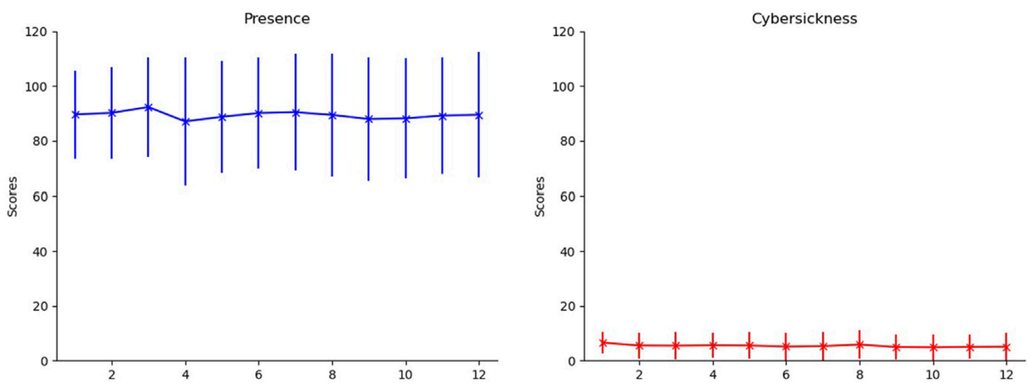


FIGURE 7
Mean and standard deviation of the scores for presence (blue) and cybersickness (red) for each section.

collected from each of the 12 sessions of VR nature immersion in order to investigate their behaviours across the duration of the protocol.

3.5 Limitations and future work

Limitations of this study include the small sample size of 20 participants and a lack of a control group. A larger study with a greater number of participants would enable the determination of a control group and would benefit the analysis of the efficacy of the intervention by the comparison of both groups. Furthermore, a group having the intervention with the olfactory stimulation compared with a group without the olfactory stimulation would provide an analysis of the separated contribution of smells to the intervention. Another limitation consisted of some participants rating the smells of the scenarios as not very realistic, mainly for the scenarios representing the hometown locations. Future studies will investigate the possibility of improvements by using different smells, such as multisensory pods, which provide a more realistic experience (Lopes et al., 2022). Lastly, while this paper has focused on the qualitative outcomes of the multisensory intervention, ongoing work includes analysis of the biosignals collected and potential quantitative outcomes on neural and bio-markers.

4 Conclusion

Although many virtual reality applications have been explored recently, most still rely on the stimulation of only hearing and vision senses. The work presented in this paper investigated the feasibility of a multisensory relaxing natural virtual reality application as a therapeutic modality for PTSD patients. Significant decreases in the severity of PTSD symptoms were seen as a result of the intervention. Ratings from CAPS-5, PCL-5, PHQ-9, and DES-II's scores significantly dropped after the intervention, and scores of CAPS-5 and DES-II also maintained a significant decrease three-month post-protocol. After the VR nature immersion, participants showed to have significantly improved their cognitive function levels of sustained attention and processing speed. However, the cognitive improvements did not persist up to the three-month follow-up measurements. The virtual reality experience was rated by participants as inducing a high level of presence with little to no negative effects of cybersickness, reinforcing its safety to use in a clinical setting.

Data availability statement

The datasets presented in this article are not yet readily available because they are being processed. The questionnaire data and the collected biosignals will be made available to the research community via the authors' website: https://musaelab.ca/resources. Requests to access the datasets should be directed to TF, tiago.falk@inrs.ca.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Research with Human Beings (CER-22-669) from INRS (University of Quebec). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

BD: Formal Analysis, Investigation, Visualization, Writing–original draft, Writing–review and editing. LP: Data curation, Writing–original draft, Writing–review and editing. ML: Writing–review and editing. M-CR: Conceptualization, Data curation, Writing–original draft, Writing–review and editing. AO: Writing–review and editing. TF: Conceptualization, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded partly by an NSERC Discovery grant to TF (RGPIN-2021-03246).

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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Aiken, M. P., and Berry, M. J. (2015). Posttraumatic stress disorder: possibilities for olfaction and virtual reality exposure therapy. *Virtual Real.* 19, 95–109. doi:10.1007/s10055-015-0260-x

Ashbaugh, A. R., Houle-Johnson, S., Herbert, C., El-Hage, W., and Brunet, A. (2016). Psychometric validation of the English and French versions of the posttraumatic stress disorder checklist for dsm-5 (pcl-5). *PloS one* 11, e0161645. doi:10.1371/journal.pone.0161645

Beard, C., Hsu, K., Rifkin, L., Busch, A., and Björgvinsson, T. (2016). Validation of the phq-9 in a psychiatric sample. *J. Affect. Disord.* 193, 267–273. doi:10.1016/j.jad.2015.12.075

Bomyea, J., and Lang, A. J. (2012). Emerging interventions for ptsd: future directions for clinical care and research. *Neuropharmacology* 62, 607–616. doi:10.1016/j.neuropharm.2011.05.028

Bouchard, S., Robillard, G., Renaud, P., and Bernier, F. (2011). Exploring new dimensions in the assessment of virtual reality induced side effects. *J. Comput. Inf. Technol.* 1, 20–32.

Carlson, E. B., and Putnam, F. W. (1993). An update on the dissociative experiences scale. *Dissociation Prog. dissociative Disord.* 6, 1, 16-27.

Cassani, R., Moinnereau, M.-A., and Falk, T. H. (2018). "A neurophysiological sensor-equipped head-mounted display for instrumental qoe assessment of immersive multimedia," in 2018 Tenth international conference on quality of multimedia experience (QoMEX) (IEEE), May 2018, Cagliari, Italy, 1–6.

Cassani, R., Moinnereau, M.-A., Ivanescu, L., Rosanne, O., and Falk, T. H. (2020). Neural interface instrumented virtual reality headsets: toward next-generation immersive applications. *IEEE Syst. Man, Cybern. Mag.* 6, 20–28. doi:10.1109/msmc.2019.2953627

Cleophas, T. J., Zwinderman, A. H., Cleophas, T. J., and Zwinderman, A. H. (2016). Non-parametric tests for three or more samples (friedman and kruskal-wallis). *Clinical data analysis on a pocket calculator: understanding the scientific methods of statistical reasoning and hypothesis testing*, Springer, Berlin, Germany, 193–197.

De Jesus, B., Jr, Lopes, M., Moinnereau, M.-A., Gougeh, R. A., Rosanne, O. M., Schubert, W., et al. (2022). "Quantifying multisensory immersive experiences using wearables: is (stimulating) more (senses) always merrier?," in Proceedings of the 2nd Workshop on Multisensory Experiences-SensoryX'22 (SBC). June 2022, Aveiro, Portugal, .

Emmelkamp, P. M., and Meyerbröker, K. (2021). Virtual reality therapy in mental health. *Annu. Rev. Clin. Psychol.* 17, 495–519. doi:10.1146/annurev-clinpsy-081219-115923

Forkus, S. R., Raudales, A. M., Rafiuddin, H. S., Weiss, N. H., Messman, B. A., and Contractor, A. A. (2022). The posttraumatic stress disorder (ptsd) checklist for dsm–5: A systematic review of existing psychometric evidence. *Clin. Psychol. Sci. Pract.* 30, 110–121. doi:10.1037/cps0000111

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., et al. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychol. Med.* 47, 2393–2400. doi:10.1017/s003329171700040x

Gualtieri, C. T., and Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, cns vital signs. *Archives Clin. Neuropsychology* 21, 623–643. doi:10.1016/j.acn.2006.05.007

Halbig, A., Babu, S. K., Gatter, S., Latoschik, M. E., Brukamp, K., and von Mammen, S. (2022). Opportunities and challenges of virtual reality in healthcare–a domain experts inquiry. *Front. Virtual Real.* 3, 14. doi:10.3389/frvir.2022.837616

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* 3, 203–220. doi:10.1207/s15327108ijap0303_3

Knaust, T., Felnhofer, A., Kothgassner, O. D., Höllmer, H., Gorzka, R.-J., and Schulz, H. (2022). Exposure to virtual nature: the impact of different immersion levels on skin conductance level, heart rate, and perceived relaxation. *Virtual Real.* 26, 925–938. doi:10.1007/s10055-021-00595-2

Kothgassner, O. D., Goreis, A., Kafka, J. X., Van Eickels, R. L., Plener, P. L., and Felnhofer, A. (2019). Virtual reality exposure therapy for posttraumatic stress disorder (ptsd): A meta-analysis. *Eur. J. psychotraumatology* 10, 1654782. doi:10.1080/20008198.2019.1654782

Kroenke, K., Spitzer, R. L., and Williams, J. B. (2001). The phq-9: validity of a brief depression severity measure. *J. general Intern. Med.* 16, 606–613. doi:10.1046/j.1525-1497.2001.016009606.x

Lee, D. J., Weathers, F. W., Thompson-Hollands, J., Sloan, D. M., and Marx, B. P. (2022). Concordance in ptsd symptom change between dsm-5 versions of the clinician-administered ptsd scale (caps-5) and ptsd checklist (pcl-5). *Psychol. Assess.* 34, 604–609. doi:10.1037/pas0001130

Littleton, A. C., Register-Mihalik, J. K., and Guskiewicz, K. M. (2015). Test-retest reliability of a computerized concussion test: cns vital signs. *Sports health* 7, 443–447. doi:10.1177/1941738115586997

Lo, J. C., Sehic, E., and Meijer, S. A. (2017). Measuring mental workload with low-cost and wearable sensors: insights into the accuracy, obtrusiveness, and research usability of three instruments. *J. cognitive Eng. Decis. Mak.* 11, 323–336. doi:10.1177/1555343417716040

Lopes, M. K., de Jesus, B. J., Moinnereau, M.-A., Gougeh, R. A., Rosanne, O. M., Schubert, W., et al. (2022). "Nat (ur) e: quantifying the relaxation potential of ultra-reality multisensory nature walk experiences," in 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE) (IEEE), October 2022, Rome, Italy, 459–464.

Marmar, C. R., Metzler, T. J., and Otte, C. (2004). *The peritraumatic dissociative experiences questionnaire*. New York, NY, USA, The Guilford Press.

Mistry, D., Zhu, J., Tremblay, P., Wekerle, C., Lanius, R., Jetly, R., et al. (2020). Meditating in virtual reality: proof-of-concept intervention for posttraumatic stress. *Psychol. Trauma Theory, Res. Pract. Policy* 12, 847–858. doi:10.1037/tra0000959

Noronha, H., and Campos, P. (2021). Harnessing virtual reality nature to promote well-being. *Interact. Comput.* 33, 353–366. doi:10.1093/iwc/iwac004

Pole, N. (2007). The psychophysiology of posttraumatic stress disorder: A meta-analysis. *Psychol. Bull.* 133, 725–746. doi:10.1037/0033-2909.133.5.725

Qureshi, S. U., Long, M. E., Bradshaw, M. R., Pyne, J. M., Magruder, K. M., Kimbrell, T., et al. (2011). Does ptsd impair cognition beyond the effect of trauma? *J. neuropsychiatry Clin. Neurosci.* 23, 16–28. doi:10.1176/appi.neuropsych.23.1.16

Ragsdale, K. A., Watkins, L. E., Sherrill, A. M., Zwiebach, L., and Rothbaum, B. O. (2020). Advances in ptsd treatment delivery: evidence base and future directions for intensive outpatient programs. *Curr. Treat. Options Psychiatry* 7, 291–300. doi:10.1007/s40501-020-00219-7

Rauch, S. A., Eftekhari, A., and Ruzek, J. I. (2012). Review of exposure therapy: A gold standard for ptsd treatment. *J. Rehabil. Res. Dev.* 49, 679–687. doi:10.1682/jrrd.2011.08.0152

Riches, S., Jeyarajaguru, P., Taylor, L., Fialho, C., Little, J., Ahmed, L., et al. (2023). Virtual reality relaxation for people with mental health conditions: A systematic review. *Soc. Psychiatry Psychiatric Epidemiol.* 58, 989–1007. doi:10.1007/s00127-022-02417-5

Rivest-Beauregard, M., Brunet, A., Gaston, L., Al Joboory, S., Trousselard, M., Simson, J.-P., et al. (2022). The clinician-administered ptsd scale for dsm-5 (caps-5) structured interview for ptsd: A French language validation study. *Psychol. Assess.* 34, e26–e31. doi:10.1037/pas0001099

Rizzo, A. S., and Shilling, R. (2017). Clinical virtual reality tools to advance the prevention, assessment, and treatment of ptsd. *Eur. J. psychotraumatology* 8, 1414560. doi:10.1080/20008198.2017.1414560

Robillard, G., Bouchard, S., Renaud, P., and Cournoyer, L. (2002). "Validation canadienne-française de deux mesures importantes en réalité virtuelle: l'immersive tendencies questionnaire et le presence questionnaire," in Poster presented at the 25e congrès annuel de la Société Québécoise pour la Recherche en Psychologie (SQRP), November 2002, Trois-Riviéres, Canada.

Schneider, M., and Schwerdtfeger, A. (2020). Autonomic dysfunction in posttraumatic stress disorder indexed by heart rate variability: A meta-analysis. *Psychol. Med.* 50, 1937–1948. doi:10.1017/s003329172000207x

Shalev, A., Liberzon, I., and Marmar, C. (2017). Post-traumatic stress disorder. *N. Engl. J. Med.* 376, 2459–2469. doi:10.1056/nejmra1612499

Taylor, S., Thordarson, D. S., Maxfield, L., Fedoroff, I. C., Lovell, K., and Ogrodniczuk, J. (2003). Comparative efficacy, speed, and adverse effects of three ptsd treatments: exposure therapy, emdr, and relaxation training. *J. Consult. Clin. Psychol.* 71, 330–338. doi:10.1037/0022-006x.71.2.330

Weathers, F. W., Bovin, M. J., Lee, D. J., Sloan, D. M., Schnurr, P. P., Kaloupek, D. G., et al. (2018). The clinician-administered ptsd scale for dsm–5 (caps-5): development and initial psychometric evaluation in military veterans. *Psychol. Assess.* 30, 383–395. doi:10.1037/pas0000486

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., Keane, T. M., et al. (1993). "The ptsd checklist (pcl): reliability, validity, and diagnostic utility," in Annual convention of the international society for traumatic stress studies, 462, San Antonio, TX, USA. October 1993.

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

TYPE Original Research
PUBLISHED 26 October 2023
DOI 10.3389/frvir.2023.1253155

OPEN ACCESS

EDITED BY
Orestis Georgiou,
Ultraleap Ltd., United Kingdom

REVIEWED BY
Juno Kim,
University of New South Wales, Australia
Hiroshi Ashida,
Kyoto University, Japan

*CORRESPONDENCE
Ramy Kirollos,
ramy.kirollos@drdc-rddc.gc.ca

RECEIVED 04 July 2023
ACCEPTED 02 October 2023
PUBLISHED 26 October 2023

CITATION
Kirollos R and Herdman CM (2023), Visual–vestibular sensory integration during congruent and incongruent self-rotation percepts using caloric vestibular stimulation.
Front. Virtual Real. 4:1253155.
doi: 10.3389/frvir.2023.1253155

Visual–vestibular sensory integration during congruent and incongruent self-rotation percepts using caloric vestibular stimulation

Ramy Kirollos[1,2]* and Chris M. Herdman[2]

[1]Toronto Research Centre, Defence Research and Development Canada, Toronto, ON, Canada,
[2]Visualization and Simulation Centre, Carleton University, Ottawa, ON, Canada

Introduction: The present study sets out to determine which sensory system mostly influences self-motion perception when visual and vestibular cues are in conflict. We paired caloric vestibular stimulation that signaled motion in either the clockwise or counter-clockwise direction with a visual display that indicated self-rotation in either the same or opposite directions.

Methods: In Experiment 1 (E1), caloric vestibular stimulation was used to produce vestibular circular vection. In Experiment 2 (E2), a virtual optokinetic drum was used to produce visual circular vection in a VR headset. Vection speed, direction, and duration were recorded using a potentiometer knob the participant controlled in E1 and E2. In Experiment 3 (E3), visual and vestibular stimuli were matched to be at approximately equal speeds across visual and vestibular modalities for each participant setting up Experiment 4 (E4). In E4, participants observed a moving visual pattern in a virtual reality (VR) headset while receiving caloric vestibular stimulation. Participants rotated the potentiometer knob while attending to visual–vestibular stimuli presentations to indicate their perceived circular vection. E4 had two conditions: 1) A congruent condition where calorics and visual display indicated circular vection in the same direction; 2) an incongruent condition where calorics and visual display indicated circular vection in opposite directions.

Results and discussion: There were equal reports of knob rotation in the direction consistent with the visual and vestibular self-rotation direction in the incongruent condition of E4 across trials. There were no significant differences in knob rotation speed and duration in both conditions. These results demonstrate that the brain appears to weigh visual and vestibular cues equally during a visual–vestibular conflict of approximately equal speeds. These results are most consistent with the optimal cue integration hypothesis.

KEYWORDS

virtual reality, visual–vestibular sensory integration, caloric vestibular stimulation, vection, optokinetic drum

1 Introduction

1.1 Vection

Vection is commonly defined as the illusory experience of self-motion while an individual is stationary (Dichgans and Brandt, 1978; Palmisano et al., 2015). An example of vection is sitting at a red light while an adjacent large vehicle such as a bus or truck pulls forward, creating the compelling illusion that the driver is moving backward despite being stationary. Circular vection is the illusion of perceived spinning about the yaw axis. Vection in all planes of motion such as illusory linear self-translation, or linear vection, and in the roll axis, referred to as tumbling, has also been produced and studied (Mach, 1875; Howard and Childerson, 1994; Riecke and Schulte-Pelkum, 2013). Vection is usually studied in the visual modality. However, there is research on vection in the auditory (Riecke et al., 2008; Keshavarz et al., 2014; Campos et al., 2018) and vestibular modalities (Fischer and Wodak, 1924; Cress et al., 1997; Fitzpatrick et al., 2002; St George et al., 2011; Fitzpatrick and Watson, 2015; Weech and Troje, 2017; Gallagher et al., 2020; Kirollos and Herdman, 2023).

Vection adds realism and immersion in training scenarios such as military and commercial flight simulation-based training, entertainment, and gaming (Riecke and Schulte-Pelkum, 2013). Understanding the behavioral (Brandt et al., 973; Gibson, 1966; Mach, 1875; Palmisano et al., 2000) and neurophysiological (Brandt et al., 1998; Nishiike et al., 2002; Kirollos et al., 2017; Berti et al., 2019) characteristics of vection has been of theoretical interest to researchers for over a century because of the sensory conflicts that can occur despite the illusion (Reason, 1978; Palmisano et al., 2000). An example of a sensory conflict from vection is any time visual displays, indicating acceleration/deceleration or direction change occurring with no corroborating information from other sensory systems, primarily the vestibular system, or *vice versa*.

1.2 Vestibular stimulation

The visual system excels in the detection of constant velocity motion and can also detect accelerations and decelerations (Howard, 1982). The vestibular system can detect accelerations/decelerations but cannot distinguish between constant velocity motion and being stationary because of the inertial properties of the endolymph fluid (Lishman and Lee, 1973). The peripheral vestibular apparatus is located in the inner ear and is made up of two otolith organs that detect linear accelerations (Purves et al., 2001). These are the saccule and the utricle. The semi-circular canals (SCCs) detect angular accelerations in the yaw, pitch, and roll axes. The horizontal SCC in each ear is positioned 30° below Earth's horizontal axis and specializes in detecting velocity change in the yaw axis (Baloh, 2003). The superior SCC is optimized for detecting the pitch-axis rotation velocity change and the posterior SCC detects the roll-axis velocity change (Rabbitt, 2019). The posterior and superior SCCs are both positioned in the vertical plane but are orthogonally configured. SCCs contain the cupular membrane in which hair cells are embedded within the endolymph fluid and housed in the ampullae. When the head rotates left to right, the endolymph fluid and embedded hair cells lag. This lag due to inertia is complimentary in the left and right horizontal SCCs and signals motion direction detection to the brain via the 8th cranial nerve (Bordoni et al., 2021).

There are various methods that can be used to stimulate and examine the peripheral vestibular organs' function without the individuals moving their heads. These include moving the individual (e.g., Barany chair and motion platform) and methods that do not require moving the individual. For instance, galvanic vestibular stimulation (GVS) uses electrodes that act on vestibular afferents when placed on the mastoid processes behind the ear. Some researchers have reported that GVS produces vestibular illusions of self-motion or vection (Fitzpatrick et al., 2002; Fitzpatrick and Day, 2004; St George et al., 2011; Fitzpatrick and Watson, 2015). However, vection generated by GVS can be brief and has been reported to produce a sudden tilting sensation rather than robust self-motion experience through space (Bense et al., 2001; Moore et al., 2011; Dilda et al., 2014; Aoyama et al., 2015).

Another method used to produce vestibular vection while an individual is stationary is by caloric vestibular stimulation (CVS). CVS alters the temperature of the endolymph fluid primarily in the horizontal SCC by administering a current of cool or warm air or water relative to body temperature via the external auditory canal (Barany, 1906). The air or water is preset to be different than body temperature, thus creating a thermal gradient that changes the endolymph fluid's density. A convection current produces a pressure change across the cupula after the endolymph fluid is sufficiently heated or cooled. The endolymph fluid shifts, resulting in perceived spinning and resembling what happens during the yaw motion of the head or vestibular vection.

CVS has seldom been used to study vestibular vection (Kirollos and Herdman, 2023). In contrast, CVS has primarily been used to assess vestibular health and function (Barany, 1906; Coats et al., 1976; Jacobson, 1993; Gonçalves et al., 2008; Sluga et al., 2021). The direct link between extra ocular muscles and SCCs allows for the use of CVS to trigger an eye movement called the vestibular ocular reflex (VOR) to assess vestibular health (Högyes, 1913; Goldberg et al., 1987). The VOR keeps the image steady on the retina when the head moves. CVS has also been used to study fluid dynamic properties in the inner ear (Meiry and Young, 1967; Kassemi et al., 2004; Kassemi et al., 2005; Santos et al., 2017; Rabbitt, 2019; Wu et al., 2021), assess spatial orientation in vestibular patients (Karnath, 1994; Moon et al., 2006), for neuroimaging of vestibular cortical regions (Frank et al., 2014; Frank and Greenlee, 2014; Frank et al., 2016; Klaus et al., 2020), and as a clinical intervention in patients with schizophrenia, psychosis, and psychopathy (Levy et al., 1983; Jones and Pivik, 1985).

1.3 Visual–vestibular sensory integration

A minority of studies on visual–vestibular sensory integration have found evidence supporting the notion of visual dominance during a conflict, specifically demonstrating that visual displays can influence posture in the absence of the vestibular input (Berthoz et al., 1975; Lee and Lishman, 1975; Lishman and Lee, 1973; Warren, 1895). Some studies have also demonstrated vestibular dominance during a conflict. For instance, Butler et al. (2010) found heading (i.e., the ability to perceive and distinguish the implied direction of self-motion from an optic flow pattern) to be more consistent with vestibular cues than with visual cues. Moreover, Harris et al. (2000) found that estimations of a person's traversed distance were more consistent with vestibular cues than with visual cues. However, most studies on visual–vestibular sensory integration have demonstrated that information from different sensory systems is fused in a statistically optimal fashion to reduce perceptual uncertainties, consistent with the optimal cue integration hypothesis (Clarke and Yuille, 1990; Telban and Cardullo, 2001; Ernst and Banks, 2002; Reymond et al., 2002; Gu et al., 2008; de Winkel et al., 2010; Fetsch et al., 2010; Jürgens and Becker, 2011; de Winkel et al., 2013; Jürgens et al., 2016; Rohde et al., 2016). The optimal cue integration hypothesis posits that the brain will rely on the sensory cue, providing the most reliable information, where "reliability" is defined as the inverse of a cue's variability (Fetsch et al., 2010). Some studies have reported deviations from the optimal cue integration hypothesis (de Winkel et al., 2010; Fetsch et al., 2012; de Winkel et al., 2015) and that cue reliability can be dynamic (Fetsch et al., 2009). Many of these studies pair a visual pattern of dots with a motion platform that moves linearly or angularly, generating congruent and incongruent conditions to study the heading and vection directions.

Weech and Troje (2017) showed that visual vection onset times are significantly reduced when noisy GVS (i.e., GVS signaling motion in no particular direction) is administered while viewing a visual display signaling vection. Visual vection has also been shown to be stronger when noisy vestibular signals are applied simultaneously compared to visual vection with no noisy GVS (Weech et al., 2018; Weech et al., 2020). Cress et al. (1997) reported significantly more convincing vection by participants when GVS that induced the roll axis tilt was combined with a visual display, indicating a tilt in the same direction compared to only viewing the visual display in the absence of GVS. The common finding between these three studies is that the reliability of the vestibular cue is decreased by noisy GVS and that the visual cues become relied upon by the individual. In all these studies, the unreliable vestibular signal generated by noisy GVS appears to cause a sensory reweighting to favor the visual cue, consistent with predictions made by the optimal cue integration hypothesis.

Kirollos and Herdman (2023) extended these findings using CVS in a recent experiment. In the first condition, participants received CVS with eyes closed. In the second condition, participants received CVS with eyes open, while they observed a stationary display signaling no self-motion in a virtual reality (VR) headset. Participants rotated a potentiometer knob that recorded the direction, speed, and duration of vection. Findings from the first condition indicated that circular vection can be induced in the vestibular system using CVS. In the second condition, participants still experienced vection despite a visual–vestibular conflict (i.e., visual cues signaling no motion while calorics acted as vestibular cues signaling spinning). However, vection was significantly shorter and slower during a conflict (Condition 2) than during no conflict (Condition 1). Kirollos and Herdman (2023) concluded, based on their results, that neither the visual nor vestibular systems dominated self-motion during a conflict and that the results were most consistent with the optimal cue integration hypothesis.

1.4 Present study and hypotheses

In the present study, we extended findings of Kirollos and Herdman (2023) by presenting participants with a visual display that moved during CVS administration. This is in contrast to the stationary display used in Kirollos and Herdman (2023). A visual–vestibular conflict was created in the present study using CVS to produce vestibular vection at approximately equal speed, as in the visual modality at the same time but in the opposite direction.

Four experiments were conducted in this paper with the same participants over 10 weeks. Perceived vection speed, direction, and duration were recorded using a potentiometer knob in Experiment 1 (E1) and Experiment 2 (E2) [see Kirollos and Herdman (2021) and Kirollos and Herdman (2023) for details and validation of this method].

In E1, participants received CVS with their eyes closed to confirm that they experienced vestibular circular vection. In E1, we hypothesized that CVS would induce vection in the clockwise (CW) direction in majority of left-ear cool air irrigations and counter-CW (CCW) vection in right-ear cool air irrigations. These findings would be consistent with the finding that cold air or water CVS relative to body temperature produces eye movements in the direction opposite to the ear being irrigated and that warm air or water CVS produces eye movements in the same direction as the ear being stimulated (Jacobson, 1993; Kirollos and Herdman, 2023). Therefore, during all cold CVS trials in this study, left-ear CVS should produce CW vection in the yaw-axis and right-ear CVS should produce CCW vection in the yaw-axis accordingly. We also hypothesized that vestibular vection speeds and durations would not differ substantially across trials because factors assumed to impact vestibular vection speed and duration, including temperature and air pressure, were held constant throughout the experiment (Wu et al., 2021).

In E2, the same participants observed a virtual optokinetic drum presented in a VR headset. They did not receive CVS. We hypothesized that the participants would experience vection in the direction indicated by the visual display in E2 at a speed that was faster than that indicated by the visual display, replicating Kirollos and Herdman (2021).

In Experiment 3 (E3), participants underwent one CVS trial at the start of the experiment. They were asked to remember their speed of spinning. The method of adjustment was then used by the participants to match the speed of vection they experienced from CVS to visual vection trials.

In Experiment 4 (E4), participants underwent two conditions: a) a congruent condition where visual and vestibular vection directions were the same and b) an incongruent condition where the visual and vestibular vection directions were different. Visual and vestibular directions in E4 were set to be the same based on the method of adjustment, matching the task performed in E3 for each participant. We hypothesized that participants would experience faster vection in the congruent condition compared to the incongruent condition. We also hypothesized that visual and vestibular cues would be used equally in deciding the self-motion direction in the incongruent condition.

FIGURE 1
ICS Air Caloric Irrigator. Photo credit: Ramy Kirollos.

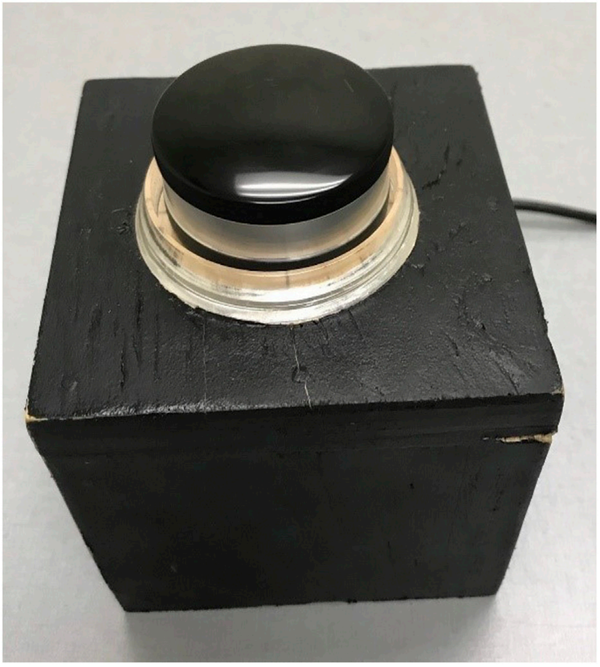


FIGURE 2
Potentiometer knob used by participants to index vection. Photo credit: Ramy Kirollos.

These findings would be consistent with the optimal cue integration hypothesis and findings from Kirollos and Herdman (2023).

2 Experiment 1: inducing vestibular vection using caloric vestibular stimulation

The objective of E1 was to assess if participants experienced vestibular vection using CVS. CVS with cold air was used to induce vestibular vection, replicating an experiment from Kirollos and Herdman (2023).

2.1 Methods

2.1.1 Participants

A total of 16 participants were recruited from Carleton University in accordance with an ethics package approved by the local university ethics board. Participants who reported visual or vestibular abnormalities, history of concussion, or who did not respond to vestibular vection were excluded from this study. Of these 16 participants, three participants did not sense vestibular vection in E1 and an additional three participants were removed because they experienced discomfort from CVS and did not wish to continue. A total of 10 participants (four female and three left-handed, M_{AGE} = 26.2, SD_{AGE} = 3.1) were included in the final analysis from E1. All participants were paid and agreed to participate in experiments 1–4.

2.1.2 Apparatus

2.1.2.1 ICS NCA 200 air caloric irrigator

The irrigator shown in Figure 1 delivered air via a glass speculum fitted with a disposable rubber tip to a participant's ear.

2.1.2.2 Control knob

The SpinTrak rotary potentiometer knob shown in Figure 2 was used to record circular vection speed, direction, and duration previously developed, validated, and used in Kirollos and Herdman (2021) and Kirollos and Herdman (2023). The knob was circular, had a diameter of 4.4 cm, and could only be turned CW or CCW indefinitely. It had a tachometer and a high-resolution pulse rate of 1,200 units over 360° for precise knob position tracking and recording. The knob was USB-integrated with custom software that logged turn rates in °/s at 75 Hz. It was housed in a custom-built wooden box and rested on the participant's stomach during testing.

2.1.2.3 Computer

The computer logging knob data comprised an Intel Core i7 processor, an NVIDIA GeForce GTX980 graphics card, and 16 GB RAM.

2.1.3 Stimuli

CVS trials were performed monaurally or one ear at a time. All CVS trials were performed at the same and constant temperature of 18°C and a constant air pressure of 10 L/min in all trials. These values were chosen to maximize chances of inducing vestibular vection. The participants underwent two left-ear and two right-ear irrigations, totaling four trials per participant. Each irrigation lasted between 90 and 180 s depending on when participants reported experiencing a robust spinning sensation in either the CW or CCW direction.

2.1.4 Procedure

The participants were tested individually in a dark and quiet room with eyes closed and wore a blindfold. They laid down in the supine posture on a table with their heads rested on a 30°-

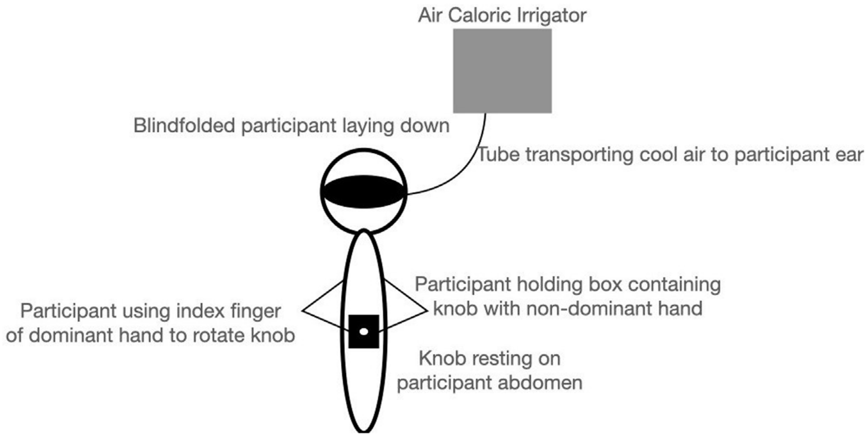


FIGURE 3
Schematic representation of a participant laying down in the supine posture blindfolded and holding a knob on their abdomen during CVS trials in E1.

angle wedge pillow to optimally stimulate horizontal SCC by ensuring it is parallel with Earth's horizontal axis (Baloh, 2003). The participants rested the potentiometer knob on their stomach and had their dominant hand's index finger positioned on the knob. They held the wooden box containing the knob with their non-dominant hand, as shown in Figure 3. The experimenter visually inspected the participants' ear for any obstruction that may reduce the efficacy of CVS before the experiment began.

The participants were asked to demonstrate CW and CCW rotation of the knob with their dominant index finger to the experimenter to avoid confusion about the direction during trial response periods. When irrigation began, the participants verbally described the self-motion direction and strength to the experimenter. Once a robust spinning sensation was reported, the participant rotated the knob in the direction and at the speed of the vection continuously for the length of the trial as long as they experienced a robust and consistent spinning sensation to match the speed of their vection. The knob had little friction, so the participant could easily continuously rotate it. The experimenter was in the room for the duration of the experimenter and observed the participant spinning the knob with their index finger. The experimenter ensured the knob direction was consistent with the verbally reported vection direction. The participants were instructed to verbally report when they no longer experienced vection and to stop spinning the knob, prompting the experimenter to end the trial. If participants reported swaying, rotation but with no precise direction, or no self-rotation, they were instructed not to rotate the knob. The participants were given at least 10-min breaks between CVS trials to allow time for endolymph fluid in the inner ear to reach the normal body temperature and resulting vestibular circular vection to subside. During breaks, the participants kept their eyes closed for the first 2 min to avoid any possible nausea and disorientation resulting from any lingering vection. This procedure was repeated four times, alternating left and right ear CVS. The experiment lasted approximately 90 min.

2.2 Results

Vection direction data were analyzed with a McNemar test. Vection speed and duration data were analyzed using within-subjects t-tests.

2.2.1 Vection direction

The participants experienced vection on 33 of 40 trials (82.5%). Of the 20 left-ear irrigation trials, the participants reported experiencing CW rotation on 12 trials and reported experiencing CCW rotation on four trials. Of the 20 right-ear irrigation trials, the participants reported experiencing CW rotation on 10 trials and reported experiencing CCW rotation on seven trials. A McNemar test revealed that there was no significant difference for the ear irrigated in the direction of perceived vection, $p = 0.50$.

2.2.2 Vection speed

The mean speed (°/s) from each trial in the left (M = 125.7, SD = 116.9) *vs.* right (M = 130.6, SD = 100.7) ear was compared with a within-subjects *t*-test. The effect of the ear irrigated on the speed of knob rotation was not significantly different ($t < 1$, df = 9).

2.2.3 Vection duration

The mean vection duration (s) from each trial in the left (M = 46.9, SD = 20.7) *vs.* right (M = 48.2, SD = 19.9) ear was compared with a within-subjects *t*-test. The effect of the ear irrigated on the duration of rotation was not significant ($t < 1$, $df = 9$).

2.3 Discussion

A difference was expected in the vection direction between the left- and right-ear irrigations. This is because according to the eye movement data in Jacobson (1993), cool air calorics in the left ear should result in vection in the CW direction and cool air calorics in the right ear should result in vection in the CCW direction, but we did not find this to be the case. This also conflicts with the findings

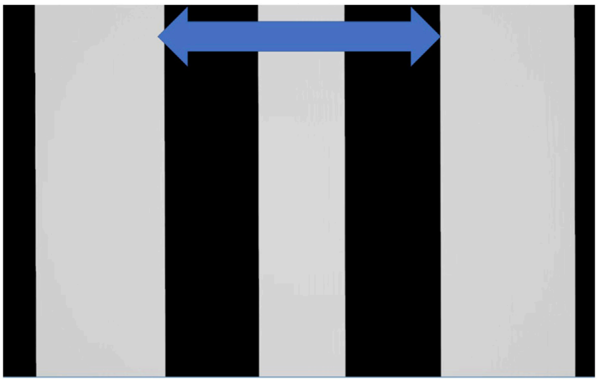


FIGURE 4
Virtual optokinetic drum presented in the VR headset in E2. The drum was presented in the VR headset and appeared in 3D while the participant lay supine with their heads tilted slightly forward on the wedge pillow. Stripes moved either CW or CCW.

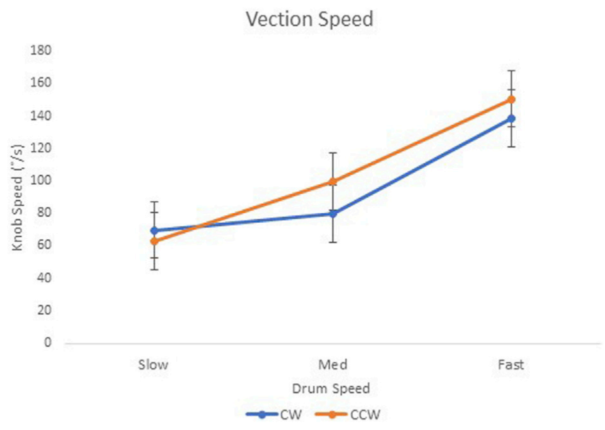


FIGURE 5
E2 knob rotation speed (°/s) when viewing three drum speeds in the clockwise ("CW") and counter-clockwise ("CCW") directions. Error bars represent 95% confidence intervals for the analysis.

from Kirollos and Herdman (2023), where cool air irrigation was used and majority of left-ear irrigations resulted in CW vection and majority of right-ear irrigations resulted in CCW vection. Although a significant difference in directions was found in Kirollos and Herdman (2023), there was variability in directions. Moreover, the sample was smaller in the current study (n = 10), compared to Kirollos and Herdman (2023) (n = 24), likely contributing to the variability in current results.

There were no significant differences between the left- and right-ear trials in the vection speed and duration. It was expected that there would be no difference in the vection speed and duration because temperature and air pressure were held constant throughout trials. Importantly, long vestibular vection durations upward of 45 s on average were observed, replicating findings from Kirollos and Herdman (2023).

We confirmed that vestibular vection can be experienced as demonstrated by the majority of trials in E1. However, the vection direction varied more than anticipated. It was important to confirm vestibular vection in E1 because vestibular vection is more variable and difficult to induce than visual vection, better ensuring that participants in E4 would experience a visual–vestibular conflict.

3 Experiment 2: inducing visual vection using a virtual reality headset

The objective of E2 was to ensure participants from E1 experienced visual vection. The 10 participants from E1 returned to the laboratory for E2. The participants viewed a vertical-striped virtual cylinder presented in a VR headset at three distinct constant velocities in the CW and CCW directions, totaling six conditions. We predicted that participants would experience visual vection and that the speed of knob rotation would be linked to the display speed. Based on findings from Kirollos and Herdman (2021), we predicted that participants would rotate the knob faster than the speed of the display and that the display speed would result in significantly different knob rotation speeds.

3.1 Methods

3.1.1 Apparatus

3.1.1.1 VR headset

An Oculus Rift DK2 VR headset provided a 110° diagonal visual angle, a native resolution of 960 × 1,080 pixels per eye, and a 75 Hz refresh rate. The left- and right-eye displays presented the same image at different perspectives, promoting a 3D perception of the stimuli. Other apparatus used were the knob and computer, as in E1.

3.1.2 Stimuli and design

The graphics display presented in the VR headset simulated a drum with a 200-cm diameter, as shown in Figure 4. When wearing the VR headset, the observer's viewpoint was set to the center of the virtual drum. The participant's body was aligned vertically with the cylinder and the stripes. Each stripe in the display corresponded to a width of 33 cm in the virtual graphics environment and subtended a horizontal visual angle of 10.85° at a virtual viewing distance of 100 cm. The vertical stripe pattern that was presented on the VR display moved in the CW or CCW directions at one of the three distinct speeds, 37.5°/s ("slow"), 56.25°/s ("medium"), and 75°/s ("fast"), resulting in six conditions. These speeds were chosen based on pilot tests from Kirollos and Herdman (2021) and replicate speeds used for experiments in that work.

The six conditions were presented six times each, totaling 36 trials that were presented in a random order to each participant. The 36 trials were split across four blocks (nine randomized trials per block). Each block lasted approximately 5 min and was followed by a 5-min break. The experiment lasted a total of 30 min.

3.1.3 Procedure

The participants laid supine in a dark room with their heads pitched forward on the 30° wedge pillow, as in E1. Once the participants were ready, the display was adjusted in the VR headset such that they were positioned in the center of the drum with the vertical lines of the drum being parallel to their body axis.

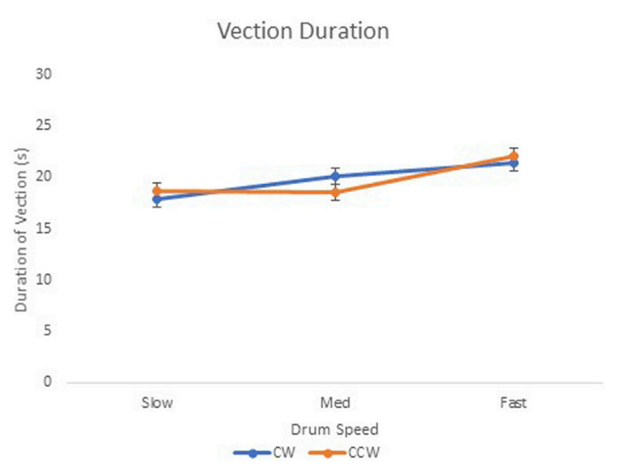


FIGURE 6
E2 vection duration (s) when viewing the three drum speeds (slow, medium, and fast) in the clockwise ("CW") and counter-clockwise ("CCW") directions. Error bars represent 95% confidence intervals for the analysis.

FIGURE 7
"Xbox 360 Controller" by Benjamin (2008). Licensed under CC BY-SA 2.0.

The VR headset display was mirrored on a monitor for the experimenter to have the same view as the participant. The participants observed 3–5 practice trials and rotated the knob at the speed, in the direction, and for the duration they experienced visual vection to become familiarized with the VR headset, stimulus, task, and potentiometer knob before the experiment began.

3.2 Results

Vection speed and duration were analyzed in separate 3 (drum speed: 37.5°/s *vs.* 56.25°/s *vs.* 75°/s) by 2 (direction: CW *vs.* CCW) repeated measures ANOVA. *Post hoc* comparisons were made using 95% confidence intervals (Jarmasz and Hollands, 2009).

3.2.1 Vection speed

Vection speed data are shown in Figure 5. There was a significant main effect of drum speed, $F(2, 18) = 4.96$, $p < .05$, and $R^2 = 0.36$, where faster drum speeds resulted in faster knob rotations. There was no significant main effect of direction ($F < 1$) on the knob rotation speed and no significant interaction between the drum speed and direction ($F < 1$) on the knob rotation speed. As shown in Figure 5, 95% confidence intervals indicated a significant difference between slow–fast and medium–fast conditions in the CW direction. In the CCW direction, there was a difference between all drum speed conditions.

3.2.2 Vection duration

Vection duration data are shown in Figure 6. There was a significant main effect for drum speed on vection duration, $F(2, 18) = 4.98$, $p < .05$, and $R^2 = .36$. The 95% confidence intervals indicated that there was a significant difference between the slow–medium and slow–fast conditions in the CW direction. There was a significant difference for slow–fast and medium–fast conditions in the CCW direction. There were no observed interaction effects between the drum speed and direction ($F < 1$).

3.3 Discussion

E2 showed that participants experienced visually induced vection when viewing a rotating virtual drum. As found previously, a faster drum speed resulted in faster and longer vection in E2 (Melcher and Henn, 1981; Palmisano and Gillam, 1998; Owens et al., 2018; Bos et al., 2019; Kirollos and Herdman, 2021). The results from E1 and E2 helped ensure that participants would perceive self-rotation in both the vestibular and visual modalities in E4.

4 Experiment 3: matching visual and vestibular vection speeds

The objective of E3 was to determine an approximately equal vection speed across the visual and vestibular modalities for each participant. This was to ensure that both visual and vestibular cues were perceived as equal in speed and, therefore, reliable when used in E4. In the current experiment, the participants from E1 and E2 underwent a vestibular–visual vection speed-matching task using the method of adjustment (Gescheider, 1997).

4.1 Methods

4.1.1 Apparatus, stimuli, and procedure

The ICS 200 Air Caloric Irrigator was set to 18°C and 10 L/min. One CVS trial was performed on each participant in either the left or right ear randomly while in the supine posture with their heads pitched forward at 30° on the wedge pillow. The irrigation lasted 90–180 s depending on when the participants reported vection, which they described to the experimenter verbally. The participants were instructed to monitor and remember the speed of vestibular vection from CVS while

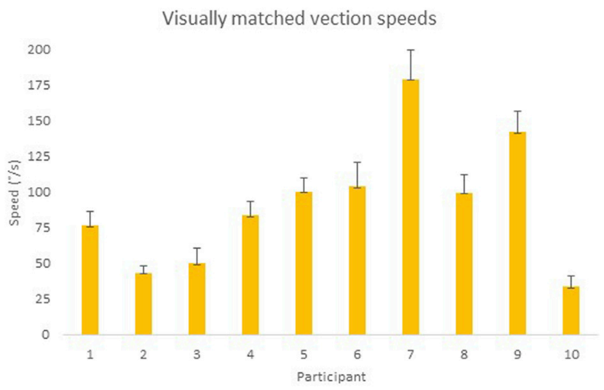


FIGURE 8
Means of the visual display speed that best represented each participant's vestibular vection experience in °/s. Error bars represent standard deviations.

their eyes were closed. The participants were then given a 10-min break.

After the CVS trial and follow-on break, the participants wore the VR headset and observed the optokinetic drum. The drum rotated in the CW or CCW directions 18 times each, totaling 36 trials. The participants were tasked with identifying a visual vection speed that best matched their vestibular vection speed from the CVS trial. Although the method used herein relies on participants remembering the stimulus, the participants were familiar with the process and experience of vestibular vection from E1. Thus, we anticipated participants could accurately perform this task, given the instructions and their previous experience with the stimuli from E1. Speeds of the drum varied from 20° to 350°/s randomly across trials. An Xbox 360 controller was used by the participant for them to select the visual drum speed that produced visual vection most similar to vestibular vection experienced during the caloric irrigation trial, as shown in Figure 7. The participants were instructed to use the left and right shoulder buttons of the Xbox 360 controller on each of the 36 trials to slow down (left shoulder button) or speed up (right shoulder button) the optokinetic drum by speed increments of 5°/s on each trial until they matched it to the vestibular vection speed experienced. When the participants identified the visual display that induced vection at the same speed experienced during CVS, they pressed the "A" button on the Xbox controller to log their response, ending the trial. The participants could only select a visual speed after a 10-s onset to ensure they experienced visual vection because the visual vection onset typically requires a few seconds to experience (Allison et al., 2012; Weech and Troje, 2017). The 36 trials were evenly split over four blocks each containing nine trials. Each block was separated by 5-min breaks to limit potential cybersickness. The experiment lasted approximately 45 min.

4.2 Results and discussion

Participants' mean visual vection speed matching their vestibular vection speed is shown in Figure 8. There was little within-subjects variability as the error bars representing standard deviations in Figure 8 indicate, implying that the participants were reasonably sure of the speed at which they felt they were moving during vestibular vection trials. The results highlight the substantial between-subjects variability in perceived self-rotation perception by air caloric vestibular stimulation, emphasizing the importance of performing the present experiment as visual–vestibular vection matching is experienced differently between individuals. The mean visual speeds for each participant obtained in this experiment were used as the preset visual display speed for that specific participant in E4. Vection after-effects were presumed to be negligible because of the random trial direction that should have cancelled out any such effects.

5 Experiment 4: self-motion perception during congruent and incongruent cues presented in the visual and vestibular modalities simultaneously

The main experiment in this study was E4. Based on the E3 results, we assumed that participants would perceive both visual and vestibular cues as equally reliable in E4 as their speeds were made similar. The goal of E4 was to determine how self-motion is decided when visual and vestibular signals are in conflict but approximately equal in terms of reliability. In the current experiment, the average visual vection speed each participant selected in E3 was set as the drum speed for that participant. There was a congruent condition and an incongruent condition. In the congruent condition, vection was signaled in the same direction in the visual and vestibular modalities. In the incongruent condition, vection was signaled in the opposite direction in the visual and vestibular modalities.

We hypothesized that the participants would experience faster and longer vection in the congruent condition than in the incongruent condition. We also hypothesized that the participants would use visual and vestibular cues approximately equally to judge their vection direction in the incongruent condition consistent with the optimal cue integration hypothesis (Fetsch et al., 2012). However, it was not clear how cue reliability might be expressed in our experiment using the potentiometer knob. For instance, it was possible that participants use the visual direction in one half and vestibular motion in the other half of incongruent trials. In addition, cue reliability in the incongruent condition could have resulted in vection cancelling out and participants not turning the knob. We did not anticipate that visual cues or vestibular cues would dominate a large majority of the incongruent condition directions as this would be inconsistent with the optimal cue integration hypothesis literature.

5.1 Methods

5.1.1 Apparatus and stimuli

The apparatus used were the Oculus VR headset, the ICS 200 Air Caloric Irrigator, and the potentiometer knob from E1 and E2. The vestibular stimulus was 18°C air caloric irrigation administered at 10 L/min. The Oculus DK2 device was used to present the optokinetic drum to the participant. The experimenter set the

TABLE 1 Vection direction frequency counts during congruent and incongruent conditions by trials across participants in E4.

Congruent (visual + vestibular)	Incongruent (direction consistent with visual vection)	Incongruent (direction consistent with vestibular vection)
15	8	8

Incongruent trials are split into responses consistent with the visual vection direction and responses consistent with the vestibular vection direction.

speed of the visual display for each participant at the start of the experiment based on the participant's mean drum speed obtained in E3. Each participant underwent two congruent trials and two incongruent trials. Left- and right-ear irrigations were alternated across trials, and the congruent and incongruent conditions were counter-balanced, as in E1. The participants were naïve as to whether they were in a congruent or incongruent trial.

5.1.2 Procedure

The procedure was identical to that in E1 (Section 2.1.4). However, when the participants verbally reported the direction of vestibular vection, the experimenter set the motion of the virtual drum in either the direction congruent with participant's reported vestibular vection direction or opposite to the direction of their vestibular vection based on whether the participant was in a congruent or incongruent condition. The participants opened their eyes and viewed the moving drum in the VR headset and rotated the knob in the direction and at the speed at which they felt they were moving. The experiment lasted approximately 90 min.

5.2 Results

Exactly equal numbers of visual and vestibular direction reports were found in the 16 incongruent trials. Vection speed and duration data were analyzed with four within-subjects t-tests. The first two t-tests compared congruent *vs.* incongruent responses for speed and duration data, respectively, to determine if the vection speed and duration were different. The third and fourth t-tests compared visual *vs.* vestibular responses for speed and duration, respectively, to determine if one modality yielded different vection results.

5.2.1 Vection direction

Two of the 10 participants did not experience vestibular vection and were, therefore, excluded from the analysis. Each of the remaining eight participants completed four trials: two congruent and two incongruent trials. This totaled 32 trials in the experiment: 16 congruent trials and 16 incongruent trials. Self-rotation was reported on 31 of the 32 trials (97%): one congruent trial did not yield a response by a participant. The vection direction data are summarized in Table 1. In the incongruent condition, the participants reported vection in the direction indicated by vestibular stimulation on eight trials and reported vection in the visual direction on the remaining eight trials. Of the eight participants, three reported vection consistent with the visual cues for vection on both incongruent trials and three reported vection consistent with the vestibular cues for vection on both incongruent trials. The two remaining participants experienced vection consistent with the visual direction of vection in one trial and based on the vestibular direction of motion in the other trial.

5.2.2 Vection speed

A within-subjects t-test for speed for the congruent (M = 177.3, SD = 145.5) *vs.* incongruent (M = 164.6, SD = 152) condition revealed no significant difference, t(7) = 5.70, and p = .587. The within-subjects t-test for vection speed responses in the incongruent condition, comparing the visual vection direction (M = 137.2, SD = 95.7) to vestibular vection direction (M = 147.8, SD = 190.5), indicated no significant difference, t(4) = −0.176, and p = .869.

5.2.3 Vection duration

A within-subjects t-test for duration (s) for the congruent (M = 99.3, SD = 34.2) *vs.* incongruent (M = 82.9, SD = 35.8) condition revealed no significant difference, t(7) = 1.194, and p = .271. A within-subjects t-test for vection duration responses in the incongruent condition, comparing the visual vection direction (M = 91.2, SD = 31.5) to vestibular vection direction (M = 69, SD = 45.7), indicated no significant difference, t(4) = 1.010, and p = .37.

5.3 Discussion

5.3.1 Vection direction

The objective of E4 was to determine the self-motion direction when the visual and vestibular systems received cues that are in conflict but approximately equally reliable. The results for the incongruent condition showed that three participants used visual cues to indicate self-motion, three participants used vestibular cues, and the remaining two participants used visual cues on one incongruent trial and vestibular cues on the other incongruent trial. Therefore, visual and vestibular self-motion were split evenly across trials and participants.

5.3.2 Vection speed and duration

We predicted that the vection speed and duration would be slower in the incongruent condition than in the congruent condition. Slower and shorter vection would be consistent with our previous findings wherein vestibular vection from CVS was experienced during the visual–vestibular conflict when showing a stationary visual display compared to when there is no conflicting visual display (Kirollos and Herdman, 2023). However, there were no significant differences for the vection speed and duration in the congruent and incongruent conditions of the current experiment. In this experiment, both the visual and vestibular cues signaled vection. This is in contrast to our previous experiment, in which the visual display did not signal vection but the vestibular stimulus did (Kirollos and Herdman, 2023). It could be the case that the stationary display in Kirollos and Herdman (2023) was a) easier to perceive as conflicting than the incongruent condition in the present study and/or b) the stationary display acted as a reference to attenuate fixations and vection. Moreover, the findings here may come as a result of the smaller sample size compared to that of Kirollos and Herdman (2023).

We also compared the vection speed and duration in the incongruent responses, categorizing responses by whether the participant vection direction was consistent with the visual or vestibular cues. We did not predict that there would be significant differences between the visual and vestibular speeds and durations because visual and vestibular speeds were set to be approximately equal. The results were consistent with this prediction.

6 General discussion

6.1 Summary and interpretation of findings

The goal of the present study was to examine how a visual–vestibular conflict is resolved during perceived self-motion. Sixteen participants were tested over the span of 10 weeks. Of these 16 participants, three participants were removed because they did not sense vestibular vection and additional three participants were removed because they experienced discomfort and did not wish to continue in E1. An additional two participants were removed in E4 because they did not sense vestibular vection.

In E1, vestibular vection was induced with CVS. Vestibular vection direction, speed, and duration were measured with the potentiometer knob. Long vestibular vection durations were observed, replicating the findings from Kirollos and Herdman (2023). However, vection direction findings did not replicate those in our previous study. In E2, a virtual optokinetic drum was used to induce and record visual vection in a VR headset. As in Kirollos and Herdman (2021), virtual drum speeds significantly impacted visual vection speeds. In E3, the goal was to identify vection speeds for the visual and vestibular systems that were approximately equal for each participant with the method of adjustment. Visual vection speeds from E3 for each participant were then used in E4.

The results from E4 showed that visual cues were used for deciding self-motion on one half of the trials and that vestibular cues were used in the other half of the trials. Therefore, there was no evidence for either visual or vestibular dominance. The results are most consistent with the optimal cue integration hypothesis, where it was predicted that participants would use the most reliable cue (visual or vestibular) in each trial to determine the self-motion direction. The equal number of responses indicating visual and vestibular cue reliance suggests that the cue matching from E3 successfully attained equal speeds of vection, as intended. The visual and vestibular vection speeds in the incongruent condition in E4 were not significantly different from each other, demonstrating that vection speeds were approximately equal across modalities.

Although not significant, vection speed was the fastest in the congruent condition and slowest in the visual incongruent condition. These data follow a trend consistent with the results in Kirollos and Herdman (2023): vection is faster when there is no visual–vestibular conflict and suggests that visual and vestibular dominance hypotheses cannot explain the results.

6.2 Relationship of findings to the relevant behavioral optimal cue integration literature

In a study on heading perception by Fetsch et al. (2009), cue reliability was directly manipulated. Fetsch et al. looked at the heading responses during a conflict with varying levels of coherence or cue reliability, where coherence was changed every trial. They found that heading responses were mediated by the level of coherence of the visual display. Therefore, the more coherent the visual display, the more heading reports were based on the visual stimulus-heading direction. The less coherent the visual display, the more the vestibular heading direction was relied upon across trials. These findings demonstrate that cue reliability is a representation of the strength of a cue. In Fetsch et al. (2009), reliability and, therefore, stimulus strength were mediated by manipulating the coherence of the elements in the optic flow display. In E4 of the current study, reliability was modulated by setting the speed to be perceptually equal across the visual and vestibular systems in all trials. The results from our current study extend findings by Fetsch et al. (2009) and others on the optimal cue integration hypothesis as they indicate that when reliability or strength is characterized by the speed of the stimuli and those speeds are approximately equal, visual and vestibular cues are used to decide the self-motion direction equally, represented by the direction frequency count.

6.3 Relationship of findings to the relevant neuroimaging literature

Debate regarding visual dominance, vestibular dominance, and optimal cue integration can analogously be found in the neuroimaging literature. A brain imaging study by positron emission topography by Brandt et al. (1998) found that during a conflict generated from visual vection, metabolic activity in vestibular cortical-processing regions such as the insula becomes inhibited. Brandt and colleagues reported that activity in visual cortical regions increased significantly during visual vection. The authors suggested that these results provide neurophysiological support for visual dominance hypotheses for resolving a conflict during vection. Nishiike et al. (2002) used a brain-imaging method called magnetoencephalography during visual vection displays and showed that both vestibular and visual regions became more active during vection. Therefore, the findings from Brandt et al. and Nishiike et al. appear to contradict one another. In a functional magnetic resonance imaging study, Kirollos et al. (2017) found that vestibular regions called the posterior insular cortex (PIC) and parieto-insular vestibular cortex (PIVC) became active during high-conflict visual displays inducing vection but not low-conflict visual displays inducing vection, potentially explaining the difference between the findings of Brandt et al. and Nishiike et al. In Kirollos et al.

(2017), high-conflict displays simulated forward vection with an up–down oscillation component, thereby creating a sustained visual–vestibular conflict throughout the display presentation. Low-conflict displays simulated smooth forward vection and therefore did not create a continuous visual–vestibular conflict. Frank et al. (2016) showed that PIC activity correlated with both visual and vestibular stimulation, whereas PIVC activity correlated with vestibular self-motion stimuli only. The results from Kirollos et al. agree with those of Nishiike et al. and Frank et al. (2016) as they all show that some vestibular processing areas (PIC and/or PIVC) become active during visual vection processing. These findings are consistent with the behavioral findings from E4 that showed that during conflict, the visual and vestibular cues are used to determine the self-motion direction [see Frank and Greenlee. (2018) for a review of visual–vestibular cortical interaction during conflict]. These findings are also consistent with a cybersickness study (Weech et al., 2018; Weech et al., 2020), indicating that unreliable vestibular signals result in a sensory down-weighting of vestibular cue reliability, producing a less sensory conflict and, therefore, less sickness. The vection literature also provides examples of sensory down-weighting of unreliable vestibular cues (Cress et al., 1997; Gallagher et al., 2020).

In summary, the findings of the present study were consistent with the neuroimaging results, showing that both visual and vestibular systems and their neural correlates are involved in processing and deciding the self-motion direction rather than cortical areas associated with one system becoming inhibited. It was unclear whether vestibular dominance or optimal cue integration explained the results during a conflict, where the visual display signaled no motion in Kirollos and Herdman (2023) as the visual cue signaled no motion but the vestibular cue did. However, in the present study, E4 showed that when the visual stimulus is matched for the vection speed to the vestibular stimulus, both visual and vestibular systems are used to determine self-motion equally.

7 Conclusions, future work, and applications

Converging evidence from our study, recent optimal cue integration findings, and adjacent literature reports on neuroimaging and motion sickness appear to oppose dominance hypotheses. Our research has demonstrated the ability to produce circular vection using CVS. However, the use of monaural CVS presents variability and proved to be a less-specific stimulus to isolate self-motion perception in horizontal SCC than expected. For instance, a limitation to our findings is that E1 results were inconsistent with the expected vestibular vection direction from our previous study and eye movement data (Jacobson, 1993; Kirollos and Herdman, 2023). Future research should replicate methods used herein using binaural CVS, stimulating one ear with cool air and the other ear with warm air simultaneously to produce more consistent vestibular vection in the yaw-axis with a larger sample. Future research can also record eye movements during CVS to correlate vection speed from the knob data with nystagmus data and compare it to visual vection eye movement data. Adding binaural CVS and eye movement recording can provide a more robust testbed to replicate and expand on the current findings. Finally, altering caloric vestibular irrigation air temperatures can help identify the correlation between the temperature speed and perceived CVS, and can produce a broader range of vestibular vection for applications requiring it, such as spatial disorientation training for aviators.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Carleton University Human Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

RK: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, visualization, writing–original draft, and writing–review and editing. CH: supervision and resources. All authors contributed to the article and approved the submitted version.

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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Allison, R. S., Zacher, J. E., Kirollos, R., Guterman, P. S., and Palmisano, S. (2012). Perception of smooth and perturbed vection in short-duration microgravity. *Exp. Brain Res.* 223 (4), 479–487. doi:10.1007/s00221-012-3275-5

Aoyama, K., Iizuka, H., Ando, H., and Maeda, T. (2015). Four-pole galvanic vestibular stimulation causes body sway about three axes. *Sci. Rep.* 5 (1), 10168–8. doi:10.1038/srep10168

Baloh, R. W. (2003). "Vestibular system," in *Encyclopedia of the neurological sciences*. Editors M. J. Aminoff and R. B. Daroff (Waltham MA: Elsevier), 661–671.

Barany, R. (1906). Untersuchungen über den vom Vestibularapparat des Ohres reflektorisch ausgelösten rhythmischen Nystagmus und seine Begleiterscheinungen. *Oscar Coblentz*.

Benjamin, N. (2008). XBOX 360 controller. Avaliable At: https://www.flickr.com/photos/28526479@N02/2902721172; Licensed under CC BY-SA 2.0.

Bense, S., Stephan, T., Yousry, T. A., Brandt, T., and Dieterich, M. (2001). Multisensory cortical signal increases and decreases during vestibular galvanic stimulation (fMRI). *J. neurophysiology* 85 (2), 886–899. doi:10.1152/jn.2001.85.2.886

Berthoz, A., Pavard, B., and Young, L. R. (1975). Perception of linear horizontal self-motion induced by peripheral vision (linearvection) basic characteristics and visual-vestibular interactions. *Exp. Brain Res.* 23 (5), 471–489. doi:10.1007/bf00234916

Berti, S., Haycock, B., Adler, J., and Keshavarz, B. (2019). Early cortical processing of vection-inducing visual stimulation as measured by event-related brain potentials (ERP). *Displays* 58, 56–65. doi:10.1016/j.displa.2018.10.002

Bordoni, B., Mankowski, N. L., and Daly, D. T. (2021). "Neuroanatomy, cranial nerve 8 (vestibulocochlear)," in *StatPearls* (Treasure Island, FL: StatPearls Publishing).

Bos, J. E., van den Berg-Kroon, E. L., Houben, M. M., and Kuiper, O. O. (2019). In-car speed estimations with real, virtual, and no view. *Displays* 58, 66–70. doi:10.1016/j.displa.2019.01.001

Brandt, T., Bartenstein, P., Janek, A., and Dieterich, M. (1998). Reciprocal inhibitory visual-vestibular interaction. Visual motion stimulation deactivates the parieto-insular vestibular cortex. *Brain a J. neurology* 121 (9), 1749–1758. doi:10.1093/brain/121.9.1749

Brandt, T., Dichgans, J., and Koenig, E. (1973). Differential effects of central versus peripheral vision on egocentric and exocentric motion perception. *Exp. Brain Res.* 16 (5), 476–491. doi:10.1007/bf00234474

Butler, J. S., Smith, S. T., Campos, J. L., and Bülthoff, H. H. (2010). Bayesian integration of visual and vestibular signals for heading. *J. Vis.* 10 (11), 23. doi:10.1167/10.11.23

Campos, J., Ramkhalawansingh, R., and Pichora-Fuller, M. K. (2018). Hearing, self-motion perception, mobility, and aging. *Hear. Res.* 369, 42–55. doi:10.1016/j.heares.2018.03.025

Clarke, J. J., and Yuille, A. L. (1990). *Data fusion for sensory information processing*. Boston, MA: Kluwer Academic, 978–971.

Coats, A. C., Hebert, F., and Atwood, G. R. (1976). The air caloric test: a parametric study. *Archives Otolaryngology* 102 (6), 343–354. doi:10.1001/archotol.1976.00780110055005

Cress, J. D., Hettinger, L. J., Cunningham, J. A., Riccio, G. E., Haas, M. W., and McMillan, G. R. (1997). Integrating vestibular displays for VE and airborne applications. *IEEE Comput. Graph. Appl.* 17 (6), 46–52. doi:10.1109/38.626969

de Winkel, K. N., Katliar, M., and Bülthoff, H. H. (2015). Forced fusion in multisensory heading estimation. *PLoS One* 10 (5), e0127104. doi:10.1371/journal.pone.0127104

de Winkel, K. N., Soyka, F., Barnett-Cowan, M., Bülthoff, H. H., Groen, E. L., and Werkhoven, P. J. (2013). Integration of visual and inertial cues in the perception of angular self-motion. *Exp. Brain Res.* 231 (2), 209–218. doi:10.1007/s00221-013-3683-1

de Winkel, K. N., Weesie, J., Werkhoven, P. J., and Groen, E. L. (2010). Integration of visual and inertial cues in perceived heading of self-motion. *J. Vis.* 10 (12), 1. doi:10.1167/10.12.1

Dichgans, J., and Brandt, T. (1978). "Visual-vestibular interaction: effects on self-motion perception and postural control," in *Perception* (Springer), 755–804.

Dilda, V., Morris, T. R., Yungher, D. A., MacDougall, H. G., and Moore, S. T. (2014). Central adaptation to repeated galvanic vestibular stimulation: implications for pre-flight astronaut training. *PLoS One* 9 (11), e112131. doi:10.1371/journal.pone.0112131

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415 (6870), 429–433. doi:10.1038/415429a

Fetsch, C. R., DeAngelis, G. C., and Angelaki, D. E. (2010). Visual–vestibular cue integration for heading perception: applications of optimal cue integration theory. *Eur. J. Neurosci.* 31 (10), 1721–1729. doi:10.1111/j.1460-9568.2010.07207.x

Fetsch, C. R., Pouget, A., DeAngelis, G. C., and Angelaki, D. E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nat. Neurosci.* 15 (1), 146–154. doi:10.1038/nn.2983

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., and Angelaki, D. E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. *J. Neurosci.* 29 (49), 15601–15612. doi:10.1523/jneurosci.2574-09.2009

Fischer, M. H., and Wodak, E. (1924). Unbekannte vestibulariseffekte bei gleichzeitiger äqualer doppelspülung. *Klin. Wochenschr.* 3 (31), 1406–1407. doi:10.1007/bf01852444

Fitzpatrick, R. C., and Day, B. L. (2004). Probing the human vestibular system with galvanic stimulation. *J. Appl. Physiology* 96 (6), 2301–2316. doi:10.1152/japplphysiol.00008.2004

Fitzpatrick, R. C., Marsden, J., Lord, S. R., and Day, B. L. (2002). Galvanic vestibular stimulation evokes sensations of body rotation. *Neuroreport* 13 (18), 2379–2383. doi:10.1097/00001756-200212200-00001

Fitzpatrick, R. C., and Watson, S. R. (2015). Passive motion reduces vestibular balance and perceptual responses. *J. physiology* 593 (10), 2389–2398. doi:10.1113/jp270334

Frank, S. M., Baumann, O., Mattingley, J. B., and Greenlee, M. W. (2014). Vestibular and visual responses in human posterior insular cortex. *J. neurophysiology* 112 (10), 2481–2491. doi:10.1152/jn.00078.2014

Frank, S. M., and Greenlee, M. W. (2014). An MRI-compatible caloric stimulation device for the investigation of human vestibular cortex. *J. Neurosci. Methods* 235, 208–218. doi:10.1016/j.jneumeth.2014.07.008

Frank, S. M., and Greenlee, M. W. (2018). The parieto-insular vestibular cortex in humans: more than a single area? *J. neurophysiology* 120 (3), 1438–1450. doi:10.1152/jn.00907.2017

Frank, S. M., Wirth, A. M., and Greenlee, M. W. (2016). Visual-vestibular processing in the human Sylvian fissure. *J. neurophysiology* 116 (2), 263–271. doi:10.1152/jn.00009.2016

Gallagher, M., Choi, R., and Ferrè, E. R. (2020). Multisensory interactions in virtual reality: optic flow reduces vestibular sensitivity, but only for congruent planes of motion. *Multisensory Res.* 1, 625–644. doi:10.1163/22134808-20201487

Gescheider, G. A. (1997). *Psychophysics: the fundamentals*. United Kingdom: Taylor & Francis Group.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin Company.

Goldberg, J., Highstein, S. M., Moschovakis, A. K., and Fernandez, C. (1987). Inputs from regularly and irregularly discharging vestibular nerve afferents to secondary neurons in the vestibular nuclei of the squirrel monkey. I. An electrophysiological analysis. *J. neurophysiology* 58 (4), 700–718. doi:10.1152/jn.1987.58.4.700

Gonçalves, D. U., Felipe, L., and Lima, T. M. A. (2008). Interpretation and use of caloric testing. *Rev. Bras. Otorrinolaringol.* 74, 440–446. doi:10.1016/s1808-8694(15)30580-2

Gu, Y., Angelaki, D. E., and DeAngelis, G. C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nat. Neurosci.* 11 (10), 1201–1210. doi:10.1038/nn.2191

Harris, L. R., Jenkin, M., and Zikovitz, D. C. (2000). Visual and non-visual cues in the perception of linear self motion. *Exp. Brain Res.* 135 (1), 12–21. doi:10.1007/s002210000504

Högyes, A. (1913). *Über den Nervenmechanismus der assoziierten Augenbewegungen: von Andreas Högyes. Übers. von Martin Sugár.* Urban & Schwarzenberg.

Howard, I. P. (1982). *Human visual orientation*. Chichester; New York: J. Wiley.

Howard, I. P., and Childerson, L. (1994). The contribution of motion, the visual frame, and visual polarity to sensations of body tilt. *Perception* 23 (7), 753–762. doi:10.1068/p230753

Jacobson, G. P., and Newman, C. W. (1993). "Background and technique of caloric testing," in *Handbook of balance function testing*. Editors G. P. Jacobson, C. W. Newman, and J. M. Kartush (Mosby, St. Louis, MO: Singular Publishing Group), 156–192.

Jarmasz, J., and Hollands, J. G. (2009). Confidence intervals in repeated-measures designs: the number of observations principle. *Can. J. Exp. Psychol.* 63 (2), 124–138. doi:10.1037/a0014164

Jones, A., and Pivik, R. (1985). Vestibular activation, smooth pursuit tracking, and psychosis. *Psychiatry Res.* 14 (4), 291–308. doi:10.1016/0165-1781(85)90097-6

Jürgens, R., and Becker, W. (2011). Human spatial orientation in non-stationary environments: relation between self-turning perception and detection of surround motion. *Exp. Brain Res.* 215 (3), 327–344. doi:10.1007/s00221-011-2900-z

Jürgens, R., Kliegl, K., Kassubek, J., and Becker, W. (2016). Optokinetic circular vection: a test of visual–vestibular conflict models of vection nascensy. *Exp. Brain Res.* 234, 67–81. doi:10.1007/s00221-015-4433-3

Karnath, H. O. (1994). Subjective body orientation in neglect and the interactive contribution of neck muscle proprioception and vestibular stimulation. *Brain* 117 (5), 1001–1012. doi:10.1093/brain/117.5.1001

Kassemi, M., Deserranno, D., and Oas, J. (2005). Fluid–structural interactions in the inner ear. *Comput. Struct.* 83 (2-3), 181–189. doi:10.1016/j.compstruc.2004.08.001

Kassemi, M., Deserranno, D., and Oas, J. G. (2004). Effect of gravity on the caloric stimulation of the inner ear. *Ann. N. Y. Acad. Sci.* 1027 (1), 360–370. doi:10.1196/annals.1324.030

Keshavarz, B., Hettinger, L. J., Vena, D., and Campos, J. L. (2014). Combined effects of auditory and visual cues on the perception of vection. *Exp. Brain Res.* 232 (3), 827–836. doi:10.1007/s00221-013-3793-9

Kirollos, R., Allison, R. S., and Palmisano, S. (2017). Cortical correlates of the simulated viewpoint oscillation advantage for vection. *Multisensory Res.* 30 (7-8), 739–761. doi:10.1163/22134808-00002593

Kirollos, R., and Herdman, C. M. (2023). Caloric vestibular stimulation induces vestibular circular vection even with a conflicting visual display presented in a virtual reality headset. *i-Perception* 14 (2), 204166952311680. doi:10.1177/20416695231168093

Kirollos, R., and Herdman, C. M. (2021). Measuring circular vection speed in a virtual reality headset. *Displays* 69, 102049. doi:10.1016/j.displa.2021.102049

Klaus, M. P., Wyssen, G. C., Frank, S. M., Malloni, W. M., Greenlee, M. W., and Mast, F. W. (2020). Vestibular stimulation modulates neural correlates of own-body mental imagery. *J. cognitive Neurosci.* 32 (3), 484–496. doi:10.1162/jocn_a_01496

Lee, D. N., and Lishman, J. (1975). Visual proprioceptive control of stance. *J. Hum. Mov. Stud.* 1 (2), 87–95.

Levy, D. L., Holzman, P. S., and Proctor, L. R. (1983). Vestibular dysfunction and psychopathology. *Schizophr. Bull.* 9 (3), 383–438. doi:10.1093/schbul/9.3.383

Lishman, J., and Lee, D. (1973). The autonomy of visual kinaesthesis. *Perception* 2 (3), 287–294. doi:10.1068/p020287

Mach, E. (1875). *Grundlinien der Lehre von den Bewegungsempfindungen*. Leipzig: Engelmann.

Meiry, J., and Young, L. (1967). Biophysical evaluation of the human vestibular system Semiannual status report. Report.

Melcher, G. A., and Henn, V. (1981). The latency of circular vection during different accelerations of the optokinetic stimulus. *Percept. Psychophys.* 30 (6), 552–556. doi:10.3758/bf03202009

Moon, S., Lee, B., and Na, D. (2006). Therapeutic effects of caloric stimulation and optokinetic stimulation on hemispatial neglect. *J. Clin. Neurology* 2 (1), 12–28. doi:10.3988/jcn.2006.2.1.12

Moore, S. T., Dilda, V., and MacDougall, H. G. (2011). Galvanic vestibular stimulation as an analogue of spatial disorientation after spaceflight. *Aviat. space, Environ. Med.* 82 (5), 535–542. doi:10.3357/asem.2942.2011

Nishiike, S., Nakagawa, S., Nakagawa, A., Uno, A., Tonoike, M., Takeda, N., et al. (2002). Magnetic cortical responses evoked by visual linear forward acceleration. *Neuroreport* 13 (14), 1805–1808. doi:10.1097/00001756-200210070-00023

Owens, D. A., Gu, J., and McNally, R. D. (2018). Perception of the speed of self-motion vs. object-motion: another example of two modes of vision? *Conscious. cognition* 64, 61–71. doi:10.1016/j.concog.2018.07.006

Palmisano, S., Allison, R. S., Schira, M. M., and Barry, R. J. (2015). Future challenges for vection research: definitions, functional significance, measures, and neural bases. *Front. Psychol.* 6, 193. doi:10.3389/fpsyg.2015.00193

Palmisano, S., Gillam, B. J., and Blackburn, S. G. (2000). Global-perspective jitter improves vection in central vision. *Perception* 29 (1), 57–67. doi:10.1068/p2990

Palmisano, S., and Gillam, B. (1998). Stimulus eccentricity and spatial frequency interact to determine circular vection. *Perception* 27 (9), 1067–1077. doi:10.1068/p271067

Purves, D., Augustine, G., Fitzpatrick, D., Katz, L., LaMantia, A., McNamara, J., et al. (2001). The Otolith organs: the utricle and sacculus. *Neuroscience* 4, 777–778. doi:10.1038/90470

Rabbitt, R. D. (2019). Semicircular canal biomechanics in health and disease. *J. neurophysiology* 121 (3), 732–755. doi:10.1152/jn.00708.2018

Reason, J. T. (1978). Motion sickness adaptation: a neural mismatch model. *J. R. Soc. Med.* 71 (11), 819–829. doi:10.1177/014107687807101109

Reymond, G., Droulez, J., and Kemeny, A. (2002). Visuovestibular perception of self-motion modeled as a dynamic optimization process. *Biol. Cybern.* 87 (4), 301–314. doi:10.1007/s00422-002-0357-7

Riecke, B. E., Feuereissen, D., and Rieser, J. J. (2008). "Auditory self-motion illusions (" circular vection") can be facilitated by vibrations and the potential for actual motion," in *Proceedings of the 5th symposium on applied perception in graphics and visualization* (ACM).

Riecke, B. E., and Schulte-Pelkum, J. (2013). Perceptual and cognitive factors for self-motion simulation in virtual environments: how can self-motion illusions ("vection") be utilized? *Hum. Walk. virtual Environ. Percept. Technol. Appl.*, 27–54. doi:10.1007/978-1-4419-8432-6

Rohde, M., van Dam, L. C., and Ernst, M. O. (2016). Statistically optimal multisensory cue integration: a practical tutorial. *Multisensory Res.* 29 (4-5), 279–317. doi:10.1163/22134808-00002510

Santos, C. F., Belinha, J., Gentil, F., Parente, M., and Jorge, R. N. (2017). An alternative 3D numerical method to study the biomechanical behaviour of the human inner ear semicircular canal. *Acta Bioeng. biomechanics* 19 (1), 3–15. doi:10.5277/ABB-00444-2015-02

Sluga, M., Hribar, M., and Battelino, S. (2021). Funkcionalni testi za oceno ravnotežnega organa. *Slovenian Med. J. - Neurobiol.* 90 (1-2), 54–73. doi:10.6016/zdravvestn.2985

St George, R. J., Day, B. L., and Fitzpatrick, R. C. (2011). Adaptation of vestibular signals for self-motion perception. *J. physiology* 589 (4), 843–853. doi:10.1113/jphysiol.2010.197053

Telban, R., and Cardullo, F. (2001). An integrated model of human motion perception with visual-vestibular interaction. AIAA Modeling and Simulation Technologies Conference and Exhibit, Warren, H. (1895). Sensations of rotation. *Psychol. Rev.* 2 (3), 273. doi:10.2514/6.2001-4249

Weech, S., and Troje, N. F. (2017). Vection latency is reduced by bone-conducted vibration and noisy galvanic vestibular stimulation. *Multisensory Res.* 30 (1), 65–90. doi:10.1163/22134808-00002545

Weech, S., Varghese, J. P., and Barnett-Cowan, M. (2018). Estimating the sensorimotor components of cybersickness. *J. neurophysiology* 120 (5), 2201–2217. doi:10.1152/jn.00477.2018

Weech, S., Wall, T., and Barnett-Cowan, M. (2020). Reduction of cybersickness during and immediately following noisy galvanic vestibular stimulation. *Exp. Brain Res.* 238, 427–437. doi:10.1007/s00221-019-05718-5

Wu, X., Yu, S., Shen, S., and Liu, W. (2021). Exploring the biomechanical responses of human cupula by numerical analysis of temperature experiments. *Sci. Rep.* 11 (1), 8208–8210. doi:10.1038/s41598-021-87730-w

Frontiers | Frontiers in Virtual Reality

TYPE Original Research
PUBLISHED 13 February 2024
DOI 10.3389/frvir.2023.1336581

OPEN ACCESS

EDITED BY
Domna Banakou,
New York University Abu Dhabi, United Arab Emirates

REVIEWED BY
Merel M. Jung,
Tilburg University, Netherlands
Marco Fyfe Pietro Gillies,
Goldsmiths University of London,
United Kingdom

*CORRESPONDENCE
John Desnoyers-Stewart,
desnoyer@sfu.ca

RECEIVED 11 November 2023
ACCEPTED 29 December 2023
PUBLISHED 13 February 2024

CITATION
Desnoyers-Stewart J, Bergamo Meneghini M, Stepanova ER and Riecke BE (2024), Real human touch: performer-facilitated touch enhances presence and embodiment in immersive performance.
Front. Virtual Real. 4:1336581.
doi: 10.3389/frvir.2023.1336581

Real human touch: performer-facilitated touch enhances presence and embodiment in immersive performance

John Desnoyers-Stewart[1]*, Margherita Bergamo Meneghini[2,3], Ekaterina R. Stepanova[1] and Bernhard E. Riecke[1]

[1]School of Interactive Arts and Technology, Simon Fraser University, Burnaby, BC, Canada, [2]Doctoral School of Aesthetics, Sciences and Technology of the Arts, Université Paris 8, Paris, France, [3]Compagnie Voix, Paris, France

Despite being an embodied medium, virtual reality (VR) prioritizes vision and sound over the other senses. While touch has been demonstrated to foster a sense of presence and embodiment, most haptic research in VR focuses on uncanny vibration motors or limited experiences of touch with simple props. Meanwhile, immersive performances such as *Eve 3.0* incorporate performer-facilitated touch in novel ways to evoke a complete and social experience of human touch in VR. In response, we conducted a mixed-methods study to investigate the experience of performer-facilitated touch in a 360° video segment from the immersive performance *Eve 3.0*. Using a 3 × 2 factorial design, we compared touch from a diary prop and performer in festival and laboratory settings. We found that performer-facilitated touch increased realistic behaviours and questionnaire measures of social presence, embodiment, and tactile realism. The setting also had a significant effect with festival participants demonstrating significantly more behaviours indicating presence, particularly in the no-touch condition. Participant descriptions reveal that in addition to touch, a rich narrative and vivid visuals of social interaction were just as important in immersing participants in the experience and making them feel present. We find that participant experiences are a co-creation situated at the intersection of artefact and context that require a willing suspension of disbelief. The authentic setting and performance artefact afforded a deep understanding of the rich and complex experience of human touch in immersive performance.

KEYWORDS

virtual reality, touch, performance, behaviour, presence, embodiment, social touch, research in the wild

1 Introduction

While virtual reality (VR) is positioned as a technology that can afford highly embodied experiences that rely on the integration of multiple senses (Gallace et al., 2012), most VR experiences focus solely on vision and sound. Touch is an important sense in the embodiment of virtual bodies (Petkova and Ehrsson, 2008; Slater et al., 2009; Kilteni et al., 2012) and their presence in virtual environments (Slater, 2009), as well as the social presence of other bodies (van Erp and Toet, 2015). While some VR systems do indeed engage the tactile sense, they typically use vibrohaptic motors and fall far short (Gallace and Girondini, 2022) of reproducing the rich experience of touch through which we ordinarily encounter the world (Field, 2014). As such, we need to explore *how rich, authentic experiences of touch in VR may afford a stronger sense of embodiment and presence*.

The incredible potential of synchronously stimulating touch and vision to alter bodily boundaries was first demonstrated by Botvinick and Cohen (1998). In their study, a rubber hand is positioned in front of the subject and their actual hand is hidden. The synchronous touch of the hidden physical hand along with visible touch of the rubber hand leads people to perceive the hand as part of their body. This mediation of the body through multisensory stimuli was demonstrated to work with video (IJsselsteijn et al., 2006) and applied to the body of a mannequin in video-based VR (Petkova and Ehrsson, 2008). Slater et al. (2009, 2010) further extended this, demonstrating that touch could elicit **body ownership** over computer-generated avatars. Body ownership is one of the components of **embodiment** in VR, which also includes self-location and agency (Kilteni et al., 2012). Self-location, or the sense of where oneself is located, and agency, have also been linked to synchronous visuo-tactile stimulus (Kilteni et al., 2012). Agency is distinct from body ownership and is closely related to visuomotor synchrony; however, it can even occur with an immovable rubber hand (Kilteni et al., 2012).

Body ownership and **presence** are closely related, and synchronous visuo-tactile stimuli may by extension elicit a sense of presence in VR (Slater et al., 2009). **Spatial presence**, defined as the "sense of being there" in a virtual environment, has often been the focus of research into the phenomenology of presence (Slater, 2009). However, we should also consider **social presence**, "the sense of being together" (Biocca et al., 2003), and **object presence**, "the subjective experience that a particular object exists in a user's environment" (Stevens et al., 2002). Embodiment and spatial presence are not required for social presence (De Greef and Ijsselsteijn, 2001; Lee, 2004). However, embodiment is likely linked to social presence, particularly in media where users are represented by avatars (Biocca, 1997). While there are many ways to measure spatial presence (Souza et al., 2021) and social presence (Oh et al., 2018), few VR studies consider **object presence**. Yet, the incorporation of physical objects increases the perceived realism of the virtual environment (Hoffman, 1998; Zhang et al., 2022; Felip et al., 2023). For synchronous tactile stimuli to affect the sense of presence in a virtual environment, it must act through specific tactile objects perceived in the virtual environment (Fulkerson, 2013). From this, we can expect synchronous touch to have a positive impact on embodiment, as well as spatial, social, and object presence.

However, the state of the art in VR tends to lack any form of touch beyond controllers. Research on social touch in VR typically employs vibrohaptics rather than direct human contact (Gallace and Spence, 2010; van Erp and Toet, 2015). Touch, especially social touch, is a complex sense deeply rooted in gender, context, and culture (Gallace and Spence, 2010; Fulkerson, 2013; Field, 2014). Even studies that facilitate touch through objects typically use pens, wands, and knives rather than skin-to-skin human contact (e.g., Petkova and Ehrsson, 2008; Slater et al., 2009). Conclusions around the incorporation of touch into VR are thus being made based on a low fidelity expression of non-human touch (e.g., Hoppe et al., 2020; Yarosh et al., 2022). Simply visually displacing controllers in VR outperformed the typical vibrohaptic paradigm (Rietzler et al., 2018), suggesting a requirement for a significant shift in haptic research.

In contrast, there are many innovative solutions arising in immersive dance and theatre. To capture the full richness of touch in VR, immersive performances, such as *The Machine to be Another* (Bertrand et al., 2014), *Eve: dance is an unplaceable place* (Bergamo Meneghini, 2019), *Draw me Closer* (Wilson, 2020), and *Delirious Departures* (Joris and Vandebroeck, 2022), incorporate performer-facilitated touch. These aim to produce a rich experience of tactility that enables a more immersive, embodied experience and stronger sense of physical and social presence. While immersive performance leverages the extensive knowledge of artists to support a compelling embodied experience, these innovative solutions are rarely evaluated systematically. A deeper understanding of the experience of touch in these performances can guide haptic research in VR by grounding it in real, physical touch. Research into experiences of performer-facilitated physical touch can encourage much more authentic perspectives on touch that integrate their rich physical and social dimensions.

One rare example is *The Machine to Be Another*, an embodied VR system that allows participants to swap bodies in VR through synchronous movement and touch (Bertrand et al., 2014). Collaço de Oliveira et al. (2016) found that *The Machine to Be Another* significantly increased the sense of presence and body agency compared to Second Life. However, this increased presence could be attributed to anything from the use of a live camera feed to the incorporation of touch and the lack of a bodily representation in the Second Life condition. Cebolla et al. (2019) evaluated the sense of embodiment as a result of visuo-haptic synchrony in the Machine to be Another. However, the level of embodiment was compared with an arbitrary score of 1 out of 7 as being "non-embodied" rather than to a specific control condition. Neither study investigates what elements might have contributed to the presence or embodiment measured.

Given that much of the innovation in performer-facilitated touch occurs in public performance spaces, there is particular value in exploring them in their authentic context. In the laboratory, the experience may become substantially different (Rogers and Marshall, 2017). While there are numerous investigations of performances *in situ* at festivals, these tend to focus on individual reflexive accounts (e.g., Popat, 2016; Wilson, 2020) or qualitative studies of the complete experience of a limited number of participants (e.g., Rostami et al., 2018; Jégo and Bergamo Meneghini, 2020). However, systematic and quantitative studies are rare. To better understand the real-world effects and generalizability of in-laboratory findings, research is required that investigates performances in an ecologically valid festival context and compares them to a lab setting.

1.1 Our research

To address this research gap and facilitate knowledge transfer between art and science, we investigated how a physical prop and performer-facilitated touch affect the experience of immersive

FIGURE 1
Eve 3.0 as seen during a performance. The performer kneels in sync with the virtual character. Following the music, she touches the immersant's hand in time with the visuals.

performance with a controlled study across laboratory and festival settings. We use a segment from *Eve 3.0*,[1] a multi-user VR performance where a performer facilitates moments of touch, as seen in Figure 1. The performer synchronizes their performance on stage with the choreography executed by characters seen in VR, touching participants at key moments and encouraging them to dance and become performers themselves. Participants also pick up a physical diary prop placed next to them when they see their virtual body reach for the diary. After hearing an excerpt narrated from the diary, it is then ripped from their hands in sync with seeing this action in VR. Later, the diary is thrown towards the participant in the 360° video but not physically. When exhibiting this performance, we often noticed that participants tried to catch the diary at that moment, something that we would not ordinarily expect to see with a 360° video. We hypothesized that added tactility may have invoked a higher level of presence that was expressed through this realistic behaviour (Slater, 2009). This behaviour was particularly interesting as it appeared to be a reflex that the participants did without thinking, often followed by laughter at their realization of what had happened.

1.2 Objectives and research questions

Inspired by this observation, we designed a convergent mixed-methods study (Creswell and Creswell, 2018) comparing three versions of the experience—A (all/full touch): the full performer-facilitated experience, B (book/diary only): using only the physical diary prop, and C (control/no touch): a typical audio-visual 360° video experience without touch. To understand the effects of performer-facilitated touch, we investigate the following research questions:

1 See a trailer of a performance at vimeo.com/834217436; for more details, visit compagnievoix.com/en/projects/creation/eve-3.

1.2.1 Quantitative research questions

How does the incorporation of a physical prop and performer-facilitated touch influence:

- the alignment of participants' behaviour with the VR experience?
- the sense of presence (social, object, and spatial)?
- the sense of embodiment?
- participants' affective state (valence and arousal)?
- the sense of connection to the virtual character?

We analysed behaviours indicative of embodiment and presence, such as moving hands along with the body in VR and attempting to catch the diary. Through validated questionnaires, we measured self-reported affective valence and arousal, embodiment, presence, immersion, and the participants' sense of connection to the main character.

1.2.1.1 Hypotheses

We anticipated that the performer-facilitated full-touch condition (A) would support a stronger sense of embodiment and presence, a stronger emotional response, and a deeper sense of connection to the VR character than the no-touch control condition (C). Similarly, we expected more people to attempt to catch the diary and move along with the visuals seen in the headset in the full-touch condition. As for the diary-only condition (B), we anticipated all variables to fall somewhere between A and C. We also anticipated the festival setting would affect our results compared to the laboratory setting due to differences in demographics, disposition, and context. Our hypotheses were as follows:

- **H1**: Setting (1.1) along with physical touch of a prop alone (1.2) and prop and performer (1.3), each increases the probability that participants will engage in behaviour that reflects presence in the virtual experience, such as trying to catch a diary thrown in the 360° video.
- **H2**: Setting (2.1) along with physical touch from a prop alone (2.2) and prop and performer (2.3), each increases spatial, social, and object presence, immersion, tactile realism, embodiment, social connection, and valence and arousal.

1. Children Playing in the Park.
Seated on a park bench, the participant is approached by a group of teenagers who are playing and dancing together. Désirée slowly approaches the participant and sets down a stack of books beside them. The first touch happens in sync with image 1, where the main character Désirée touches the participant on the shoulder.

2. Picking up the Diary.
The participant hears the instruction ``look at your hands, take my diary'' as they see their body in the 360° video bring their hands up and begin reaching towards a diary next to them. A physical diary is aligned with the video. The participant picks up the diary and holds on to it in their lap.

3. A Story Unfolding through Touch
With the diary in hand, the voice of a young Désirée can be heard speaking lines from the diary, as if the voice was coming from within the diary itself. As the final word is spoken, Désirée notices the participant holding her diary and runs to grab it. The physical diary is synchronously pulled from the participant's hands as Désirée pulls it away.

4. Standing Up and Dancing
Désirée rejoins the other teenagers with her diary. Another character walks up to the participant and helps them stand up from their seat. The performer again synchronizes their touch, pulling the participant a few steps forward before lifting their hands up in the air and beginning to dance with their hands, encouraging them to follow along.

5/6. Playful Moments of Touch
The other characters each approach the participant in turn. First two boys walk up dancing, giving a fist bump on each shoulder. Then, two girls walk up and take the participant by the shoulders, turning and walking with them about a meter.

7. The Diary is Stolen
One of the characters steals the diary from Désirée and begins throwing it to the others, playing 'keep away' as Désirée tries to recover it. Finally, one of the characters throws it to the ground in front of the participant as seen in. This is the moment we noticed many people attempting to catch the thrown diary, although no actual diary is thrown.

8. Helping Désirée
Désirée kneels down to pick up the diary. She looks up to the participant and reaches for help. She opens the diary and dances with it, now sharing her story openly with the participant through the same voice emanating from the diary. In the full performance, an image in the diary becomes the next scene but here the scene fades to black and ends.

FIGURE 2
Narrative structure and moments of touch in the 360° video. Only 2 and 3 are enacted in the diary-only condition, while none are in the no-touch condition. No actual diary is ever thrown in 7.

1.2.2 Qualitative research questions

We also asked open-ended questions to gain a deeper qualitative understanding of the rich complexity of participant experiences. Our qualitative research questions were

- What aspects of the experience do participants describe as contributing to their experience of touch, embodiment, and presence in *Eve 3.0*?
- Why do participants try to catch the diary, or not, when it is thrown towards them in *Eve 3.0*?

1.2.3 Mixed research question

The qualitative findings are used to explain the quantitative results, connecting them to the behavioural and questionnaire measures to develop a holistic understanding and triangulate the data:

- How do participants' descriptions of their experience align and explain the behavioural and questionnaire outcomes?

2 Materials and methods

2.1 Artefact: *Eve 3.0* performance 360° video segment "Dear Diary"

The full version of *Eve 3.0* is performed in front of an audience where six participants are brought on stage and each sees a different story in VR. Each story follows a character as they detail their struggle with anxiety, addiction, depression, paranoia, obsession, and jealousy. Participants are given a VR headset and asked to align their bodies with the one seen in a stereoscopic 360° video. Throughout the 360° video, the performer makes contact with the participants in sync with the visuals, encouraging them to become immersed and to move and dance. The full performance continues with a motion-captured dance, with each dance reflecting movements inspired by each character's struggles, and ends with an open, improvised dance encouraged through lively pass-through visuals.

Rather than being designed specifically for this study, we take an unmodified segment of a complete artwork, created "for art's sake." The performance has been minimally adjusted to stay true to the authentic artefact while enabling quantitative analysis. We focused on the 360° video as it offered an opportunity to investigate the behaviours regularly seen in public performances. It also provided an easily repeatable segment that could be run under different conditions simultaneously. The video segment is 9 min long. Moments of touch are synchronized with a subtle audio cue in the soundtrack allowing the researchers to consistently facilitate the moments of contact.

We specifically used the 360° video featuring Désirée, who struggles with addiction, as this story had the most consistently relatable narrative and best audio recording quality. The narrative is presented in English for this study.

2.1.1 Narrative and key moments of touch in "Dear Diary"

In the 360° video, participants embody the character Eve. This performance includes not only the physical moments of touch but also invites participants to participate in a rich narrative, as seen in Figure 2.

2.1.2 Hardware

Due to hardware availability, different headsets were used. Three Oculus Go headsets were used in the festival setting with Showtime VR[2] to synchronize the videos. Two Meta Quest 2s and one Meta Quest 1 were used in the laboratory setting, synchronized using custom *Eve 3.0* software. The Quest 1 was alternated between conditions each session to avoid affecting the results for any one condition. A Meta Quest 2 was used for all convenience samples. All headsets were used to view the same three degrees of freedom, stereoscopic, 360° video, minimizing any differences in the hardware. While the resolutions and refresh rates differ between the headsets, the limiting factor was the video quality (4096 × 4096 @ 60 fps) which, at a 90° field of view (FOV), produces a 1024 × 1024 stereoscopic image.

2.2 Participants

This study was approved by the Simon Fraser University Research Ethics Board (#30001617) and the Université Paris 8 Ethics Committee (#CE-P8-2023-06-3). Participants gave their informed consent and were allowed to withdraw from the study until their data were de-identified. Participants could ask questions about the research and performance after completing the study and sign up to receive the study results.

2.2.1 Recruitment

By conducting our research in both a laboratory and festival setting, we aimed to improve ecological validity, recruit a more diverse sample, and investigate whether laboratory studies are representative of a real-world application. 108 participants were recruited through three methods: 1) through a performance at the 2023 Cinedans festival in Amsterdam[3] (festival, $n = 50$); 2) through Simon Fraser University's School of Interactive Arts Research Participation System in Surrey, BC, Canada (lab, $n = 47$); and 3) through convenience sampling (convenience, $n = 11$). Convenience sampled participants were recruited in a wide variety of locations during visits to labs and universities to ensure a more representative sample (Calgary, AB, Canada, $n = 4$; Portland, OR, United States, $n = 1$; Vienna, Austria, $n = 2$; Santiago de Compostela, Spain, $n = 3$; Vancouver, BC, Canada, $n = 1$). Laboratory and convenience participants were combined for analysis due to similarities in context and the small number of convenience participants.

2.2.2 Exclusions

Seven additional participants were excluded from the study. Three participants in the festival setting did not complete the optional survey and provide consent. Three convenience participants were excluded due to asynchronous physical touch. One laboratory participant was excluded because they removed the headset during critical moments of the experience.

2.2.3 Demographics

Seventy-two (66.7%) participants were identified as women, 35 (32.4%) as men, and one (0.926%) as gender-fluid. Furthermore, 70.7% of laboratory participants identified as women compared to 62.0% of festival participants. The participants were between 18 and 74 years of age and had a wide range of VR experience, from none ($n = 16$) to some with over 100 times using VR ($n = 5$) and 360° video ($n = 2$). Participants had a wide range of movement practice experience, from 0 to 40 years (mean = 6.72 years). Their current

2 showtimevr.eu/

3 cinedans.nl/p/dear-diary-23

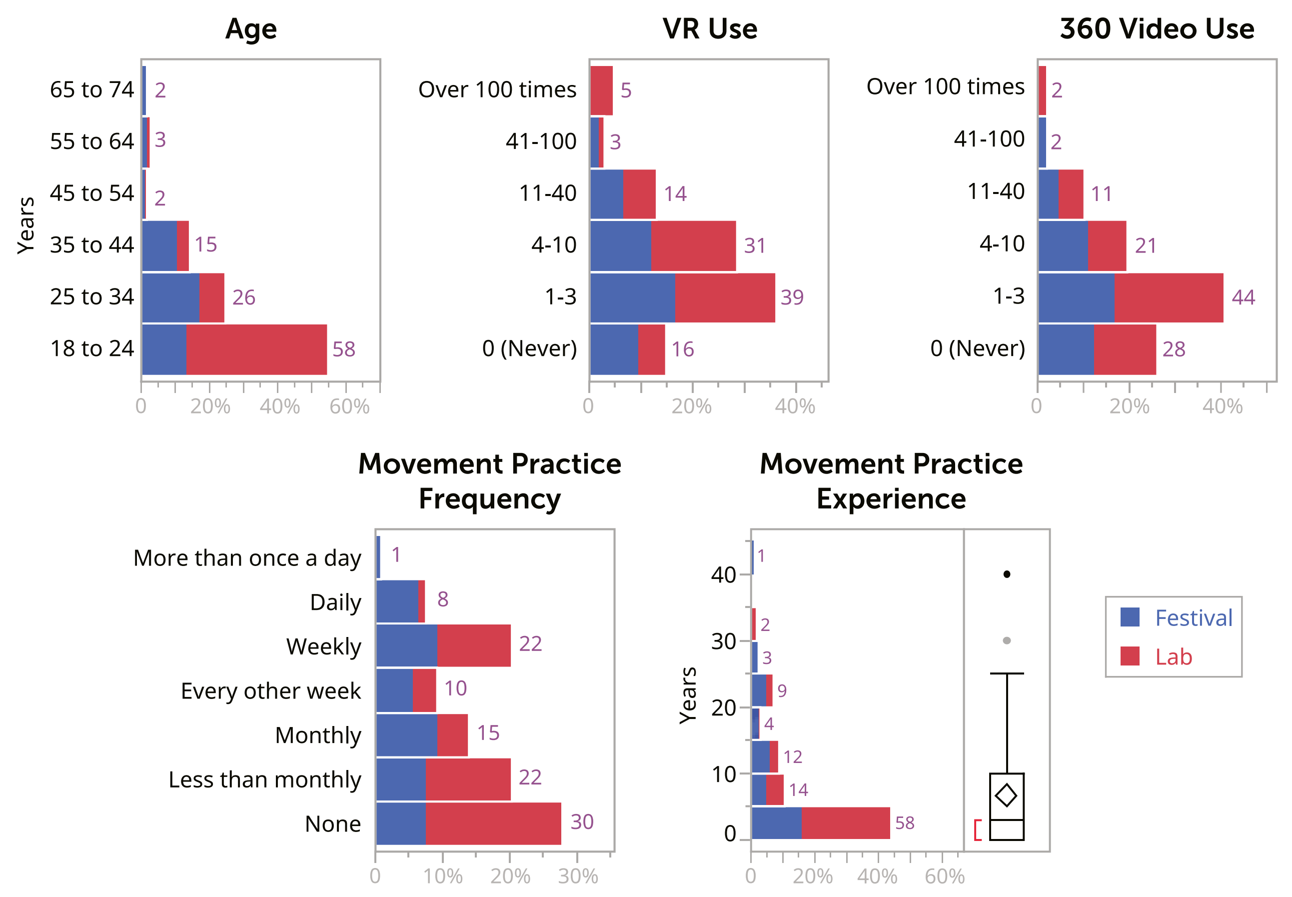


FIGURE 3
Histograms of participant demographics.

movement practice frequency varied from none (n = 30) to daily (n = 9). Movement practice was defined in the questionnaire to include informal and professional practices such as dance, tai chi, capoeira, or yoga but left open to interpretation by participants; for more details, see Figure 3.

2.3 Experimental design and procedure

The study follows a 2 × 3 factorial between-subjects design. The independent variables are different levels of physical touch (A: full touch, B: diary only, and C: no touch), and setting (festival and laboratory). The overall structure is shown in Figure 4A. Each session lasted 30 min that included setup, the 9-min 360° video, and 15 min of questionnaires.

2.3.1 Settings

Participants in the festival setting attended Cinedans in Amsterdam alongside other performances, films, and VR experiences. The experience was included in the festival guide and online, and participants could walk up to participate in the experience. They could try the experience without participating in the study. The context can be seen in Figure 4B. In the laboratory setting, participants signed up in advance and came to a black box studio laboratory at a scheduled time. The laboratory setting can be seen in Figure 4C. Convenience participants primarily engaged with the experience in university lab similar to the one seen here.

2.3.2 Condition assignment and pre-experience questionnaire

Upon entering the lab or visiting the festival booth, participants were assigned a condition and participant number. We sought to stratify participants across conditions by assigning them sequentially to the next condition pseudo-randomly based on different alternating factors such as positioning from left to right or right to left to reduce possible bias in selecting the condition. When fewer than three participants were present, we alternated conditions to maintain a balanced number of participants in each condition. Participants were given their participant number and asked to complete a digital consent form and demographic survey on their smartphone or a researcher-provided laptop.

2.3.3 Conditions

Upon completing the demographic survey, participants were asked to sit in one of three chairs set up in a line to allow the researcher to facilitate multiple experiences simultaneously. The same chair position was used for each condition to ensure consistent video analysis, as shown in Figure 4. Similar participant numbers were run through each condition in each setting and overall: full touch (A) = 35 (festival = 18, lab = 17); diary only (B) = 38 (festival =

FIGURE 4
Study design and settings showing the layout where participants were seated. The conditions were routinely seated in the same location from left to right: full touch, diary-only, and no touch. **(A)** Overview of study design showing when each measure was collected. **(B)** Festival setting at Cinedans 2023 in Amsterdam, Netherlands. The facilitator helps the full touch participant to stand up. Festival-goers can be seen lounging at a cafe in the background. **(C)** Lab setting at Simon Fraser University in Surrey, Canada. Participants are lined up at the beginning of the study in the black box laboratory. The diaries have not been placed yet.

17, lab = 21); no touch (C) = 35 (festival = 15, lab 20). Discrepancies in participant numbers primarily arose due to exclusions.

Each participant was asked to sit in a chair with a low table or chair to their right. Once all participants were comfortably seated and ready with their headset on, the facilitator started the experience. The facilitator asked participants to line themselves up with what they saw to ensure their body was aligned. The exact diary seen in the 360° video was then carefully placed aligned with its visual location next to participants in the diary-only (B) and full-touch (A) conditions. The control condition (C) was provided with no prop, of no touch, and simply watched the 360° video. In the diary-only condition (B), only a physical diary was used. In the full-touch condition (A), a diary was provided, and the facilitator touched the participants in sync with the characters in the 360° video to produce the complete experience as designed.

Participants viewed the 360° video segment simultaneously allowing for consistent facilitation video analysis using the music to identify the moment in the experience. The facilitator performed six moments of touch shown in Figure 2. In the diary-only condition, the diary was placed next to the participant and then removed from their hands at the right moment during the experience. When a participant did not reach over to pick up the diary or struggled to find it, the facilitator brought it closer or placed it in their hands.

2.3.4 Video recording and behavioural measures

The entire experience was video-recorded for analysis of any significant behaviours. Facilitators also recorded in a notebook whether participants attempted to catch the diary and any interesting behaviours to include in the analysis. Such behaviours

TABLE 1 Questions from the post-experience questionnaire.

Measure	Construct	Question
Affect grid (Russell et al., 1989)	Valence and Arousal	Please rate how you are feeling right now by clicking on the grid.
Temple presence inventory (Lombard et al., 2009)	Spatial Presence	How much did it seem as if the objects and people you saw had come to the place you were?
		How much did it seem as if you could reach out and touch the objects or people you saw?
		How often when an object or person seemed to be headed towards you did you want to move to get out of its way?
		To what extent did you experience a sense of being there inside the environment you saw?
		How often did you want to or try to touch something you saw?
	Social Presence: Parasocial	How often did you have the sensation that people you saw could also see you?
		How much did it seem as if you and the people you saw were together in the same place?
	Social Presence: Active	How often did you smile in response to someone you saw in immersive environment?
	Engagement: Immersion	To what extent did you feel mentally immersed in the experience?
	Perceptual Realism: Touch	Overall, how much did touching the things and people in the environment you saw feel like it would if you had experienced them directly?
Peck and Gonzalez-Franco (2021)	Embodiment	When I was holding the diary (as seen in the above image), I felt as if the body I saw holding it was my body
Schubert et al. (2001)	Object Presence	When the diary was being thrown around in front of me (as in the above image), I felt like the diary I saw was physically there in the immersive experience.
IOS (Aron et al., 1992)	Social Connection	Please click on the picture below that best describes your relationship with the main character in the immersive experience (pictured above—the girl who held the diary up to you at the end).
Quality	Touch Quality	If you did not feel any physical touch, please select N/A. If you felt any touch during the experience, please rate how well synchronized those moments were for you.
	Technical Quality	Please rate the technical quality of your experience where Very Good means no technical issues and Very Poor means you experienced many issues such as skipping or blurry images, sound issues, and tracking problems.
	Technical Issues	If you encountered any technical issues, please briefly list them below.
Open-ended qualitative	Perceived Catch	Did you try to catch the diary when it was thrown to you? Why or why not?
	Significant Moments	What interesting or significant moment(s) stood out to you? Please describe each moment. (What happened? What did you think/see/feel/hear/do in that moment? What made it significant for you?)
	Diary Moment	Thinking back to the moment when you saw the diary thrown to you, what was that experience like? What did you think/see/feel/hear/do in that moment?

that align with what is experienced in VR are indicative of presence (Slater, 2009) and object presence (Reiner and Hecht, 2009).

2.3.5 Post-experience questionnaire

After completing the VR experience, participants' headsets were removed. Participants were asked not to remark on the experience or ask questions until they completed the survey. The survey was conducted in a public festival setting, so emphasis was placed on keeping the survey concise for reliable completion. With a 93.9% completion rate, we struck a good balance. The full list of questions is shown in Table 1 and included in the Supplementary Material.

A continuous Affect Grid (Russell et al., 1989) was used to measure valence and arousal upon completing the experience. While there are many different questionnaires for measuring presence (Souza et al., 2021), we chose the Temple Presence Inventory (TPI) for its adaptability and inclusion of social presence and immersion (Lombard et al., 2009). Select items from the TPI (Lombard, 2013) were used to measure spatial and social presence (active social and parasocial dimensions), as well as immersion and perceptual realism. Items were selected to fit the research questions of the study and ensure they were sensible, given the constraints of the 360° video. For example, we did not ask *"How often did it feel as if someone you saw/heard in the environment was talking directly to you?"* because there is only internal dialogue. We excluded the social richness and social realism dimensions because the narrative was relatively dreamlike and the experience was not interactive.

We used a single-item embodiment question based on Peck and Gonzalez-Franco (2021) and a single-item object presence measure based on Schubert et al. (2001). Similar to Stevens et al. (2002), we rephrased the question to refer to the object rather than environment presence. For each question we asked participants to refer to a specific instance for consistency. Participants then identified their connection to the main character using the Inclusion of Other in Self (IOS) scale (Aron et al., 1992).

FIGURE 5
Moments in the 360° video that we coded for participant behaviours in addition to Figure 2. **(A)** The children lead the participant in a dance, while their representation in the 360° video follows along. **Followed Hands** was coded for participants following along with the movements this section. **(B)** The children fight over a diary right in front of the participant. Many participants were seen **Backing Away** at this moment, especially if they were standing. **(C)** The children pass the diary to each other. We noticed almost every single participant **Tracking the Diary as it was Thrown Around** with their head in this moment, suggesting it was the centre of their attention.

Finally, participants filled out questions about technical and touch quality before answering four open-ended questions. Namely, the open-ended questions inquired about participants' memory and reasoning for trying to catch the diary; the most significant moment; and their experience of the diary being thrown towards them. All Likert-scale questions are on a 7-point rating scale, except the quality questions which were out of 5.

2.4 Analysis

JMP Statistical Discovery Software[4] was used to analyse quantitative data. Behaviour coding was combined with questionnaire results by the participant number in Microsoft Excel and imported into JMP. Qualitative analysis was conducted using NVivo.[5]

2.4.1 Behaviour coding and analysis

Each researcher recorded perceived attempts to catch the diary at the end of each session. Additional interesting behaviours were noted through observation and selected for coding before the quantitative analysis began. In total, we observed whether participants

- reached for the diary (Figure 2.2),
- stood up (Figure 2.4),
- followed hand movements seen in VR (Figure 5A),
- backed away at a moment when the children came close while fighting over the diary (Figure 5B),
- tracked the diary as it was thrown around (Figure 5C),
- tried to catch the diary when it was thrown to them (Figure 2.7), and
- reached to help the girl as she stood up (Figure 2).

These behaviours were identified to indicate that participants were deeply engaged or feeling a strong sense of presence as suggested by Slater (2009). Behavioural measures have been previously used to measure object presence (Reiner and Hecht, 2009) and demonstrate the social presence of avatars (Bailenson et al., 2003). The video data were independently coded by the first two authors and then reviewed to ensure consensus for the final coding. Participants were also asked whether they tried to catch the diary. We reviewed any video data where there was a mismatch between our code and participant perceptions.

To analyse the relationship between these behaviours and the different settings and conditions, we ran a full-factorial logistic regression for each behaviour. In cases where a significant main effect was found, we conducted a *post hoc* χ^2 (chi-squared) test between each level. We used likelihood ratio χ^2 for main effects and interactions, while *post hoc* tests are Wald-based odds ratio χ^2 tests.

2.4.2 Quantitative questionnaire responses

Two-way ANOVA was used to analyse all Likert-scale questionnaire responses. According to Harpe (2015), Likert-scale items can be analysed as continuous data provided that conditions are met, such as having at least five numerical response categories. This also aligns with how the Temple Presence Inventory (Lombard et al., 2009), IOS (Aron et al., 1992), and Affect Grid (Russell et al., 1989) are typically analysed to facilitate meta-analysis. ANOVA is also highly robust to violations of assumptions that include skewness and non-normality (Carifio and Perla, 2008). Tukey's honest significance test was used for *post hoc* tests.

4 www.jmp.com/

5 lumivero.com/products/nvivo/

We also conducted correlation analyses between specific behaviours and questionnaire results. We used Pearson correlations for relationships between questionnaire answers. Pearson correlations are robust against violations of normality and can be used with Likert-data (Carifio and Perla, 2008). We used point-biserial correlation to identify relationships between dichotomous behaviour coding and questionnaire measures (Field, 2013).

2.4.3 Qualitative questionnaire responses

The open-ended survey questions were analysed through template analysis (King, 2012). The first three authors reviewed all qualitative data and made notes on prominent themes and possible codes for the template. Then, the researchers discussed their observations to derive a draft template. This template was informed by the research questions, knowledge of relevant phenomena (i.e., embodiment, presence, etc.), and observations from the data collection and familiarization stages. Then, the second and third authors divided the components of the template and individually coded the data, revising the template as necessary. The first author reviewed the final coding and contributed to the overarching themes. Finally, the first author connected quotes to relevant quantitative data to understand how participants' interpretations connected to our measurements.

2.4.3.1 Researcher reflexivity

Qualitative approaches acknowledge the key role that the researcher's background and expertise play in analysis (Creswell and Poth, 2018). To support the transparency of our qualitative analysis, we present brief statements from each researcher who engaged with the qualitative data outlining the lens through which we engage with participants' descriptions.

John Desnoyers-Stewart is an artist/researcher with a background in engineering who has exhibited over 40 interactive immersive performances and installations since 2018. His practice-based research investigates underexplored possibilities of VR to support embodiment, social connection, and self-expression. For *Eve 3.0*, he contributed to the experience design, developed the software, and created the real-time graphics.

Margherita Bergamo Meneghini is a dancer and choreographer who led the creation of *Eve 3.0*, building upon the successes of *Eve: dance is an unplaceable place (2019)*. Her research is oriented towards audience participation, applying elements from dance movement therapy in combination with immersive technologies.

Ekaterina R. Stepanova is a cognitive scientist and a VR researcher, with over 8 years of research experience. Informed by embodied cognition, phenomenology, and social psychology, her work explores how VR can foster a felt experience of connection.

3 Results

Our results are presented below. In line with Cumming (2014) and APA recommendations, we report 95% confidence intervals for all quantitative measurements [in square brackets]. Marginal significance levels are also analysed to reduce dichotomous reliance on the *p*-value in evaluating results and focus on the story told by the measurements and effect size (Cumming, 2011). We begin with behavioural measures, questionnaire measures, and correlation results, followed finally by qualitative themes. Raw questionnaire data and coded video data are included in Supplementary Material. Three participants' data have been removed from the data set as they did not consent to share their data.

3.1 Summary of quantitative findings

Given the complexity of the data presented below (Figures 7, 9, 10), we present an overview of the significant main effects found in this study in Figure 6. Condition had a significant effect on all behaviours except reaching for the diary, while the setting had an effect on all but backing away. Age, movement experience, and frequency were all significantly higher in the festival setting. Condition significantly affected parasocial presence, tactile realism, embodiment, and touch quality, while setting significantly affected social connection (IOS), valence, arousal, and technical quality.

3.2 Behavioural measures

As shown in Table 2, the logistic regressions demonstrated significant main effects for both condition and setting across almost all behaviours. The only exceptions were reaching for the diary, which only showed a marginally significant effect for condition, and backing away, which was not significantly affected by the setting. Detailed descriptive statistics and *post hoc* test results for each section can be seen in Figure 7. Unless noted otherwise, logistic regression *p*-values are based on likelihood ratio χ^2 tests while *post hoc* tests are Wald-based odds ratio χ^2 tests with one degree of freedom.

3.2.1 Reaching for the diary

We compared only two conditions for reaching for the diary because up to this moment, the no-touch and diary-only conditions were identical. No touch and diary only are combined into "No Touch Before Diary" and compared to "Touch Before Diary" (full touch). A single touch on the participant's shoulder led to a marginally significant increase in participants reaching to pick up the diary. As seen in Figure 7, this had a marginally significant effect on festival participants (χ^2 = 3.50, p = 0.061) but not on lab participants. Overall, festival participants were significantly more likely to reach for the diary (66.0% [52.2%, 77.6%]) than laboratory participants (38.6% [27.1%, 51.6%]).

3.2.2 Standing up

We did not conduct inferential tests for standing up because participants in the full-touch condition were physically compelled to stand up. However, as shown in Figure 7, significantly more festival participants (n = 7, 46.7%) stood up unprompted in the no-touch condition compared to only a single (5%) laboratory participant (χ^2 = 5.98, p = 0.014).

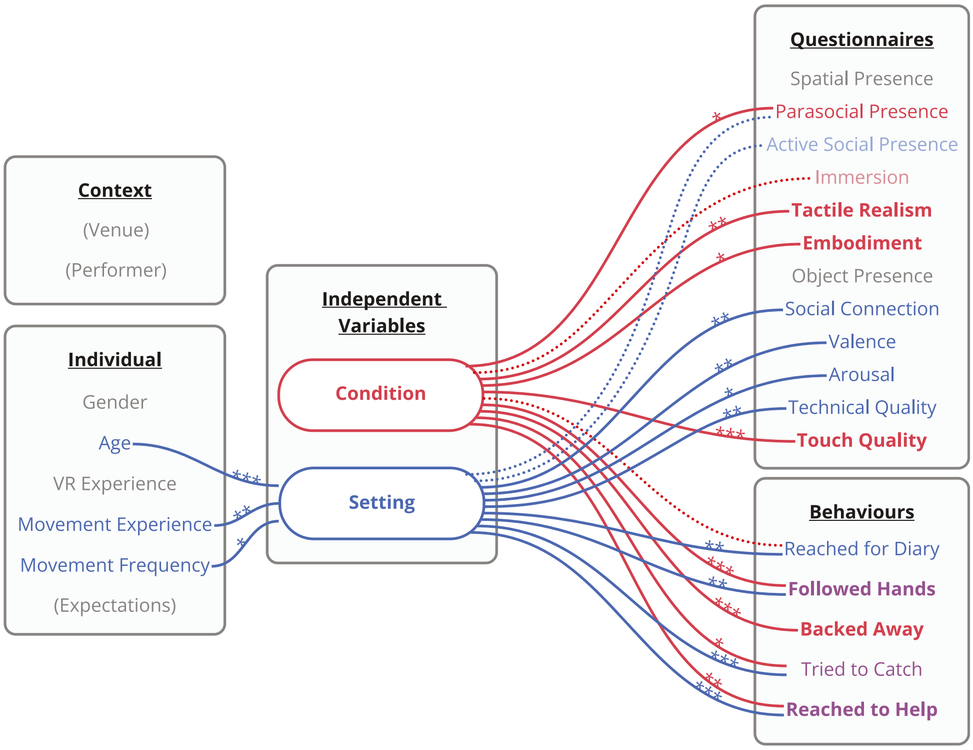


FIGURE 6
Overview of main effects (dashed line: $0.05 < p < 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$). Bold text indicates a medium effect size or greater. Colour indicates which independent variable had a significant effect. Elements in brackets are theorized but not measured.

TABLE 2 Table of results from logistic regressions on key observed behaviours.

Behavioural measures	Condition			Setting			Condition × setting		
	$\chi^2(1)$	p	w	$\chi^2(1)$	p	w	$\chi^2(1)$	p	w
Reached for diary	2.74	<u>0.0981</u>	<u>0.159</u>	9.43	**0.0021**	<u>0.296</u>	1.74	0.1874	0.127
	$\chi^2(2)$	p	w	$\chi^2(1)$	p	w	$\chi^2(2)$	p	w
Followed hands	37.88	**<0.0001**	**0.592**	10.18	**0.0014**	**0.307**	3.36	0.1861	0.176
Backed away	14.31	**0.0008**	**0.364**	0.72	0.3954	0.082	0.43	0.8845	0.063
Tried to catch	8.39	**0.0151**	<u>0.279</u>	13.75	**0.0002**	**0.357**	6.54	**0.0380**	<u>0.246</u>
Reached to help	9.67	**0.0079**	<u>0.299</u>	13.75	**0.0002**	**0.357**	5.54	<u>0.0626</u>	<u>0.227</u>
	$F(2, 106)$	p	η_p^2	$F(1, 106)$	p	η_p^2	$F(2, 106)$	p	η_p^2
Behaviour count	13.85	**<0.0001**	**0.176**	15.58	**0.0001**	**0.099**	1.78	0.1733	0.023

Bold text indicates a significant result ($p < 0.05$) or a medium effect size ($w > 0.3$ or $\eta_p^2 > 0.06$) or greater. Underlined text indicates marginal significance ($0.05 < p < 0.10$) or a small effect size ($w > 0.1$ or $\eta_p^2 > 0.01$).

3.2.3 Following hands

In all, 100% of full touch participants were seen following the hands seen in the 360° video (Figure 5A). While the facilitator pushed full touch participants' hands up as they let go in line with the video, we only coded instances where participants continued to move their hands or did so unprompted. There was a large ($w > 0.5$) effect for condition and medium ($w > 0.3$) effect for setting (Cohen, 1988). While laboratory participants steadily increased from one condition to the next (Figure 7), festival participants were significantly more likely to follow the hands in the no-touch condition compared to laboratory participants ($\chi^2 = 8.52$, $p = 0.0035$).

3.2.4 Backing up

In all, 63.6% ($n = 35$) of participants who were standing ($n = 55$) stepped back from the children fighting over the diary (Figure 5B), regardless of the condition. We also included seated participants who

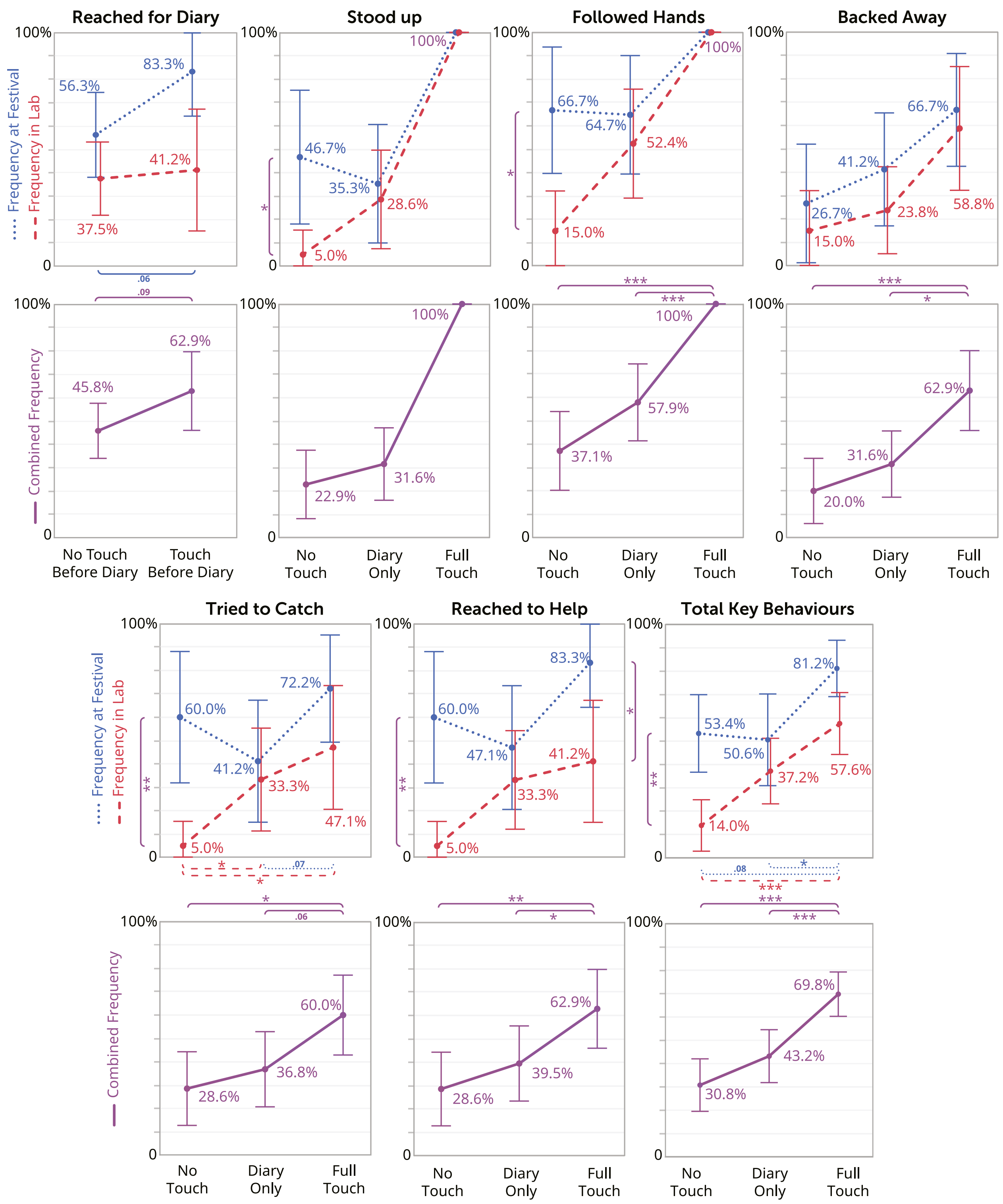


FIGURE 7
Frequency of observed behaviours. Error bars show the 95% confidence interval. Bars show significant pairwise comparisons (*p < 0.05, **p < 0.01, ***p < 0.001). The top row is separated by setting, while the bottom is the total.

visibly shifted their weight back in their chair (n = 6). As shown in Figure 7, there was a medium effect for condition with full touch participants being 3.619 [1.35, 9.62] times more likely to back away than diary only participants (χ^2 = 9.11, p = 0.0025) and 6.67 [2.26, 19.7] times more than no touch participants (χ^2 = 10.39, p = 0.0013). More festival participants (46.0% [33.0%, 59.6%]) exhibited this behaviour than laboratory participants (31.0% [20.6%, 43.8%]); however, we found no significant difference (χ^2 = 0.71, p = 0.398).

3.2.5 Tracking the diary as it is thrown around

A total of 105 (97.2%) participants were observed tracking the diary with their heads as it was thrown around in the video (Figure 5C). The three participants who did not follow the diary were in the no-touch condition in the laboratory. We did not conduct a logistic regression given the prevalence of this behaviour.

3.2.6 Trying to catch the diary

We found a significant but small ($w > 0.1$) effect for condition and condition × setting on trying to catch the diary (Table 2). We also found a significant medium effect of setting ($w > 0.3$). As shown in Figure 7, festival participants (58.0% [44.2%, 70.6%]) were much more likely to try catching the diary than laboratory participants (27.6% [17.8%, 40.2%]). This was particularly notable in the no-touch condition where participants were 28.5 [2.97, 273] times more likely to try catching the diary in the festival setting than in the laboratory (χ^2 = 8.43, p = 0.0037). Overall, participants in the full-touch condition were 5.41 [1.43, 20.5] times more likely to try catching the diary than in the no-touch setting (χ^2 = 6.18, p = 0.013). In the laboratory setting, participants in the diary-only (χ^2 = 4.00, p = 0.046) and full-touch (χ^2 = 6.20, p = 0.013) conditions were significantly more likely to try catching the diary than those in the no-touch condition. Meanwhile, in the festival setting, we found a marginally significant increase from diary-only condition to full touch (χ^2 = 3.31, p = 0.069).

An additional 17 participants thought they tried to catch the diary despite not physically responding (16.5% of total), while two tried to catch the diary but thought they had not (1.94%), However, most participants' perceptions aligned with the observed behaviour (n = 84, 81.6%).

3.2.7 Helping the girl stand up

Both condition and setting significantly affected the proportion of participants who reached to help the girl stand up (Table 2). Participants in the full-touch condition were significantly more likely to reach and help than in the diary-only (χ^2 = 3.90, p = 0.048) or no-touch (χ^2 = 7.29, p = 0.007) conditions (Figure 7). Participants in the festival setting (64.0% [50.1%, 75.9%]) were much more likely to reach out and help than laboratory participants (25.9% [16.3%, 38.4%]) (χ^2 = 8.43, p = 0.0037). While no interaction was found, Figure 7 clearly shows that festival participants were significantly more likely to reach out to help than laboratory participants in both the no-touch condition (χ^2 = 8.43, 0.0037) and the full-touch condition (χ^2 = 6.01, p = 0.0142).

3.2.8 Total key behaviours

We added the total number of key behaviours exhibited by each participant to investigate the overall difference between conditions. This included reaching for the diary, following hands, backing away, trying to catch the diary, and reaching to help (n = 5). We left out standing up since it was physically compelled in the full-touch condition and tracking the diary since nearly every participant engaged in this behaviour.

A two-way ANOVA revealed significant main effects for condition and setting with no interaction (Table 2). There was a strong effect for condition ($\eta_p^2 > 0.14$) and medium effect for setting ($\eta_p^2 > 0.06$) (Cohen, 1988; Miles and Shevlin, 2001). Participants in the full touch condition engaged in significantly more key behaviours than both the diary only and no touch conditions as shown in Figure 7. Festival participants engaged in significantly more key behaviours (*mean* = M = 3.12 [2.65, 3.59]) than laboratory participants (M = 1.76 [1.33, 2.19]). Tukey's honest significance test found the total number of behaviours observed in the no-touch condition was significantly higher at the festival than in the laboratory (p < 0.0001).

3.2.9 Behavioural hypotheses findings

These results support H1.1 that setting affects the probability of all realistic behaviours except backing away. H1.2 was not supported as we found only one significant result (trying to catch in the lab) for diary only compared to no touch. The data does however support H1.3 that full touch that includes the diary and performer-facilitated touch increases the probability of all realistic behaviours compared to no touch and using the diary only. Two key patterns are visible in nearly all of the graphs in Figure 7:

1. In the laboratory setting, there is a steady increase in realistic behaviour with additional touch starting from almost none in the no-touch condition.
2. In the festival setting, there is an elevated baseline of realistic behaviour for the no-touch condition. The diary-only condition typically performed worst.

We explore possible explanations in the Discussion section.

3.3 Questionnaire measures

As shown in Table 3, two-way ANOVAs showed significant main effects for condition on parasocial presence, tactile realism, embodiment, and touch quality, while setting significantly affected social connection (IOS), valence, arousal, and technical quality. We found no significant results for condition or setting on object presence. Detailed descriptive statistics and *post hoc* test results for each section can be seen in Figure 9. Unless noted otherwise, *post hoc* tests are Tukey's honest significance tests.

3.3.1 Demographics

We compared demographics between settings to identify any significant differences. Gender, VR use, and 360° video use were not significantly different. A contingency analysis showed that festival participants (M = 33) were significantly older than laboratory participants (M = 25) [$\chi^2(5)$ = 24.2, p < 0.0002]. Festival participants (median = monthly) had a more frequent movement practice than lab participants (median < monthly) [$\chi^2(6)$ = 16.6, p < 0.05]. Festival participants also had a significantly longer history of movement practice (M = 9.1 years) than lab participants (M = 4.7 years) [$t(101)$ = 2.72, p < 0.01]. These practices included dance, yoga, and similar activities although some participants may have included various sports. The mean age difference was 8 years, while the difference in movement practices was 4.4 years. Thus, the difference in movement experience is partly attributable to age and partly to more substantial movement practices among festival participants. This aligns with expectations since attendees of a dance festival would likely have a stronger interest in movement practice generally.

TABLE 3 Results from 2-way ANOVAs on questionnaire results.

Questionnaire measures	Condition			Setting			Condition × setting		
	$F(2, 102)$	p	η_p^2	$F(1, 102)$	p	η_p^2	$F(2, 102)$	p	η_p^2
TPI									
Spatial presence	1.93	0.149	0.036	2.42	0.123	0.022	1.25	0.291	0.023
Parasocial	3.28	**0.042**	0.058	2.96	0.088	0.026	0.72	0.487	0.013
Active social	1.21	0.301	0.027	3.11	0.081	0.027	0.59	0.532	0.009
Immersion	2.47	0.090	0.045	1.09	0.298	0.010	0.55	0.578	0.010
Tactile realism	5.02	**0.008**	**0.088**	0.12	0.728	0.001	0.92	0.402	0.016
Embodiment	3.67	**0.029**	**0.061**	1.34	0.249	0.011	3.60	**0.031**	**0.060**
Object presence	1.97	0.145	0.037	0.00	0.988	0.000	0.47	0.628	0.009
IOS	0.08	0.923	0.002	4.48	**0.037**	0.042	2.56	0.082	0.048
Valence	2.02	0.138	0.035	8.30	**0.005**	**0.072**	1.04	0.356	0.018
Arousal	0.04	0.958	0.001	5.10	**0.026**	0.048	2.38	0.098	0.045
Tech quality	1.22	0.299	0.022	7.53	**0.007**	**0.068**	3.82	**0.031**	0.003
	$F(1, 70)$	p	η_p^2	$F(1, 106)$	p	η_p^2	$F(1, 70)$	p	η_p^2
Touch quality	16.67	**0.0001**	**0.197**	0.16	0.694	0.002	0.20	0.658	0.002

Bold text indicates a significant result ($p < 0.05$) or a medium effect size ($\eta_p^2 > 0.06$) or greater. Underlined text indicates marginal significance ($0.05 < p < 0.10$) or a small effect size ($\eta_p^2 > 0.01$).

3.3.2 Temple Presence Inventory

As shown in Table 3, there was a significant main effect of condition on the parasocial presence (actor within medium) and tactile perceptual realism dimensions of the TPI. *Post hoc* tests revealed a marginally significant increase parasocial presence for the full-touch condition compared to diary only (p = 0.090) and no touch (p = 0.059). There was a medium effect on tactile realism which was significantly higher for the full-touch condition than was for both the diary-only (p = 0.019) and no-touch (p = 0.020) conditions. Interestingly, the pattern seen in the behavioural results was reversed for perceptual realism. At the festival, perceptual realism increased steadily with the increase in touch, while in the laboratory, the diary-only condition is rated the lowest; for details, see Figure 8. Condition had a marginally significant effect on the immersion (engagement) dimension but *post hoc* tests were non-significant.

Setting had a marginally significant effect on the parasocial and active social dimensions. The mean parasocial score was higher at the festival (5.72 [5.36, 6.08]) than the laboratory (5.37 [5.05, 5.69]). The mean active social score was also higher at the festival (4.94 [4.48, 5.40]) than the laboratory (4.03 [3.59, 4.48]). All other TPI results were non-significant.

3.3.3 Embodiment

We found a medium main effect for condition and a medium interaction effect of condition × setting on embodiment. Overall, the full-touch condition was significantly higher than the no-touch condition (p = 0.022). In the laboratory, embodiment was significantly higher in the full-touch condition than in the diary-only (p = 0.038) and no-touch (p = 0.042) conditions. The diary-only condition also showed significantly higher embodiment at the festival than the lab (p = 0.027). This result is particularly interesting, given the inverted relationship with behavioural measures in the festival setting; for details, see Figure 9.

3.3.4 Social connection

The Inclusion of Other in Self (IOS) scale showed a significant but small main effect of setting (Table 3). Festival participants' IOS score towards the main character in the 360° video was higher on average at 3.56 [3.02, 4.10] than at 3.29 [2.83, 3.755] for laboratory participants (p = 0.0367). There was also a marginally significant interaction of setting × condition; however, *post hoc* tests did not reveal any significant differences between specific conditions and settings. This marginally significant interaction may be explained by the contrast between the steady increase in IOS scores across conditions in the laboratory and the steady decrease in the festival setting, as seen in Figure 9.

3.3.5 Affect

There was a medium effect of condition on valence, with festival participants reporting a more positive valence (M = 0.749 [0.698, 0.800]) than laboratory participants (M = 0.611 [0.542, 0.681]). We also found a marginally significant difference between the festival and laboratory settings in the no-touch condition (p = 0.0534). Laboratory participants reported a significant but subtly higher level of arousal (M = 0.570 [0.514, 0.626]) than festival participants (M = 0.531 [0.472, 0.590]). A marginal interaction effect was detected for arousal, but we found no results from *post hoc* tests. However, as shown

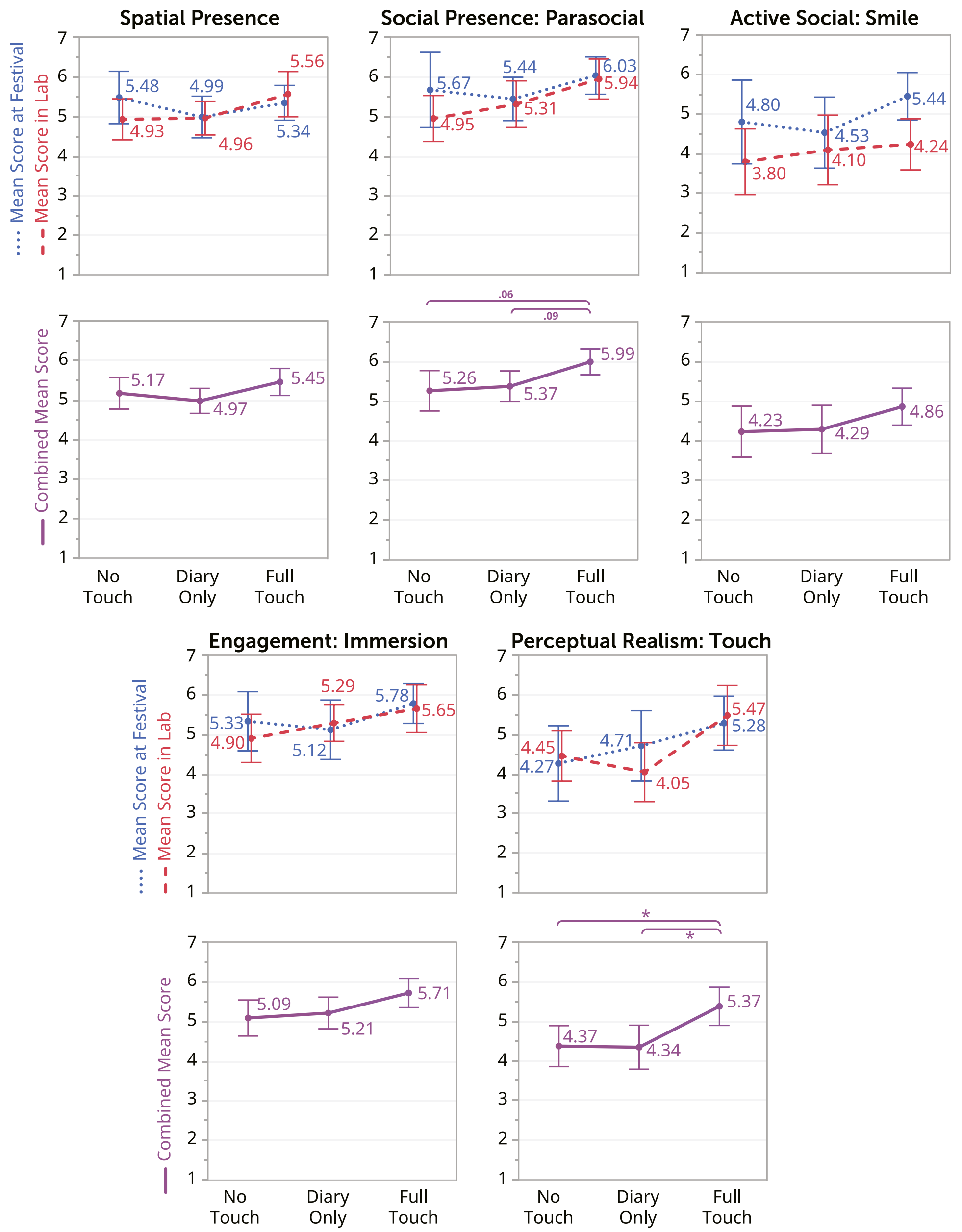


FIGURE 8
Score of Likert scale questions from the Temple Presence Inventory. Error bars show the 95% confidence interval. Bars show significant pairwise comparisons (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). p-values specified for marginal results. The top row is separated by setting, while the bottom is the total.

in Figure 9, most of the difference between settings can be attributed to the no-touch setting where festival participants were generally more relaxed. Taken together, festival participants reported a more relaxed and positive affect in the no-touch condition, aligning closely with laboratory participants in other conditions.

3.3.6 Quality of experience

We found a medium effect of setting on the perceived technical quality and a significant but almost undetectable interaction effect between condition and setting. Festival participants rated the technical quality significantly higher (M = 4.10 [3.91, 4.29]) than laboratory participants (M = 3.96 [3.75, 4.18]). This small difference can be entirely attributed to the no-touch condition which was higher at the festival than in the laboratory, as seen in Figure 10 with marginal significance (p = 0.075).

For touch quality, we removed the no-touch condition since judgements of touch quality would be baseless. We found a strong effect of condition on touch quality with full touch scoring significantly higher than the diary-only condition (p = 0.0001), as shown in Figure 10. We found no significant difference between the lab setting (M = 3.82 [3.47, 4.16]) and festival (M = 3.92 [3.57, 4.27]). Touch quality may have been worse in the diary-only condition for two reasons:

1. There was only one moment of touch for participants to base their judgement of tactile quality on.

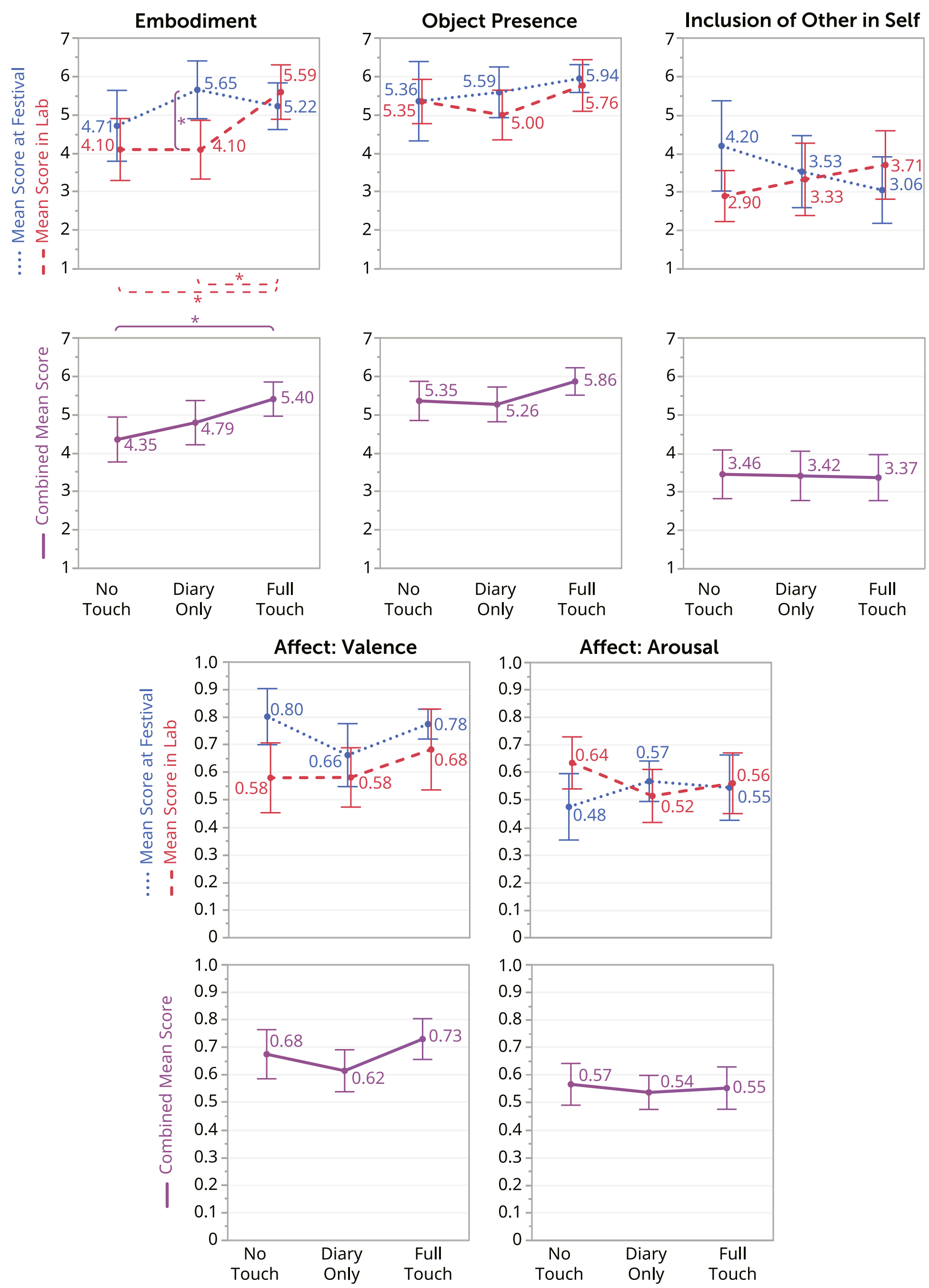


FIGURE 9
Score of Likert scale questions and Affect Grid from the questionnaire. Error bars show the 95% confidence interval. Bars show significant pairwise comparisons (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). The top row is separated by setting, while the bottom is the total.

2. Participants perceived other moments of visual touch as missing tactility, and thus a lower quality.

3.3.7 Questionnaire hypothesis findings

The results partially support H2.1 and H2.3 that setting and the full-touch condition increase the scores of some questionnaire-based indicators of presence, while others were inconclusive. The full-touch condition incorporating a prop and performer increased participants' sense of parasocial presence (actor within medium), tactile realism, and embodiment. The festival setting significantly increased participants' sense of social connection (IOS), valence, and arousal. The data does not support H2.2 as we found no measurable difference between the no-touch and diary-only conditions.

3.4 Correlation results

To better understand how participants' questionnaire results correlated with their behaviour, we ran a multivariate correlation

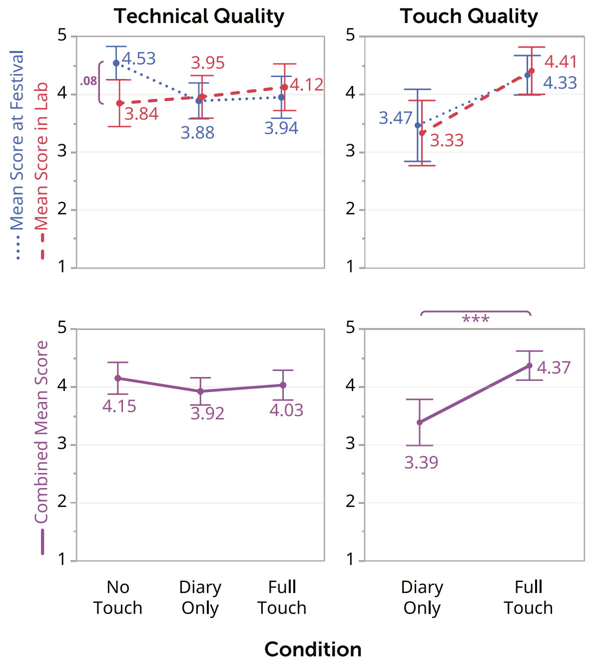


FIGURE 10
Score of touch and technical quality scores from the questionnaire. Error bars show the 95% confidence interval. Bars show significant pairwise comparisons (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). The top row is separated by setting, while the bottom is the total.

analysis between select behavioural and questionnaire measures. All notable correlations were positive unless noted otherwise.

3.4.1 Prior experience

We investigated correlations of reaching for the diary with participants' prior VR and 360° use as this was the first element to diverge from ordinary VR and 360° video. As such, their behaviour up to this point should correlate with their prior experience. We ran a contingency analysis that showed that reaching for the diary increased from 36.5% of participants who had never used 360° video to 66.7% of participants who had used it more than three times [$\chi^2(5) = 12.3$, $p = 0.031$]. However, there was no significant correlation between overall VR use and reaching for the diary [$\chi^2(5) = 6.26$, $p = 0.282$]. This suggests that participants with more 360° video, but not VR experience, were more likely to reach for the diary.

3.4.2 Behaviour correlations with questionnaire results

Point-biserial regression was used to identify correlations between dichotomous behaviour variables and questionnaire results. The correlation results are summarized in Table 4. p-values can be misleading when analysing correlations (Cumming, 2011), and effect size is a significantly better indicator (Field, 2013). Therefore, we present only the effect size with confidence intervals and look for at least a weak ($r > 0.1$) correlation whose confidence interval does not cross zero.

Backing away was weakly correlated with spatial presence but showed no other clear correlations. Trying to catch the diary was moderately ($r > 0.3$) correlated with object presence. It also showed weak correlations with the TPI active dimension of social presence and participants' valence. Reaching to help the girl stand up showed the strongest correlation results. We found moderate correlations between reaching to help and both object presence and valence. There were weak correlations between reaching to help and all dimensions of the TPI except tactile realism. Embodiment was also weakly correlated with reaching to help. While the 95% confidence interval for IOS just barely crosses zero, there may be a weak correlation present here as well. In addition to the results shown in Table 4, following along with the hands in the 360° video (Followed Hands) was weakly correlated with embodiment ($r(106) = 0.212[0.023, 0.387]$). The total behaviour count was also moderately correlated with object presence ($r(106) = 0.296$ $[0.112, 0.460]$) and valence ($0.323[0.141, 0.484]$).

3.4.3 Questionnaire result correlations

Pearson's r was used to identify correlations between questionnaire results, as shown in Table 5. We did not assess correlations between the TPI dimensions, as this was already addressed in the development of the questionnaire (Lombard et al., 2009). Embodiment was strongly correlated ($r > 0.5$) with tactile realism and moderately correlated with the spatial, parasocial presence, and immersion dimensions of the TPI but not with active social presence. Object presence was strongly correlated with spatial presence, weakly correlated with active social presence, and moderately correlated with the other TPI dimensions. Embodiment and object presence were also moderately correlated. The IOS results were moderately correlated with embodiment, object presence, and both the spatial and tactile dimensions of the TPI. IOS also showed weak correlations with the parasocial and immersion dimensions but not with active social. We also found a correlation between the TPI active social dimension, based on whether participants smiled, and valence ($r(106) =$ 0.292 $[0.107, 0.457]$). Although this correlation was weak, it was the strongest correlation of any variable with valence.

3.5 Qualitative themes

Here, we present the key themes identified in participants' qualitative responses (Section 2.4.3). We include exemplary quotes with additional quotes in the Supplementary Material. Relevant quantitative measures provide context to illustrate how qualitative themes relate to quantitative results.

Participants described their experience in VR, reactions to feeling touch, and other significant moments. By considering participants' reflections, we can understand what felt important for them and interpret which elements of the VR experience may have elicited, supported, or inhibited aspects of their phenomenological experience. The themes identified tell a nuanced story about each participant's unique experience that contributed to their sense of presence in VR. These themes are interwoven with the notions of embodiment, touch, emotions, and the social and narrative components of the experience. We outline the themes through three intertwined facets of the complex experience of presence: physical touch supports embodiment, object, and spatial presence (Section 3.5.1); eye contact, proximity, and engagement support social presence (Section

TABLE 4 Correlations between behaviours and questionnaire results.

Behavioural correlations	Backed away		Tried to catch		Reached to help	
	r(106)[a]	(95% *CI*)	*r*(106)[a]	(95% *CI*)	*r*(106)[a]	(95% *CI*)
TPI						
Spatial	0.227	(0.040, 0.399)	0.138	(−0.053, 0.318)	0.260	(0.074, 0.428)
Parasocial	0.087	(−0.104, 0.272)	0.106	(−0.084, 0.290)	0.247	(0.061, 0.417)
Active social	0.027	(−0.163, 0.215)	0.216	(0.028, 0.389)	0.205	(0.017, 0.379)
Immersion	0.146	(−0.044, 0.326)	0.154	(−0.036, 0.334)	0.235	(0.049, 0.406)
Tactile realism	0.047	(−0.143, 0.234)	0.061	(−0.129, 0.247)	0.163	(−0.027, 0.341)
Embodiment	0.144	(−0.047, 0.325)	0.067	(−0.124, 0.254)	0.222	(0.033, 0.395)
Object presence	0.119	(−0.073, 0.302)	**0.306**	(0.123, 0.469)	**0.305**	(0.122, 0.468)
IOS	−0.043	(−0.232, 0.149)	0.078	(−0.115, 0.265)	0.164	(−0.028, 0.344)
Valence	0.106	(−0.087, 0.291)	0.257	(0.069, 0.426)	**0.341**	(0.160, 0.499)

Bold text indicates a moderate correlation (> 0.3). Underlined text indicates at least a weak correlation (> 0.1), where the CI does not cross 0.
[a]Denotes average degrees of freedom. Degrees of freedom varied between 103 and 107 depending on the questionnaire answer as not all respondents answered all questions.

TABLE 5 Correlations between questionnaire results.

Questionnaire correlations	Embodiment		Object presence		IOS	
	r(106)[a]	(95% *CI*)	*r*(106)[a]	(95% *CI*)	*r*(106)[a]	(95% *CI*)
TPI						
Spatial	**0.458**	(0.293, 0.596)	**0.554**	(0.407, 0.673)	**0.356**	(0.177, 0.512)
Parasocial	**0.405**	(0.233, 0.553)	**0.479**	(0.318, 0.613)	0.248	(0.060, 0.419)
Active social	0.161	(−0.030, 0.340)	0.285	(0.101, 0.451)	0.164	(−0.027, 0.344)
Immersion	**0.359**	(0.182, 0.514)	**0.446**	(0.280, 0.586)	0.266	(0.079, 0.434)
Tactile realism	**0.521**	(0.368, 0.647)	**0.486**	(0.327, 0.619)	**0.305**	(0.121, 0.468)
Embodiment	—	—	**0.380**	(0.202, 0.530)	**0.315**	(0.131, 0.477)
Object presence	**0.378**	(0.202, 0.53)	—	—	**0.304**	(0.120, 0.469)

Bold text indicates a moderate correlation (> 0.3). Underlined text indicates at least a weak correlation (> 0.1), where the CI does not cross 0.
[a]Denotes average degrees of freedom. Degrees of freedom varied between 103 and 107 depending on the questionnaire answer as not all respondents answered all questions.

3.5.2); and narrative immerses participants in the story (Section 3.5.3). We then consider participants' descriptions of the diary, which fuses all three facets into a rich experience.

We refer to participants by code numbers. The first two letters indicate the location:

- CD: Cinedans Festival, Amsterdam, Netherlands
- SF: Simon Fraser University, Surrey, Canada
- CA: University of Calgary, Canada
- SC: University of Santiago de Compostela, Spain
- PL: Portland, OR, United States.

The two-digit code indicates the session number which included up to three participants. The last letter indicates the condition: A = all/full touch, B = book/diary only, and C = control/no touch. Relevant quantitative measures are indicated by the following codes:

- SP: Spatial Presence (1–7)
- SoP: Social Presence (Parasocial) (1–7)
- SoA: Social Presence (Active) (1–7)
- EI: Engagement/Immersion (1–7)
- TR: Tactile Realism (1–7)
- Em: Embodiment (1–7)
- OP: Object Presence (1–7)
- IOS: Social Connection (1–7)
- Ar: Arousal (0–1.0)
- Va: Valence (0–1.0)
- TeQ: Technical Quality (1–5)
- ToQ: Touch Quality (1–5).

For example, a participant in the third session at Cinedans in the diary-only condition with a tactile realism score of 6 would be indicated as (CD03B—TR: 6).

3.5.1 Physical touch supports embodiment, and spatial and object presence

Participants described how physical touch contributed to spatial presence, embodiment, and object presence. Picking up the physical diary supported participants' sense of object presence which increased their spatial presence. "*I reached out to touch the books that were placed next to me, they were actually there which made me think that I'm fully immersed in the VR environment*" (SF07B—TR: 6, OP: 7, SP: 6). Interestingly, several participants described this moment of presence as **surreal**: "*The most surreal part of the experience was when I was physically interacting with the objects that were in the environment which made me really feel like I was there*" (SF10A—TR: 7, OP: 7, SP: 6.6). The merging of the physical and virtual worlds through touch felt not only surprisingly realistic but also uncannily impossible and therefore surreal.

Physical presence often had a rich social and emotional component. For instance, CA03A described how the touch they felt elicited a complex emotional response: "*The two boys touching me on the shoulder. It was a strange moment, it seemed equal parts threatening and kind*" (SoA: 7, SoP: 7, EI: 7). Physical touch stimulated a further desire to interact with the environment physically. Participants described a desire to touch, to be touched, to interact with the diary, and to take part in the action seen in VR: "*. . .each time the people in the environment approached me . . .I felt a not insignificant desire to physically respond and interact in kind*" (PL01B).

Inconsistent touch took some participants out of the experience. Some participants in the diary-only condition developed an expectation for physical touch that went unfulfilled, impeding their sense of presence: "*Later the girl held my hands but in reality I didn't feel anyone grabbing my hands which really reduced my immersion*" (SF07B—ToQ: 2). However, SF07B still rated their spatial and social presence highly, stating that "*the visuals were extremely realistic*" (SP: 6, SoP: 7, EI: 5).

While the diary supported presence, when immersants reached for it and did not "catch" it, it sometimes made them more aware of the mediation of their experience: "*My first response was to reach out and grab it, but after realizing it was not real, I laughed*" (SF01B—EI: 7). These are examples of **bifurcation**, a simultaneous awareness of two parallel dimensions (Morie, 2007).

3.5.2 Eye contact, proximity, and engagement support social presence

A crucial component of presence is the social presence of other people. In the 360° video, characters acknowledge the presence of the participant and invite them to join in. This created some of the most compelling moments as a majority of the participants reported a strong reaction to the characters making eye contact or getting close to them: "*. . .each time a person looked directly at me or approached me. That elicited the strongest feelings for me. It was eerie how much it felt like they were looking directly at me*" (CA04C—SoP: 5, SoA: 5). **Eye contact** and **proximity** were the most commonly mentioned significant moments of social presence.

3.5.2.1 Eye contact acknowledges immersant's presence

Eye contact produced a spectrum of reactions, with many participants recognizing the pivotal role it played in social presence: "*The eye contact was powerful throughout the experience and contributed a lot to the sense that I was part of the experience, rather than an observer*" (SC01B—SoP: 6.5, SoA: 6). This led participants to feel the social presence of others and being noticed by them: "*I felt like the people are actually there and are noticing me*" (SF07B). Proximity and touch seemed to compound this experience. Eye contact sometimes felt "*intimate*" (SC01B) and even "*uncomfortable*" (SF06A) and could elicit a sense of "*awkwardness*" (SF06B), "*stress*" (CA01B), or "*anxiety*" (SF08B). While eye contact sometimes felt intense and uncomfortable, this intensified the experience, making participants more engaged and immersed: "*The direct eye contact and abstract movements made me feel uneasy but even more immersed*" (SF01B).

3.5.2.2 Proxemics pushed people away but made social presence palpable

Another contributing factor to social presence was the proximity of the characters to participants. Being directly approached stimulated a strong sense of being together with the characters: "*I thought that I was watching a film . . . but after the character approached me I realized I was part of it too*" (SF12C—SoP: 6). A particular moment when all of the characters get close to the participant often led to an embodied response from participants trying to create more distance, as seen in Section 3.2. "*Having people in the video get close to my space. I instinctively would react to move out of the way*" (SF19B—SoP: 5.5). Similarly to the direct eye contact, moments of close interpersonal proximity were sometimes associated with fear: "*I think that it was interesting to see how other people started getting closer to me . . . It felt like a threat mentally so my hands were a bit shaking*" (SF13A—SoP: 5.5). Participants sometimes tried to remove themselves from the experience to overcome this discomfort: "*. . . I had to close my eyes as consciously I knew it's just video but I wanted to move away from them . . .*" (SF11C—SoP: 2.5, SoA: 2, EI: 5).

The discomfort of intense eye contact was mitigated by positive facial expressions: "*When the two boys came up close and danced, I still wanted to take a step back, but their attitude was so jovial that it really made me smile*" (CA03A—SoP: 7, SoA: 6). Characters **smiling** pulled participants into the experience with many reporting a desire to smile back: "*The moments where characters smiled or clearly addressed me with an expression or gesture made me feel strongly included in the action*" (CD12A—SoP: 7, SoA: 5, EI: 6). Strong emotional reactions also supported the feeling of presence: "*The first person who made eye contact changed my feeling of observing the scene to being really present. It's quite intense*" (SC01A—SP: 4.4, SoP: 5.5, SoA: 5). This also demonstrates the dynamic nature of the experience, changing from one moment to the next. Participants felt that the sustained eye contact was very realistic, while sometimes too much: "*the main character was looking right at my eyes . . . It gave me a bit of anxiety because it felt too realistic*" (SF08B—SoP: 6.5, SoA: 5).

3.5.2.3 Included through an invitation to participate

When the characters invited participants to join, it deepened participants' sense of presence, engagement, and embodiment: "*When the two girls helped me get up and come dance with them. It felt like I was actually there and included. It changed from watching to (inter)acting*" (CD06C—SP: 6.6, Em: 6). However, the sense of social presence was most strongly supported by participants' own desire to participate: "*I wanted to move and dance with the people in the VR, I wanted to smile back and respond*" (CA01B—SoP: 6.5, SoA: 6).

Eye contact and proximity elicited diverging feelings of either **intrusiveness or inclusiveness**. Intrusiveness happened when the VR characters approached directly and did not engage the participant, who remained an observer: "*Whenever the characters looked directly at me I felt as if I was intruding their space. Over time that changed and they started to include me* . . ." (SF05A). Inclusiveness occurred when participants were invited to participate, finding their place in the narrative (e.g., CD06C). Characters smiling created a more positive disposition: "*When they came close with the two of them, smiling, that felt nice and included*" (CD07B—SoP: 5.5, SoA: 5, Va: 0.69).

When eye contact and proximity led to a feeling of inclusiveness, participants reported a sense of social connection: "*Dancing with the characters. I felt connected to them and felt like they can actually see me*" (SF17A—SoP: 5.5, IOS: 1). This supported immersion into the virtual world: "*It felt like I was being led into another world, one that those kids are seeing and feeling*" (SF15B—SP: 6.4, SoP: 5, EI: 5, IOS: 6).

3.5.3 Narrative immerses participants through role play and a willing suspension of agency

Participants willingly suspended their disbelief and sought their **role** in the narrative: "*The set-up makes me try to play a role and kind of obey the role*" (CD13B—EI: 6). Participants often pondered about the appropriateness of their interpretation and actions: "*It made me wonder what my role in this narrative was supposed to be*" (CA03A—EI: 7). They often followed what was expected from their role, taking the cues from the virtual body provided to them: "*I think I was trying to match my apparent hands*" (PL02A—EI: 4, Em: 2). This willing engagement immersed participants without necessarily evoking a strong feeling of embodiment: "*The way I had to follow my own arms was a strong experience, I was trying to imitate myself*" (CD03C—EI: 6, Em: 3). By moving along with what they saw in VR, participants became more involved and immersed in the story: "*Arm waves made me feel that I'm in same dimension*" (CD09B—SoP: 5.5, EI: 5, Em: 5).

3.5.3.1 A willing suspension of agency: allowing the virtual body to lead you

Since the experience was a 360° video and could not respond to participants' actions, it might be assumed that participants feel little to no agency. However, many participants' descriptions indicate the contrary, due to their willing **suspension of agency** akin to the concept of willingly suspending disbelief (Coleridge, 1817). Participants engaged with the narrative by suspending their agency and submitting to what their virtual body was doing, gaining an embodied involvement in the story: "*it felt more like the VR body was controlling my actual body than the other way around*" (CD06C—EI: 6, Em: 6). Participants often felt surprised by how the experience pulled them in: "*I was intrigued by how I* . . . *was trying to move my body simultaneously with the virtual body*" (CD04A—EI: 4, Em: 2). However, participants were sometimes frustrated by the lack of agency in the 360° video: ". . .*the experience was wanting me to move when I wanted to stay still and forcing me to stay still when I wanted to move* . . ." (CA05B—EI: 5, Em: 5).

3.5.3.2 Identities shared and split

Participants felt like they were a character playing a role in the story, performing the actions they expected of their character. Some participants felt like they *were* a different virtual character: "*I suppose I didn't really associate myself in the experience with my own identity. I got thinking about whether the children were seeing someone male or female, young or old*" (CA04C—EI: 6). Other participants retained their own identity in the virtual space. This discrepancy sometimes inhibited participants' experiences as they felt out of place in the story: "*When two girls came very close to me it felt awkward, because of the age difference* . . . *Like I wasn't the right character for the role, maybe?*" (CD02A—SoP: 5, EI: 5). In addition to age, male participants CA05B and PL01B were particularly uncomfortable because of the gender interaction as they felt they should not be interacting with a young girl in such a seemingly intimate encounter in a park.

3.5.4 The physical diary fuses touch, narrative, and social significance

The diary was central to the experience. It combined the physical, social, and narrative elements, all intertwined through a single prop.

The diary's virtual and physical existence caused participants to reflect on what was real or not: "*The most surreal part of the experience was when I was physically interacting with the objects that were in the environment which made me really feel like I was there*" (SF10A—SP: 6.6, TR: 7, Em: 7, OP: 7). Participants stated that during their experience in VR, it felt "*real*" (SF01B), but in retrospect, they knew it could not be: "*[I'm] not sure to what extent I am trying to justify trying to catch the book after the fact. It did make me laugh that I tried to catch it*" (CA06A—OP: 6).

3.5.5 Subjective explanations for trying to catch the diary

Participants' reactions to the diary being thrown towards them exemplify how the physical, social, and narrative components of the experience come together to produce an observable physical response. Reflecting on their experience, participants provided a range of explanations for why they think they tried to catch the diary ranging from a simple reflex to reasons based on the meaning embedded in the diary.

3.5.5.1 Because it is an instinctive response

Many participants stated that they attempted to catch the diary reflexively: "*It was automatic. I will always try to catch something if it is thrown to me!*" (SC01B—OP: 7). Surprised by this reflex, participants realized how immersed they were: "*Instinctively I felt I had to catch the book, that made me realise how much I felt [immersed] into the environment*" (CD13A—EI: 6, OP: 7).

3.5.5.2 Because it was expected

Others anticipated the diary to be thrown to them narratively: "*I thought they're going to throw me the book so I should be ready to catch it*" (CD16B—TR: 7). The physical interaction with the diary in the full-touch condition set the expectation that a physical diary might be thrown: ". . .*since I had a physical diary to pick up at the beginning, I thought that* . . . *I would get to touch it physically*"

(SF10A—TR: 7, OP: 7). For some participants, their inability to catch the diary imbued with emotions: "*I felt slightly disappointed when I didn't catch the book as if I truly missed it*" (SF09A—OP: 7).

3.5.5.3 Because I was immersed

Some participants attributed their actions to feeling immersed in the experience: "*[The] very immersive experience led to my natural bodily reflexes to act immediately*" (SF14A—EI: 6, OP: 6). Many explain that it "*felt real.*" Social presence was also a factor: "*Because I felt like . . . I was [one] of them, so I tried catching the diary . . .*" (CD18A—SoP: 6.5, SoA: 7, OP: 6). Conversely, participants who did not try to catch the diary felt like spectators watching a film. "*[I] didn't feel like it was my own body but watching as in the same perspective*" (SF16B—Em: 3, OP: 4).

3.5.5.4 Because the diary was significant

The need to catch the diary was often rooted in participants' relationship to it. Some felt that the diary was theirs, while others felt they should not touch it since it belonged to the girl. For instance, SF17A describes a sense of personal connection to the diary: "*I had the urge to grab the diary like it was mine*" (OP: 7). Conversely, CD12A felt that it was inappropriate for them to grab someone else's diary: "*I wanted to pick it up–but . . . it wasn't for me to touch*" (SoP: 7, OP: 7). Participants felt that the narrative importance of the diary made them try to catch it: "*I tried to catch the diary because . . . it was something important*" (SF15B—OP: 6). The story filled the diary with significance: "*Through the narration . . . we understand the weight of what is inside the book*" (CD12A).

3.5.5.5 Because I felt empathy for the girl

Some participants felt a sense of emotional connection to the main character and wanted to help her: "*I felt a lot of love toward her and wanted to find a way to help*" (CA04C—IOS: 3). They expressed a physical urge to help: "*I felt an overwhelming sense to intervene/ empathy for the main character*" (CD12A—IOS: 2), or at least to comfort her: "*I felt bad for the main girl because everyone was throwing her stuff around. I wanted to give her a hug*" (CD16C—IOS: 6). Participants who felt empathy did not necessarily feel connected to the girl sometimes, as indicated by low IOS results. Those who did feel connected mentioned that this feeling was fostered by the intimate thoughts shared from the diary: "*It felt like the main character was telling everything to me that was in the diary and her thoughts, that made me feel super connected with her*" (SF08A—SoP: 7, IOS: 6).

4 Discussion

This study systematically explored how performer-facilitated touch in controlled laboratory and ecological festival settings affected participants' presence in VR. We found that both setting and touch impacted participants' experiences of presence and embodiment in a 360° video performance, *Eve 3.0*, as demonstrated through behaviour and self-report. Here, we interpret the quantitative and qualitative results and how they relate, discussing the underlying factors contributing to our findings. We discuss the values and challenges of conducting research in the wild and embracing the complexity and diversity of experiences that arise. We conclude with research directions towards a deeper understanding of physical human touch in immersive experiences.

4.1 Festival setting and performer-facilitated touch increase realistic behaviour, presence, and embodiment

As hypothesized, performer-facilitated touch increased the probability of realistic behaviour in response to a 360° video. As Slater (2009) claims, this realistic behaviour reflects a stronger sense of presence in VR, suggesting that touch indeed enhances presence, aligning with previous research (e.g., IJsselsteijn et al., 2006; Petkova and Ehrsson, 2008; Slater et al., 2009; for a review, see Souza et al., 2021). Previous research showed that touching a physical prop increased presence and realism in VR (Hoffman, 1998; Zhang et al., 2022; Felip et al., 2023). Unlike these studies, we did not find a significant difference between the no-touch control and diary-only conditions. However, we did see a significant increase from diary only to full touch in all behaviours. As hypothesized, we observed a steady increase between conditions in the laboratory setting, suggesting that touching a single prop likely lies somewhere between experiences with no touch and those incorporating multiple moments of human touch.

Interestingly, participants were more likely to engage in all of the observed behaviours in the festival setting, particularly in the no-touch condition. While the laboratory results suggest an incremental benefit with increased touch, the festival results suggest that consistency is more important. The physical diary might set up an expectation for future touch that is then unfulfilled, reducing presence. This difference may relate to different expectations. Festival participants came to the dance festival to enjoy it. Thus, they were likely eager to engage and follow along. Festival participants in the full-touch condition were also more likely to reach for the diary after a single shoulder tap, suggesting that they may have been more attentive to touch. On the other hand, laboratory participants, mostly university students pursuing a technical degree, may have anticipated an innovative technological solution. Touching the diary therefore incrementally enhanced their experience over 360° video.

4.1.1 Touch leads to complex experiences that support realistic behaviours resulting in complex experiences

In line with Souza et al. (2021), we found behavioural measures were more reliable than questionnaires. However, as Souza et al. identified, it can be challenging to understand behaviour without additional measures. We found that performer-facilitated touch increased participants' sense of social presence, tactile realism, and embodiment. These results make clear sense as the performer's touch should feel realistic and mostly facilitate social touch between different characters and participants' bodies. The lack of results around spatial presence and immersion is reasonable as their high scores likely represented a ceiling effect from the realistic visuals and engaging narrative. We were surprised by the lack of results for object presence, given the clear effects on tactile realism and embodiment. However, this may have resulted from the diary's narrative significance, leading its presence to be high across all conditions and reinforcing the social dimension of presence. This combination of questionnaire results along with the increased behaviours at the festival indicates that the experiences evoked by touch were complex.

To unpack the complexity of participants' experiences, we conducted a correlation analysis. The moderate correlation between object presence and trying to catch the diary indicates that this behaviour coincided with a high level of object presence, regardless of what led to that level of object presence. This suggests that the reflex to catch the diary might be a good indicator of object presence. While backing away was weakly correlated with spatial presence, this was more likely an indicator of participants restoring a comfortable level of proximity to the characters in line with the equilibrium theory (Argyle and Dean, 1965; Bailenson et al., 2001). While backing away may indicate social presence, some participants were instead excited to be included in the circle.

Reaching to help correlated weakly with most measures of presence, suggesting that this behaviour might indicate overall presence. While this gesture might indicate a connection to the person being helped, there was no clear correlation for IOS, suggesting participants instead followed their virtual body. Meanwhile, the correlation between following hands and embodiment suggests that following hand movements led to embodiment, or *vice versa*. While we cannot infer causation, participants' dispositions and experiences likely led to realistic behaviours that in turn reinforced their experience through a feedback loop. This is perhaps most clearly demonstrated by the moderate correlation of valence with several behaviours. Happier participants may have been more likely to willingly suspend their disbelief and engage, creating a feedback loop that reinforced that willing engagement.

Strong correlations between questionnaire results suggest that the different measures of presence all were closely interrelated. While these correlations are not enough to unpack a detailed model of their relationships, we can begin to see how deeply interwoven each aspect is. For example, tactile realism may facilitate object presence and embodiment and thereby extend to spatial presence. These aspects also seem to be modulated by other factors such as social dimensions that are deeply rooted in the participants' cultural context and past experiences.

4.1.2 Triangulating methods towards a multi-faceted understanding

To better understand participants' experiences, we triangulated our quantitative and qualitative results. Generally, participants' descriptions align with questionnaires while capturing additional significant elements such as eye contact and proximity. They give insights into what might be happening behind the numbers and unpack the unfolding of the experience over time, giving specificity to what moments led to that experience. Qualitative results suggested that immersion may have been influenced by the rich narrative more than by touch. They also showed that eye contact and proximity with characters led to social presence.

We also found conflicts that indicate complexity or issues with interpreting questions. For example, some participants reported a low sense of embodiment despite following the hands suggesting that embodiment required more than aligning with the body. We also found that participants who mentioned empathy and connection gave highly variable IOS scores for social connection. The clearest conflict was between participants' perception of catching the diary versus the observed behaviour. In all, 18.5% of participants either misremembered ($n = 14$) or could not remember ($n = 6$) whether they tried to catch the diary. This demonstrates the utility of behavioural measures to corroborate with participant's experiences or overcome fallible memories (Henry et al., 1994).

4.2 Richness of experience

Our qualitative results suggest that beyond touch, many components contributed to an overall richness of experience, allowing participants to become immersed. While Slater and Wilbur (1997) described immersion as a property of technology, Ermi and Mäyrä (2005) proposed it as the confluence between design elements, and Vidyarthi (2012) posited it as an active co-creation in the mind of the spectator. Our findings align with Ermi and Mäyrä's (2005) sensory and imaginative dimensions of immersion. A combination of sensory and narrative richness led participants to become immersed and feel present.

4.2.1 Richness of tactility: physical interaction

In the full-touch condition, this experience evoked rich tactility comparable to physical reality. Participants experienced moments of touch completely in intricate detail, making the experience feel real. It encouraged participants to become immersed and feel present in the environment along with the objects and people they touched and saw.

4.2.2 Richness of visuals: social interaction

The compelling visual richness of *Eve 3.0* was captured in participants' descriptions, high questionnaire scores, and surprisingly realistic behaviours in the no-touch condition. Even without touch, the visual experience was compelling, but together they produced an experience that drew participants in deeper. Participants spoke to the power of eye contact and proximity, suggesting that the rich visuals evoked an intimate social presence that overwhelmed some as they stepped back to restore a comfortable level of proximity (Argyle and Dean, 1965; Bailenson et al., 2003).

4.2.3 Richness of story: meaningful engagement

Revealing the contents of a personal diary, *Eve 3.0* presented a rich and emotional story. The action of the characters is equally rich, as they dance in a story that unfolds through layered movement and keeps participants engaged. Beyond touch and visuals, rich narrative elements are critical to immersing participants in the story world and might amplify the presence facilitated through the senses. The diary's focal point in the story ties the rich imaginative and sensory elements of immersion into a single object, leading participants to reach out to catch it for a variety of reasons.

4.3 A willing suspension of disbelief based on individual and experience, context, and content

The rich tactile, visual, and narrative elements of *Eve 3.0* led to a variety of experiences and behaviours from participants. Participants' experiences and behaviours are situated by their past, dispositions, memories, context, expectations, and intentions. The outcome is the result of the interaction between the VR artefact and participants' active role in willingly suspending their disbelief and engaging in meaning-making within the narrative as they co-construct the experience. The settings were a confluence of different cultures, contexts, expectations, and more; all of which played a role in deciding their resulting experience and behaviour.

While our study focused on how tactility could support the sense of presence, participants often described willingly "playing a role." Role-playing might be a useful conceptualization of embodiment in VR that recognizes participants' awareness (Quaglia and Holecek, 2018) and agency over such experiences, and how attention can shift between physical and virtual or even be experienced simultaneously (Morie, 2007).

4.4 Meaningful discomfort

While many experiences were positive, participants also described discomfort. Discomfort was often associated with a lack of agency or social discomfort. Some felt that their body did not represent them. Other participants mentioned discomfort around interacting with school-age children. Indeed, giving participants a clearer role to play from the outset might have helped, for example, telling them who they were and why they were there. We do indeed introduce participants to their character "Eve" in the full performance.

While physical touch from an unseen facilitator might be expected to increase discomfort, it seems that this novel experience drew participants in. Participants became more immersed in their role, and no participant reported discomfort from the performer's touch itself. Elements that did make some participants uncomfortable also drew others into the experience. For example, while eye contact and proximity felt intense, it increased immersion and connected participants to the characters. Discomfort around agency left participants feeling trapped, wanting to help but unable to do so. However, these moments were vital to the story, and built tension that contributed to a meaningful experience. This is a challenge to balance in immersive storytelling, especially with pre-recorded 360° video. Discomfort is not inherently bad and can be vital to producing a meaningful experience (Benford et al., 2012).

4.5 Reciprocal benefits of bringing art and science together: authentic experience, authentic setting, and authentic results

Rooting research in artistic performance allows us to capture the richness and authenticity of an experience intended for real world use. This means that the visual and narrative quality were very high, with a focus on rich storytelling rather than manipulating a particular variable. This ensures a complete experience that best represents how the technology is used outside of the laboratory, especially in an ecological setting like the dance festival. While this can lead to noisier data that can be challenging to interpret, the results we do find are made more meaningful by their authenticity. Moreover, unexpected results are often the most interesting. As we see here, this can lead to more questions than answers, and carefully designed laboratory-based research is required to reciprocally inform such studies. Nonetheless, we demonstrated how a festival can provide a relatively controlled setting that invites a vastly different and potentially more representative demographic.

Bringing art and science together also leads to more ethical research through direct benefits to participants. Research can unfortunately treat participants as subjects rather than collaborators, sometimes even causing harm to evoke a behavioural response (e.g., Petkova and Ehrsson, 2008). By contrast, participants thanked us for their experience and shared profound moments, leaving us feeling elevated and enthusiastic. Despite not being required to participate in the study, the minimal dropout rate of 2.6% (n = 3) speaks of participants' enthusiastic participation.

4.6 Limitations: it's complicated

While conducting our research using an authentic experience in an ecologically valid setting led to rich and interesting results, it also limits how our results can be interpreted. Our settings included many inseparable elements such as demographics, context, facilitator, and headsets used. As a result, it is challenging to identify what differences between the settings may have contributed to the results. However, we simplified and aligned our procedures to limit the impact of any differences beyond the context and participant demographics. As with most VR studies, participants are not necessarily representative of the general population. However, by studying two decidedly different demographics and contexts, our study captures just how much the results can be affected.

As with any VR study, it is difficult to measure complex phenomenological experiences that vary over time through post-experience questionnaires. Limited research indeed shows that embodiment changes dynamically (Keenaghan et al., 2020). For example, some participants described negative experiences in one moment, followed by positive ones in another. It becomes unclear which moment(s) that participants' scores reflect, even when we ask them to reflect on a specific one. Our questionnaires were also limited by a lack of standardized questions that fit a 360° video experience. However, behavioural measures and qualitative responses helped overcome this.

4.7 Future work and research directions

Most systems for social touch fall far short of the richness of human touch (Gallace and Girondini, 2022). Human touch is inextricably imbued with sociocultural meanings, not only making social touch a more potent experience but also introducing a broad range in experiences, depending on how participants perceive such touch. As we found in Section 3.5.3, even touching the diary was imbued with rich social significance through the narrative. Future research should continue exploring the interplay between physical tactility and sociocultural meaning, especially in mediated VR experiences.

While questionnaires are used commonly in VR research, as we found here, they have many limitations. They face challenges around interpretations of language used in the survey, have limited comparability across different media, and are particularly limited in capturing the dynamic experience of presence or even a singular past moment. As suggested by Souza et al. (2021), behavioural measures can be more reliable and may offer a way to analyse presence over time. However, they are still best supported by measures that give insights into participants' interpretations. To reduce recall errors and misinterpretation while digging deeper into the connections between behaviour and experience, we recommend a combination of behavioural observation and micro-phenomenology (Petitmengin et al., 2018). Micro-phenomenology is based on an explication interview and allows deeper access to an authentic singular past experience through the process of evocation. With a highly trained interviewer, it affords a fine-grained exploration of the elements of a very short singular experience. Simultaneous behavioural observations can capture the experience over time and help corroborate and interpret the interview results. This approach could capture the rich temporal nature of presence, measuring presence while also keeping the rich diversity of individual experiences in view.

Here, we began to unpack the complexity of the experience of presence. From our correlation analysis and qualitative results, it seems that many aspects of the phenomenology of VR that are often studied in isolation (e.g., embodiment, spatial presence, social presence) are deeply intertwined. Mediation analysis could help unpack the deeper structure of how these elements are intertwined, keeping in mind that this structure likely depends on many factors embedded in each individual, technology, and media artefact. For this reason, it is also clear that more research is required that establishes a reciprocal relationship between art and science, with each contributing to a deeper and richer understanding of the phenomenology of VR experiences.

4.8 A rich experience of real human touch

This study uncovered how tactile cues and human interaction contribute to a more compelling embodied immersive experience. With the growing use of performer-facilitated touch in immersive performance, it provides a timely guide to artists and researchers alike to better understand how incorporating physical human touch can affect embodiment and presence in the immersive performance. In addition, we hope that our use of pre-recorded 360° video opens other researchers to the potential of an often overlooked technology. Our participants had deep experiences of immersion and presence that real-time computer graphics cannot yet replicate. Even in a pre-recorded 360° video, the incorporation of physical touch is deeply interwoven with other considerations such as the individual experience, the context in which it is viewed, and the sensory and narrative richness of the experience beyond touch.

We hope this study can inform the exploration of novel solutions that root virtual experiences of tactility in the full richness offered by human touch. VR research surrounding touch, especially social touch, should consider as its basis, a full physical experience of touch in its entirety rather than being constrained to what is currently possible with haptic technology. This can help identify important gaps in research that often perpetuate significant assumptions about touch, a sorely under-investigated and misunderstood sense (Gallace and Spence, 2010). As mixed reality becomes more prevalent, this approach also gains importance as our physical and virtual worlds begin to blend into one. Research into real human touch will continue to guide VR towards rich and meaningful 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.

Ethics statement

The studies involving humans were approved by the Simon Fraser University Research Ethics Board and Université Paris 8 Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individuals for the publication of any potentially identifiable images or data included in this article.

Author contributions

JD-S: conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, visualization, writing–original draft, and writing–review and editing. MB: conceptualization, formal analysis, investigation, resources, writing–original draft, and writing–review and editing. ES: conceptualization, data curation, formal analysis, investigation, methodology, writing–original draft, and writing–review and editing. BR: conceptualization, methodology, resources, supervision, and writing–review and editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This research was supported in part by the Social Science and Humanities Research Council of Canada (R632343, R640228). The artefact we researched, *Eve 3.0*, benefited from state aid managed by the French National Research Agency under the Future Investments program bearing the reference ANR-17-EURE-0008. The artwork was produced with the financial assistance of the European Union. Open access publication fees funded by Simon Fraser University.

Acknowledgments

The authors thank Daniel González-Franco for contributing the VR interaction concept and supporting the filming of the 360° video. They also thank Veronica Boniotti for concept development, scriptwriting, and stage assistance and Mark Lee for production and scriptwriting. They especially thank the whole production team, especially the 360° video dancers: Maude Doublet-Viguié, David Fleury, Alexi Mondet, Lola Montaut, Charline Regnier, Mélissa Vaurin, and Camille Nerry. *Eve 3.0* was produced by Compagnie Voix with support from Digital Rise, the Seesaw Project, and Laval Virtual Center. Research venues were provided by Cinedans Amsterdam, Simon Fraser University, University of Calgary, Intangible Realities lab 2.0 at the University of

Santiago de Compostela, and more. For full credits, visit compagnievoix.com/en/projects/creation/eve-3.

Conflict of interest

The performance artefact studied here was conceived by the second author (MB) and produced by Compagnie Voix, a non-profit production association that MB belongs to as artistic director and choreographer. Compagnie Voix, along with the first (JD-S) and second (MB) authors, may receive royalties and other compensation for public performances of Eve 3.0. However, this experimental performance was conceived to investigate new forms of immersive performance, and we do not feel that fair compensation for performances has affected the results presented here. We present a transparent and honest account of our findings in the interest of sharing knowledge across art and science.

The remaining 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

Author disclaimer

The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2023.1336581/full#supplementary-material

References

Argyle, M., and Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry* 28, 289–304. doi:10.2307/2786027

Aron, A., Aron, E. N., and Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *J. Personality Soc. Psychol.* 63, 596–612. doi:10.1037/0022-3514.63.4.596

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2001). Equilibrium theory revisited: mutual gaze and personal space in virtual environments. *Presence Teleoperators Virtual Environ.* 10, 583–598. doi:10.1162/105474601753272844

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2003). Interpersonal distance in immersive virtual environments. *Personality Soc. Psychol. Bull.* 29, 819–833. doi:10.1177/0146167203029007002

Benford, S., Greenhalgh, C., Giannachi, G., Walker, B., Marshall, J., and Rodden, T. (2012). "Uncomfortable interactions," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), CHI '12, New York, NY, USA, May 5 - 10, 2012, 2005–2014. doi:10.1145/2207676.2208347

Bergamo Meneghini, M. (2019). "Performance "Eve, dance is an unplaceable place"," in IEEE GEM 2019, New Haven, CT, 18-21 June 2019, 6.

Bertrand, P., Gonzalez-Franco, D., Cherene, C., and Pointeau, A. (2014). ""the machine to be another": embodiment performance to promote empathy among individuals," in AISB 2014 50th Annual Convention of the AISB, London, UK, 1st-4th April 2014. 5.

Biocca, F. (1997). "The cyborg's dilemma: embodiment in virtual environments," in Proceedings Second International Conference on Cognitive Technology Humanizing the Information Age (Aizu-Wakamatsu City, Japan: IEEE Comput. Soc), 12–26, Aizu-Wakamatsu City, Japan, 25–28 August 1997. doi:10.1109/CT.1997.617676

Biocca, F., Harms, C., and Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: review and suggested criteria. *Presence Teleoperators Virtual Environ.* 12, 456–480. doi:10.1162/105474603322761270

Botvinick, M., and Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature* 391, 756. doi:10.1038/35784

Carifio, J., and Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Med. Educ.* 42, 1150–1152. doi:10.1111/j.1365-2923.2008.03172.x

Cebolla, A., Herrero, R., Ventura, S., Miragall, M., Bellosta-Batalla, M., Llorens, R., et al. (2019). Putting oneself in the body of others: a pilot study on the efficacy of an embodied virtual reality system to generate self-compassion. *Front. Psychol.* 10, 1521. doi:10.3389/fpsyg.2019.01521

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. 2. New York: Routledge. doi:10.4324/9780203771587

Coleridge, S. T. (1817). "Chapter XIV," in *Biographia literaria* (Chicago, Illinois: Poetry Foundation).

Collaço de Oliveira, E., Bertrand, P., Lesur, M. E. R., Palomo, P., Demarzo, M., Cebolla, A., et al. (2016). "Virtual body swap: a new feasible tool to Be explored in health and education," in 2016 XVIII Symposium on Virtual and Augmented Reality (SVR). 81–89, Gramado, Brazil, June 21 2016 to June 24 2016. doi:10.1109/SVR.2016.23

Creswell, J. W., and Creswell, J. D. (2018). *Research design: qualitative, quantitative, and mixed methods approaches*. 5. Thousand Oaks, California: SAGE Publications, Inc.

Creswell, J. W., and Poth, C. N. (2018). *Qualitative inquiry & research design: choosing among five approaches*. 4. Thousand Oaks: SAGE Publications, Inc.

Cumming, G. (2011). *Understanding the new statistics: effect sizes, confidence intervals, and meta-analysis*. New York: Routledge. doi:10.4324/9780203807002

Cumming, G. (2014). The new statistics: why and how. *Psychol. Sci.* 25, 7–29. doi:10.1177/0956797613504966

De Greef, P., and Ijsselsteijn, W. A. (2001). Social presence in a home tele-application. *CyberPsychology Behav.* 4, 307–315. doi:10.1089/109493101300117974

Ermi, L., and Mäyrä, F. (2005). "Fundamental components of the gameplay experience: analysing immersion," in *Changing views: worlds in play*. Editors S. de Castell and J. Jenson (Vancouver, Canada: DiGRA), 14.

Felip, F., Galán, J., Contero, M., and García-García, C. (2023). Touch matters: the impact of physical contact on haptic product perception in virtual reality. *Appl. Sci.* 13, 2649. doi:10.3390/app13042649

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics: and sex and drugs and rock 'n' roll/andy Field*. 4. Thousand Oaks, California: Sage.

Field, T. (2014). *Touch*. 2. Cambridge, MA: MIT Press.

Fulkerson, M. (2013). *The first sense: a philosophical study of human touch*. Cambridge, UNITED STATES: MIT Press.

Gallace, A., and Girondini, M. (2022). Social touch in virtual reality. *Curr. Opin. Behav. Sci.* 43, 249–254. doi:10.1016/j.cobeha.2021.11.006

Gallace, A., Ngo, M. K., Sulaitis, J., and Spence, C. (2012). "Multisensory presence in virtual reality: possibilities & limitations," in *Multiple sensorial media advances and applications: new developments in MulSeMedia (IGI global). 1–38.*

Gallace, A., and Spence, C. (2010). The science of interpersonal touch: an overview. *Neurosci. Biobehav. Rev.* 34, 246–259. doi:10.1016/j.neubiorev.2008.10.004

Harpe, S. E. (2015). How to analyze Likert and other rating scale data. *Curr. Pharm. Teach. Learn.* 7, 836–850. doi:10.1016/j.cptl.2015.08.001

Henry, B., Moffitt, T. E., Caspi, A., Langley, J., and Silva, P. A. (1994). On the "remembrance of things past": a longitudinal evaluation of the retrospective method. *Psychol. Assess.* 6, 92–101. doi:10.1037/1040-3590.6.2.92

Hoffman, H. (1998). "Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments," in Proceedings of the IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180) (Atlanta, GA, USA: IEEE Comput. Soc), 59–63, Atlanta, GA, USA, March 1998. doi:10.1109/VRAIS.1998.658423

Hoppe, M., Rossmy, B., Neumann, D. P., Streuber, S., Schmidt, A., and Machulla, T.-K. (2020). "A human touch: social touch increases the perceived human-likeness of agents in virtual reality," in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), CHI '20, 1–11, New York, NY, USA, April 25-30, 2020. doi:10.1145/3313831.3376719

Ijsselsteijn, W. A., de Kort, Y. A. W., and Haans, A. (2006). Is this my hand I see before me? The rubber hand illusion in reality, virtual reality, and mixed reality. *Presence Teleoperators Virtual Environ.* 15, 455–464. doi:10.1162/pres.15.4.455

Jégo, J.-F., and Bergamo Meneghini, M. (2020). "Let's resonate: how to elicit improvisation and letting go in interactive digital art," in Proceedings of the 7th International Conference on Movement and Computing (Jersey City/Virtual NJ USA: ACM), Jersey City, NJ, USA, July 15-17, 2020, 1–8. doi:10.1145/3401956.3404194

Joris, E., and Vandebroeck, I. (2022). "Delirious Departures," in *ACM SIGGRAPH 2022 immersive pavilion* (New York, NY, USA: Association for Computing Machinery). SIGGRAPH '22, 1–2. doi:10.1145/3532834.3536222

Keenaghan, S., Bowles, L., Crawfurd, G., Thurlbeck, S., Kentridge, R. W., and Cowie, D. (2020). My body until proven otherwise: exploring the time course of the full body illusion. *Conscious. cognition* 78, 102882. doi:10.1016/j.concog.2020.102882

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

King, N. (2012). "Qualitative organizational research: core methods and current challenges," in *Qualitative organizational research: core Methods and current challenges* (55 City Road: SAGE Publications, Inc.), 426–450. doi:10.4135/9781526435620

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

Lombard, M. (2013). *Measuring presence: the Temple presence inventory*.

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. 15, Valencia, Spain, 12-15 May 2015.

Miles, J., and Shevlin, M. (2001). *Applying regression & correlation: a guide for students and researchers*. London; Thousand Oaks, Calif: Sage Publications.

Morie, J. F. (2007). Performing in (virtual) spaces: embodiment and being in virtual environments. *Int. J. Perform. Arts Digital Media* 3, 123–138. doi:10.1386/padm.3.2-3.123_1

Oh, C. S., Bailenson, J. N., and Welch, G. F. (2018). A systematic review of social presence: definition, antecedents, and implications. *Front. Robotics AI* 5, 114. doi:10.3389/frobt.2018.00114

Peck, T. C., and Gonzalez-Franco, M. (2021). Avatar embodiment. A standardized questionnaire. *Front. Virtual Real.* 1. doi:10.3389/frvir.2020.575943

Petitmengin, C., Remillieux, A., and Valenzuela-Moguillansky, C. (2018). Discovering the structures of lived experience. *Phenomenology Cognitive Sci.* 18, 691–730. doi:10.1007/s11097-018-9597-4

Petkova, V. I., and Ehrsson, H. H. (2008). If I were you: perceptual illusion of body swapping. *PLOS ONE* 3, e3832. doi:10.1371/journal.pone.0003832

Popat, S. (2016). Missing in action: embodied experience and virtual reality. *Theatre J. Baltim.* 68, 357–378. doi:10.1353/tj.2016.0071

Quaglia, J. T., and Holecek, A. (2018). "Lucid virtual dreaming: antecedents and consequents of virtual lucidity during virtual threat," in 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (Tuebingen/Reutlingen, Germany: IEEE), Reutlingen, Germany, 18-22 March 2018, 1–72. doi:10.1109/VR.2018.8446546

Reiner, M., and Hecht, D. (2009). Behavioral indications of object-presence in haptic virtual environments. *CyberPsychology Behav.* 12, 183–186. doi:10.1089/cpb.2008.0141

Rietzler, M., Geiselhart, F., Frommel, J., and Rukzio, E. (2018). "Conveying the perception of kinesthetic feedback in virtual reality using state-of-the-art hardware," in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18 (Montreal QC, Canada: ACM Press), Montreal, QC, Canada, April 21 - 26, 2018, 1–13. doi:10.1145/3173574.3174034

Rogers, Y., and Marshall, P. (2017). Research in the wild. *No. 37 in synthesis lectures on human-centred informatics (morgan & claypool).*

Rostami, A., Rossitto, C., and Waern, A. (2018). "Frictional Realities: enabling immersion in mixed-reality performances," in Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video - TVX '18 (SEOUL, Republic of Korea: ACM Press), SEOUL, Republic of Korea, June 26-28, 2018, 15–27. doi:10.1145/3210825.3210827

Russell, J., Weiss, A., and Mendelsohn, G. (1989). Affect grid: a single-item scale of pleasure and arousal. *J. Personality Soc. Psychol.* 57, 493–502. doi:10.1037/0022-3514.57.3.493

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

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

Slater, M., Pérez Marcos, D., Ehrsson, H., and Sanchez-Vives, M. V. (2009). Inducing illusory ownership of a virtual body. *Front. Neurosci.* 3, 214–220. doi:10.3389/neuro.01.029.2009

Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLOS ONE* 5, e10564. doi:10.1371/journal.pone.0010564

Slater, M., and Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence Teleoperators Virtual Environ.* 6, 603–616. doi:10.1162/pres.1997.6.6.603

Souza, V., Maciel, A., Nedel, L., and Kopper, R. (2021). Measuring presence in virtual environments: a survey. *ACM Comput. Surv.* 54, 1–37. 1–163:37. doi:10.1145/3466817

Stevens, B., Jerrams-Smith, J., Heathcote, D., and Callear, D. (2002). Putting the virtual into reality: assessing object-presence with projection-augmented models. *Presence Teleoperators Virtual Environ.* 11, 79–92. doi:10.1162/105474602317343677

van Erp, J. B. F., and Toet, A. (2015). Social touch in Humanâ€"Computer interaction". *Front. Digital Humanit.* 2. doi:10.3389/fdigh.2015.00002

Vidyarthi, J. (2012). *Sonic cradle: evoking mindfulness through 'immersive' interaction design*. Surrey, BC, Canada: Master's thesis, Simon Fraser University.

Wilson, H. R. (2020). New ways of seeing, feeling, being: intimate encounters in virtual reality performance. *Int. J. Perform. Arts Digital Media* 16, 114–133. doi:10.1080/14794713.2020.1770531

Yarosh, S., Wang, X., and Yao, Y. (2022). Perceptions of visual and multimodal symbolic mediated social touch: role of technology modality, relationship, and task emotional salience. *Int. J. Human-Computer Stud.* 159, 102757. doi:10.1016/j.ijhcs.2021.102757

Zhang, Y., Wang, Y., and Liu, Q. (2022). "Touch the history in virtuality: combine passive haptic with 360° videos in history learning," in 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW). 824–825, Christchurch, New Zealand, March 12 2022 to March 16 2022. doi:10.1109/VRW55335.2022.00263

frontiers | Frontiers in Virtual Reality

TYPE Case Report
PUBLISHED 13 March 2024
DOI 10.3389/frvir.2024.1284299

Check for updates

OPEN ACCESS

EDITED BY
Orestis Georgiou,
Ultraleap Ltd., United Kingdom

REVIEWED BY
Justin Maximilian Mittelstädt,
German Aerospace Center (DLR), Germany
Tiago Henrique Falk,
Université du Québec, Canada

*CORRESPONDENCE
Sara A. Freedman,
✉ sara.freedman@biu.ac.il

RECEIVED 28 August 2023
ACCEPTED 26 February 2024
PUBLISHED 13 March 2024

CITATION
Freedman SA, Dayan E, Senitsky M, Bellin E, Attias Y, Richman T, Eitan R and Tatsa-Laur L (2024), Case report: the addition of olfaction to virtual reality enhanced exposure therapy for PTSD.
Front. Virtual Real. 5:1284299.
doi: 10.3389/frvir.2024.1284299

Case report: the addition of olfaction to virtual reality enhanced exposure therapy for PTSD

Sara A. Freedman[1]*, Ehud Dayan[2], Michal Senitsky[1], Elisheva Bellin[1], Yossi Attias[1], Tzvi Richman[1], Renana Eitan[3] and Lucian Tatsa-Laur[4]

[1]School of Social Work, Bar Ilan University, Ramat Gan, Israel, [2]Sonarion Ltd, Jerusalem, Israel, [3]Tel Aviv Medical Center—Ichilov, Tel Aviv, Israel, [4]Health Department, Medical Corps, Israel Defense Forces, Tel Aviv, Israel

Posttraumatic stress disorder (PTSD) occurs in approximately 20% of individuals following exposure to a potentially traumatic event. Re-experiencing in the form of intrusive memories is a hallmark feature, and often includes sensory elements, including odor. A small body of studies have examined the relationship between PTSD and odor evoked memories. Patients avoid smells that remind them of their traumatic event, and odors are known triggers for intrusive memories and distress. PTSD is associated with higher ratings of odor intensity and unpleasantness however accuracy of odor detection is often impaired. It has been suggested that adding trauma-related odors will enhance treatment, however little data exists. This case study illustrates the addition of trauma specific olfaction in Virtual Reality (VR) enhanced exposure therapy for combat related PTSD. Treatment was successful, the patient no longer met criteria for PTSD at the end of treatment, and self-report symptoms scores indicated clinically significant reductions. The patient indicated at the end of treatment that he would not have been helped had the therapy not used VR, and the addition of odors helped him process the traumatic memories. Future research should assess the addition of olfaction in a systematic way, in order to assess its specific impact in the effectiveness of PTSD treatment.

KEYWORDS

PTSD, virtual reality, olfaction, exposure therapy, case report PTSD, case study

Introduction

"Smell is a potent wizard that transports us across thousands of miles and all the years we have lived. . .. odors, instantaneous and fleeting, cause my heart to dilate joyously or contract with remembered grief".

This quote from Helen Keller (Keller, 1910) expresses an experience that is common to us all: being taken back in time, vividly remembering a seemingly forgotten past event after encountering an odor. This is known as the *Proust effect*, following the example the writer Proust gave of the smell of biscuits that took him back to a lucid childhood memory (Proust and Sturrock, 2003).

Odor evoked memories are autobiographical memories, often from the first decade of life. Perceptual elements of memory include all sensory information: visual, tactile, auditory and olfactory. Studies have shown that olfactory memories are the most likely to be related to remembering emotionally laden events (Toffolo et al., 2012). Odor associated with a stressful situation has been shown in rats to lead to conditioned fear responses to the odor by itself (Bombail, 2019). These findings are also explained by the neuroanatomy of emotion and olfaction: the olfactory system and the limbic system have a direct interconnection (Kontaris et al., 2020).

Posttraumatic stress disorder (PTSD) occurs in approximately 20% of individuals following exposure to a potentially traumatic event (Kessler et al., 2017). Symptoms of PTSD include four constellations: re-experiencing of the traumatic event via flashbacks, nightmares and intrusive memories, avoidance of trauma cues, negative emotions and cognitions and hyperarousal (Schaal et al., 2015). Re-experiencing in the form of intrusive memories are suggested to be fear memories that have not extinguished and are a hallmark feature of the disorder (Duek et al., 2020).

A small body of studies that have examined the relationship between PTSD and odor evoked memories. Widespread anecdotal clinical experience exists regarding patients' avoidance of smells that remind them of their traumatic event, and odors are known triggers for intrusive memories and distress. As described a century ago by a First World War poet and sufferer of PTSD: "... any unusual smell, even a sudden strong smell of flowers in a garden, was enough to send me trembling." (Graves, 1960). An early research paper examining aspects of traumatic memories shows that they included odors 77% of the time (Van der Kolk and Fisler, 1995). A study of refugees showed that 45% experienced panic attacks triggered by odor, and many of these involved flashbacks to the traumatic event (Hinton et al., 2004). Between 4%–10% of assault victims with PTSD reported smells or tastes related to intrusions (Michael et al., 2005). A more recent paper describes three case studies that report olfactory triggers for traumatic memories (Vermetten and Bremner, 2003).

Experimental studies of combat related trauma survivors have shown that PTSD is associated with higher ratings of odor intensity and unpleasantness, and odors are more likely to act as triggers for posttraumatic symptoms (Dileo et al., 2008; Cortese et al., 2015; Tjalvin et al., 2017; Bedwell et al., 2018; Wilkerson et al., 2018). One study of PTSD following child maltreatment showed similar results (Croy et al., 2010). Similar to studies of other types of psychopathology such as depression, levels of odor detection in general are impaired in PTSD relative to healthy controls (Dileo et al., 2008; Cortese et al., 2015). Taken together these studies indicate that PTSD symptom levels are related to increased sensitivity to unpleasant trauma-related odors, which cause more distress, while at the same time showing decreased sensitivity to odors in general. These studies clearly point to the conclusion that olfactory memories may play an important role in PTSD.

PTSD is a chronic and costly disorder (Kessler, 2000) and current effective treatments help only between 60%–70% of people (Bisson et al., 2019). The most effective PTSD treatments involve interventions that target traumatic memories, with a goal of moving from a perceptual sensory memory to one that is more verbal (Foa and Kozak, 1986). These trauma-focused cognitive behavioural treatments (TF-CBT) often ask the patient to reconstruct the narrative of the traumatic event, telling this story repeatedly. It is important that the patient feels emotionally connected to this narrative, so the patient is requested to close their eyes, and use first person, present tense, in order to facilitate this. However, this key component of therapy can be challenging: some patients do not have clear memories of the event, while others find it hard to retell the narrative of these memories. Often, with events that involve atrocities or unusual elements, the patient finds it hard to find the appropriate language. Therefore, ways to embellish this type of therapy with agents that will enhance memory retrieval or reconsolidation such as D-cycloserine (Cukor et al., 2009; de Kleine et al., 2015) and intranasal oxytocin (Flanagan et al., 2018) have been examined.

The use of Virtual Reality (VR) to enhance this type of therapy is well established (e.g., ('Skip' and Shilling, 2017; Rizzo et al., 2009; Skip et al., 2010; Rothbaum et al., 2001)) and allows the patient to retell the narrative while inside a headset, experiencing scenes that are similar to the traumatic event. VR enhanced therapy is effective, and can overcome some barriers to treatment, such as avoidance, as well as allowing the patient easier access to emotions when retelling the narrative.

It has been suggested that adding odors that are related to the traumatic event will improve treatment outcomes, and this has been carried out both in hypnotherapy (Abramowitz and Lichtenberg, 2010) as well as in the context of Virtual Reality enhanced Exposure Therapy (Van Veelen et al., 2021; Lo and Hurley, 2022), This case study describes the treatment of combat related PTSD using VR enhanced exposure therapy with the addition of olfaction.

Case description

Patient Y

Y was a 25-year-old single man, with no history of psychiatric problems. He was an army veteran, who had served in compulsory service from the ages of 18–21. During the last year of his service he was serving as a combat medic during a month-long conflict, when he was exposed to several potentially traumatic events. Y developed symptoms of PTSD immediately following the conflict. During the next 4 years, his symptoms worsened but he did not seek help during this time. He self-referred for treatment 4 years later. Figure 1 illustrates the timeline of his history and treatment.

Current treatment

The treatment took place as part of an ongoing randomized clinical trial comparing Prolonged Exposure with Virtual Reality enhanced Exposure Therapy, for combat related PTSD. The study was approved by the Helsinki Committee, Jerusalem Mental Health Center and the Medical Unit, IDF (NCT-03000478). Participants were assessed pre and post treatment by an assessor blind to treatment condition.

Assessment measures

A clinical assessment was carried out by a trained clinician before and after treatment. This included self-report measures:

PTSD symptom levels were measured using the PCL5 (Blevins et al., 2015), a 20 item self-report questionnaire covering all PTSD symptoms.

Depression symptom levels were measured using the BDI (Beck et al., 1996) a 21 item self-report questionnaire.

10.3389/frvir.2024.1284299

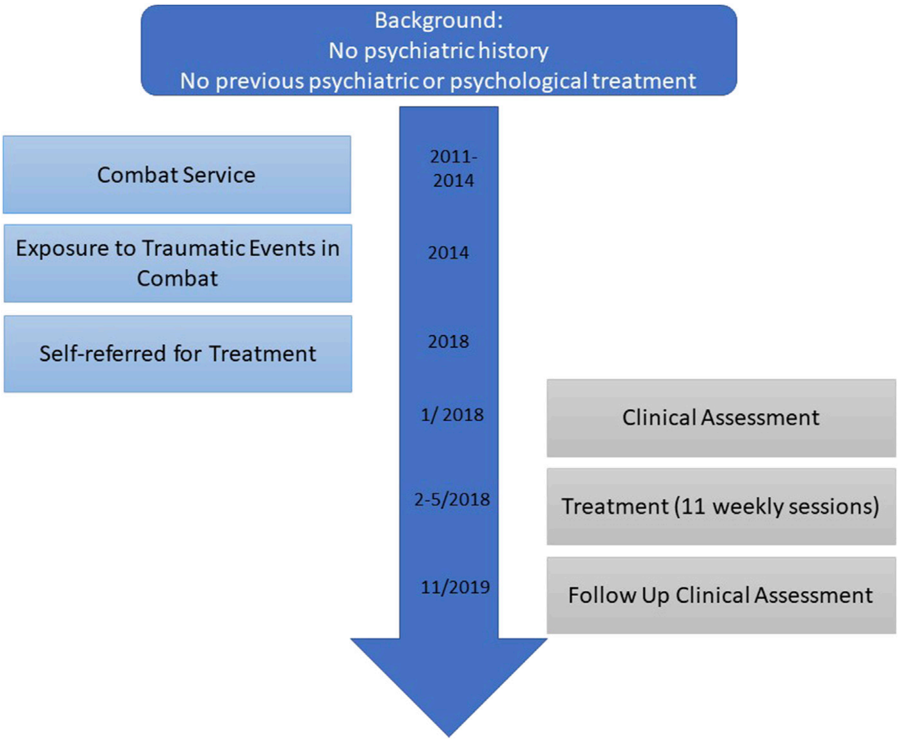


FIGURE 1
Timeline of treatment.

FIGURE 2
Scene from virtual Azza.

Dissociation levels were measured using the DES (Sheehan et al., 1998) a 28 item self-report questionnaire.

In addition, clinical interviews were carried out: the MINI (Sheehan et al., 1998) o assess all current and past Axis I psychiatric disorders, and the CAPS5 (Weathers et al., 2013) to assess PTSD intensity and presence.

Instruments

The Virtual Reality system used Oculus Rift VR goggles with 110° diagonal fields of view per eye and 1080 by 120 pixels resolution per eye (90 hz). The goggles were plugged into a workstation with Intel Core i5 6500 (3.2 GHz) 8 GB Memory 1 TB HDD NVIDIA GeForce GTX

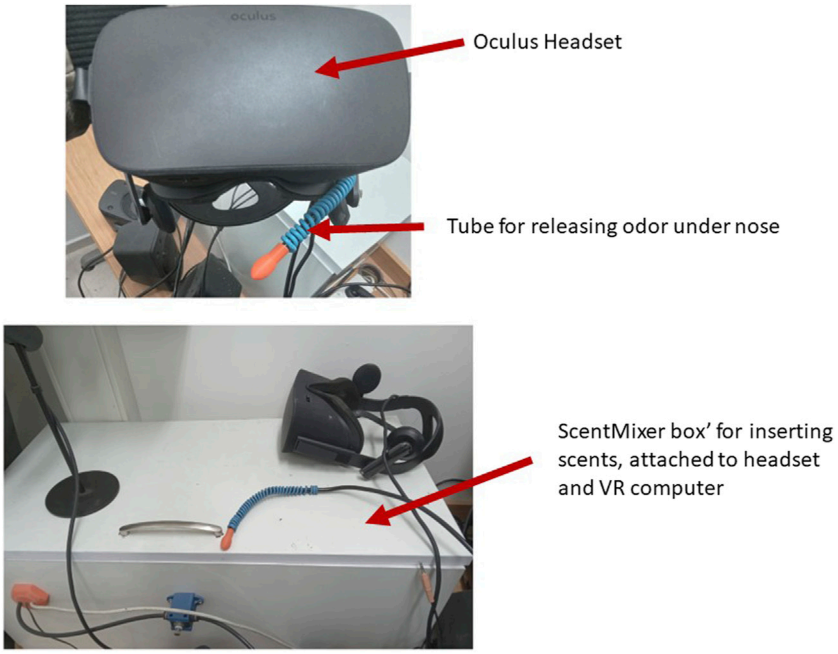


FIGURE 3
Scent mixer.

970 4 GB GDDR5. The software (Virtual Azza) was developed by Sonarion Ltd. using Unity and used urban fighting scenes. Some of the scenes contained soldiers, military vehicles and tank turrets. The OS was windows 10- home edition - 64 bit (Figure 2).

Olfaction System: The ScentMixer system designed by Sonarion Ltd. (http://sonarion.com) was used in this project (Figure 2). The system enables the therapist to dictate odor release manually or automatically. When using the auto mode, the odor intensity varies in relation to the position of the smell source in the VR scene. The odor flow releases short (few seconds) airbursts in ~300 mHg pressure producing an airflow of 0.03LPM/s. As can be seen in Figure 3, the odor is released a few centimeters from the patient nose via a tube attached to the VR headset. The odor is then washed away by fresh air. Different scents can be added, dependent on the patient's individual narrative; in the case presented here gun oil was used.

Procedure

All participants, including Y, signed written informed consent. Y also signed permission to publish details of his treatment. Y was randomized to receive Virtual Reality enhanced treatment. All therapists had received training in Cognitive Behavior Therapy, in Prolonged Exposure, in the use of VR and the VR enhanced exposure therapy protocol and received weekly supervision from SAF. Sessions were videoed and reviewed for the purpose of supervision.

Assessment

During the initial assessment, Y reported that he was single and not working. Apart from the combat related incidents, he also reported being exposed to a natural disaster. In the 4 years that had passed since the conflict, Y found it increasingly difficult to function. At the time of the assessment, he had not been working for several months. Y reported the main traumatic incident he had experienced was the recovery of several bodies of soldiers who had been killed. He had no recollection of where or when this had exactly taken place. Y had not talked about his experiences with anyone. As can be seen (Table 1), Y reported initial high levels of PTSD, dissociation and depression. He reported no other psychiatric diagnoses and no previous history of psychiatric problems or treatment.

Treatment

Following the assessment, Y began weekly treatment, receiving 11 sessions of Virtual Reality Enhanced Treatment. The VR protocol follows the Prolonged Exposure (PE) protocol closely. The therapy includes weekly sessions of 90 min, and the patient is given exercises to complete between sessions. The first session includes psychoeducation about PTSD and breathing retraining. In session two, *in vivo* exposure is introduced, and a hierarchy of avoided situations is built. The patient begins this exposure as homework assignments, and each subsequent session this is discussed and further exposure assignments are agreed upon. From session three, imaginal exposure is begun. This consists of 45–60 min of the patient reliving the traumatic memory, telling the narrative in first person and present tense, over and over again. The patient is asked every 5 minutes to rate his Subjective Units of Distress (SUDS, from 0 to 100). Sessions are recorded, and the patient is requested to listen to the whole session once between sessions, and the imaginal exposure daily. The VR protocol differs from the PE protocol in one

TABLE 1 PTSD, depression and dissociation measures.

	Pre-treatment	Week 1	Week 6	Week 11	Post treatment
CAPS5	39				16
PCL5	43	36	24	6	
BDI	15	14	7	5	
DES		20	3.9	3.2	

respect: the patient carries out the imaginal exposure whilst using the VR headset. Thus, Y received 11 weekly session, and began the narrative exposure with VR from session three, as per the PE protocol.

Treatment course

Y's initial session concentrated on psychoeducation and breathing retraining. In the second session Y and the therapist focused on building *in vivo* exposure hierarchy. These exposures included listening to radio and television news, talking to friends from his army days and talking to his friends about his experiences.

In this session, Y was also introduced to the VR equipment, including a neutral scenario in VR. This is done to ensure that the patient has no negative side effects precluding the use of the VR, such as cybersickness. It also gives the patients an opportunity to become accustomed to the VR room, which looks different to a normal clinic room, before the session where imaginal exposure takes place.

In session three, Y was introduced to the rationale for imaginal exposure, and began the exposure while using the VR headset. The VR has five conflict-based scenarios, these are not hierarchical (i.e., one is not considered more anxiety provoking than another) but reflect different scenes the veteran may have encountered. The therapist and patient decide each session on which scenario to use, and can change between scenarios if this will help the patient feel more connected. In this session, within minutes of beginning to re-tell the traumatic incident, he began crying heavily. There were few details given, and at several points he said he could not continue. His SUDS remained high during these sessions.

During session four, Y reported during the processing that the imaginal exposure that it "was less hard this time" and he noticed a difference in how he saw the event. He was less shocked by the story and mentioned that the more he got used to the story, the more he was able to remember that he is no longer in the traumatic event.

By session five, Y was concentrating on hotspots during the imaginal exposure (specific points in the narrative that are associated with high distress), and his overall levels of distress were lower. Y noted in this session that he required fewer breaks when telling the story.

In session six, the importance of olfaction became more apparent. Y did not cry when discussing going through the pockets of the soldiers who died but reported how hard it was to manage the smells related to the incident. He reported experiencing olfactory flashbacks and was very distressed by these.

During sessions six to eight, Y concentrated on retelling the hotspots (the hardest parts of the narratives); when he reached the part of the narrative where he had reported difficulty regarding smell, the therapist added the smell of gun oil. The scent mixer is loaded with the scent (which is actual gun oil) before the start of the session, and the therapist presses a command on the desktop to release the smell. The smell was released each time the patient reached the part of the narrative when the odor was salient. In each subsequent session (after session 6), the smell was added during imaginal exposure.

By session 8, Y reported significantly lower SUDS, no higher than 30. In the last session, Y related the whole narrative, not just hotspots, and reported SUDS of 0.

Treatment outcome

As can be seen in Table 1, Y's symptoms reduced drastically during the treatment, and the clinical assessment at the end of treatment showed that he no longer met criteria for PTSD.

Patient's perspective

A qualitative interview was carried out following completion of therapy. The patient was interviewed by a research team member who had no knowledge of the treatment and its outcome.

Y explained that the VR component of the treatment was salient in its success. There were several components of the VR that he felt specifically contributed to the effectiveness of the treatment. Firstly, Y had never spoken of his experiences with anyone, and was reluctant to even tell his therapist. He felt that the VR enabled this, since he could not actually see the therapist during the imaginal exposure although he knew she was present. In this way, his story "went out" (in his words) without him needing to make eye contact. He would not, he felt, have been able to tell the narrative to a therapist without the VR, and likened it to the invisibility cloak in Harry Potter. Y knew he was not alone, and that the therapist was there, however the VR gave him privacy and control. He described sometimes closing his eyes when he needed to, adding to his feeling of control.

Secondly, Y felt the visual scenes of combat were very helpful. Y had no clear memory of some of the events he had experienced, and the VR allowed him to remember with more clarity. He was reassured that the scenes were clearly related to his combat experiences but were not so realistic that they would be hard to differentiate from reality.

Thirdly, Y described the addition of the odors as extremely important. During the therapy, Y began experiencing olfactory flashbacks. He felt, that the addition of an odor that was part of these flashbacks, that he smelt at the point of the narrative when he had encountered the odor during the traumatic event, was very important to his ability to process the event. This combination of the visual component with the olfactory within the VR was challenging for him, however he described that it allowed him to retrieve the traumatic memory, and learn to cope with it. Y had not talked about the events for the 4 years between their occurrence and the therapy partly because he had very unclear memories of the events. The VR enhanced therapy allowed him to remember in much better detail what he had experienced. He was sure that the visual and olfactory cues were essential in this process.

Discussion

Evidence based interventions for PTSD demonstrate effectiveness for the majority of patients (Forbes et al., 2020). However, many patients avoid therapy, dropout or remain symptomatic (van Minnen et al., 2002; Hoge et al., 2004) and therefore methods that can enhance treatment effects are important. This paper describes the use of Virtual Reality in Exposure Therapy treatment of a patient with combat related PTSD where treatment was enhanced using olfaction during imaginal exposure. As with most VR enhanced therapies, the content of the VR scenarios was similar to the narrative of the patients but does not correspond exactly. Although it has been suggested that individualized VR scenarios are important (see (Van Veelen et al., 2021), both very realistic although generic scenarios (Skip et al., 2010) as well as abstract scenarios (Botella et al., 2015) have been found to be effective and at the theory underlying exposure based treatment for PTSD places importance on the ability of the patient to emotionally process the traumatic memory (Foa and Kozak, 1986) and employs techniques to enhance this, for instance asking the patient to retell the narrative using first person, present tense. Since this process can be challenging for patients, methods to enhance prolonged exposure have been examined (e.g., (Flanagan et al., 2018). Using Virtual Reality, and adding sensory elements such as odor, are additional ways to increase emotional connections. It remains unclear to what degree VR in general, and odor in particular, contribute to this process. One study has shown that VR helped patients who had not responded to previous trauma-focused treatments, possibly indicating that the VR plays an important role in enhancing emotional processing (Difede et al., 2007). Moving from case studies to systematic randomized controlled trials that tease out different therapeutic techniques is necessary to answer this question, and future studies that address this are needed.

This case study does, however, show the way in which odor can be incorporated into VR enhanced PE. The addition of olfaction may add to the efficiency and depth to which a survivor is able to recall memories and engage in the imaginal exposure of the therapy in deeper ways. Further, the addition of olfaction can allow for a more holistic and full-body experience when trying to reimagine and recreate in one's mind the events of the past: this need to be further explored.

val.

Limitations

Several limitations are noted in the context of this study. Firstly, as a case study, only one case is explored here. A larger sample size may increase generalizability of the case findings. Further, while qualitative data can offer an important perspective on the connection between olfaction deepening the imaginal exposure process, additional data is needed on the significant impact that olfaction can have on the process, comparing subjects where odor was used during treatment as compared to subjects who did not receive the addition of odor. These findings could strengthen the impact that odor has on the treatment.

Conclusion

This case study describes VR enhanced treatment for PTSD, with the addition of odor. While it is not possible to delineate the contribution of each element of the treatment to its success, the patient was confident that the VR and the odor were essential for treatment to be effective. Further studies should examine the addition of odor in treatment in a systematic manner.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Helsinki Committee, Jerusalem Mental Health Center and Helsinki Committee, Medical Unit, IDF. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article. Written informed consent was obtained from the participant/patient(s) for the publication of this case report.

Author contributions

SF: Conceptualization, Methodology, Writing–original draft. ED: Software, Writing–original draft. MS: Investigation, Writing–original draft. EB: Writing–original draft, Writing–review and editing. YA: Project administration, Writing–original draft. TR: Investigation, Writing–original draft. RE: Supervision, Writing–review and editing. LT-L: Resources, Supervision, Writing–original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded by a grant from the IDF Medical Corps.

Conflict of interest

Author ED was employed by Sonarion Ltd.

The remaining 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.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abramowitz, E. G., and Lichtenberg, P. (2010). A new hypnotic technique for treating combat-related posttraumatic stress disorder: a prospective open study. *Int. J. Clin. Exp. Hypn.* 58 (3), 316–328. doi:10.1080/00207141003760926

Beck, A. T., Steer, R. A., and Brown, G. K. (1996). *BDI-II manual*. San Antonio, TX: The Psychological Corporation.

Bedwell, J. S., Bohil, C. J., Neider, M. B., Gramlich, M. A., Neer, S. M., O'Donnell, J. P., et al. (2018). Neurophysiological response to olfactory stimuli in combat veterans with posttraumatic stress disorder. *J. Nerv. Ment. Dis.* 206 (6), 423–428. doi:10.1097/nmd.0000000000000818

Bisson, J. I., Berliner, L., Cloitre, M., Forbes, D., Jensen, D. K., Lewis, C., et al. (2019). The International society for traumatic stress studies new guidelines for the prevention and treatment of posttraumatic stress disorder: methodology and development process. *J. Trauma. Stress* 32 (4), 475–483. doi:10.1002/jts.22421

Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., and Domino, J. L. (2015). The posttraumatic stress disorder checklist for DSM-5 (PCL-5): development and initial psychometric evaluation. *J. Trauma. Stress* 28, 489–498. doi:10.1002/jts.22059

Bombail, V. (2019). Perception and emotions: on the relationships between stress and olfaction. *Appl. Animal Behav. Sci.* 212 (February 2018), 98–108. doi:10.1016/j.applanim.2018.12.013

Botella, C., Serrano, B., Baños, R. M. G. P. A., and García-Palacios, A. (2015). Virtual reality exposure-based therapy for the treatment of post-traumatic stress disorder: a review of its efficacy, the adequacy of the treatment protocol, and its acceptability. *Neuropsychiatr. Dis. Treat.* 11, 2533–2545. doi:10.2147/ndt.s89542

Cortese, B. M., Leslie, K., and Uhde, T. W. (2015). Differential odor sensitivity in PTSD: implications for treatment and future research. *J. Affect. Disord.* 179 (2015), 23–30. doi:10.1016/j.jad.2015.03.026

Croy, I., Schellong, J., Joraschky, P., and Hummel, T. (2010). PTSD, but not childhood maltreatment, modifies responses to unpleasant odors. *Int. J. Psychophysiol.* 75 (3), 326–331. doi:10.1016/j.ijpsycho.2010.01.003

Cukor, J., Spitalnick, J., Difede, J. A., Rizzo, A., and Rothbaum, B. O. (2009). Emerging treatments for PTSD. *Clin. Psychol. Rev.* 29 (8), 715–726. doi:10.1016/j.cpr.2009.09.001

de Kleine, R. A., Smits, J. A. J., Hendriks, G.-J., Becker, E. S., and van Minnen, A. (2015). Extinction learning as a moderator of D-cycloserine efficacy for enhancing exposure therapy in posttraumatic stress disorder. *J. Anxiety Disord.* 34, 63–67. doi:10.1016/j.janxdis.2015.06.005

Difede, J., Cukor, J., Jayasinghe, N., Patt, I., Jedel, S., Spielman, L., et al. (2007). Virtual reality exposure therapy for the treatment of Posttraumatic stress disorder following September 11, 2001. *J. Clin. Psychiatry* 68, 1639–1647.

Dileo, J. F., Brewer, W. J., Hopwood, M., Anderson, V., and Creamer, M. (2008). Olfactory identification dysfunction, aggression and impulsivity in war veterans with post-traumatic stress disorder. *Psychol. Med.* 38 (4), 523–531. doi:10.1017/s0033291707001456

Duek, O., Raphael, T., Robert, S., Fried, E. I., and Harpaz-Rotem, I. (2020). Network analysis of PTSD and depressive symptoms in 158, 139 treatment - seeking veterans with PTSD. *Depress. Anxiety* 38 (October), 1–9. doi:10.1002/da.23112

Flanagan, J. C., Sippel, L. M., Wahlquist, A., Moran-Santa Maria, M. M., and Back, S. E. (2018). Augmenting Prolonged Exposure therapy for PTSD with intranasal oxytocin: a randomized, placebo-controlled pilot trial. *J. Psychiatric Res.* 98, 64–69. doi:10.1016/j.jpsychires.2017.12.014

Foa, E. B., and Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychol. Bull.* 99 (1), 20–35. doi:10.1037//0033-2909.99.1.20

Forbes, D., Bisson, J. I., Monson, C. M., and Berliner, L. (2020), *Effective treatments for PTSD: practice guidelines from the international society for traumatic stress studies*. New York, NY, USA: Guilford.

Graves, R. (1960). *Goodbye to all that*. 6th. Harmondsworth, UK: Penguin Books.ed.

Hinton, D., Pich, V., Chhean, D., and Pollack, M. (2004). Olfactory-triggered panic attacks among Khmer refugees: a contextual approach. *Transcult. Psychiatry* 41, 155–199. doi:10.1177/1363461504043564

Hoge, C. W., Castro, C. A., Messer, S. C., McGurk, D., Cotting, D. I., and Koffman, R. L. (2004). Combat duty in Iraq and Afghanistan, mental health problems, and barriers to care. *N. Engl. J. Med.* 351, 13–22. doi:10.1056/nejmoa040603

Keller, H. (1910). *The World I live in*. New York, NY, USA: The Century Co.

Kessler, R. C. (2000). Posttraumatic stress disorder: the burden to the individual and to society. *J. Clin. Psychiatry* 61 (Suppl. 5), 4–14.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Benjet, C., Bromet, E. J., Cardoso, G., et al. (2017). Trauma and PTSD in the WHO World mental health surveys. *Eur. J. Psychotraumatology* 8 (Suppl. 5), 1353383. doi:10.1080/20008198.2017.1353383

Kontaris, I., East, B. S., and Wilson, D. A. (2020). Behavioral and neurobiological convergence of odor, mood and emotion: a review. *Front. Behav. Neurosci.* 14 (March), 35–15. doi:10.3389/fnbeh.2020.00035

Lo, W., and Hurley, R. A. (2022). Extended reality Technologies: expanding therapeutic approaches for PTSD. *J. Neuropsychiatry Clin. Neurosci.* 34, A4–A5. doi:10.1176/appi.neuropsych.21100244

Michael, T., Ehlers, A., Halligan, S. L., and Clark, D. M. (2005). Unwanted memories of assault: what intrusion characteristics are associated with PTSD? *Behav. Res. Ther.* 43 (5), 613–628. doi:10.1016/j.brat.2004.04.006

Proust, M., and Sturrock, J. (2003). *Search of lost time*. London, UK: Penguin.

Rizzo, A., Newman, B., Parsons, T., Reger, G., Holloway, K., Gahm, G., et al. (2009). "Development and clinical results from the virtual Iraq exposure therapy application for PTSD," in 2009 Virtual Rehabilitation International Conference, Haifa, Israel (IEEE), 8–15.cited 2023 Dec 14 Available from: http://ieeexplore.ieee.org/document/5174198/

Rothbaum, B. O. B. B. O., Hodges, L. F. L. F., Ready, D., Graap, K., Alarcon, R. D., Bo, R., et al. (2001). Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. *J. Clin. Psychiatry* 62, 617–622. doi:10.4088/jcp.v62n0808

Schaal, S., Koebach, A., Hinkel, H., and Elbert, T. (2015). Posttraumatic stress disorder according to DSM-5 and DSM-IV diagnostic criteria: a comparison in a sample of congolese ex-combatants. *Eur. J. Psychotraumatology* 6, 24981–24988. doi:10.3402/ejpt.v6.24981

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry* 59, 22–57.1998

Skip, R. A., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., et al. (2010). Development and early evaluation of the *Virtual Iraq/Afghanistan* exposure therapy system for combat-related PTSD. *Ann. N. Y. Acad. Sci.* 1208 (1), 114–125. doi:10.1111/j.1749-6632.2010.05755.xOct

Skip, R. A., and Shilling, R. (2017). Clinical Virtual Reality tools to advance the prevention, assessment, and treatment of PTSD. *Eur. J. Psychotraumatology* 8 (Suppl. 5), 1414560. doi:10.1080/20008198.2017.1414560Oct 27

Tjalvin, G., Magerøy, N., Bråtveit, M., Lygre, S. H. L., Hollund, B. E., and Moen, B. (2017). Odour as a determinant of persistent symptoms after a chemical explosion, a longitudinal study. *Ind. Health* 55 (2), 127–137. doi:10.2486/indhealth.2016-0155

Toffolo, M. B. J., Smeets, M. A. M., and van den Hout, M. A. (2012). Proust revisited: odours as triggers of aversive memories. *Cognition Emot.* 26 (1), 83–92. doi:10.1080/02699931.2011.555475

Van der Kolk, B., and Fisler, R. (1995). Dissociation and the fragmentary nature of traumatic memories: overview and exploratory study. *J. Trauma. stress* 8 (4), 505–525. doi:10.1007/bf02102887

van Minnen, A., Arntz, A., and Keijsers, G. P. J. (2002). Prolonged exposure in patients with chronic PTSD: predictors of treatment outcome and dropout. *Behav. Res. Ther.* 40 (4), 439–457. doi:10.1016/s0005-7967(01)00024-9Apr

Van Veelen, N., Boonekamp, R. C., Schoonderwoerd, T. A. J., Van Emmerik, M. L., Nijdam, M. J., Bruinsma, B., et al. (2021). Tailored immersion: implementing personalized components into virtual reality for veterans with post-traumatic stress disorder. *Front. Virtual Real* 2, 740795. doi:10.3389/frvir.2021.740795Sep 30

Vermetten, E., and Bremner, J. D. (2003). Olfaction as a traumatic reminder in posttraumatic stress disorder: case reports and review. *J. Clin. Psychiatry* 64 (2), 202–207. doi:10.4088/jcp.v64n0214

Weathers, F. W., Blake, D. D., Schnurr, P. P., Kaloupek, D. G., Marx, B. P., and Keane, T. M. (2013). The clinician-administered PTSD scale for DSM-5 (CAPS-5). Available from: [Assessment] Available from

Wilkerson, A. K., Uhde, T. W., Leslie, K., Freeman, W. C., LaRowe, S. D., Schumann, A. Y., et al. (2018). Paradoxical olfactory function in combat Veterans: the role of PTSD and odor factors. *Mil. Psychol.* 30 (2), 120–130. doi:10.1080/08995605.2018.1425063

TYPE Systematic Review
PUBLISHED 24 April 2024
DOI 10.3389/frvir.2024.1252539

Check for updates

OPEN ACCESS

EDITED BY
Marta Matamala-Gomez,
University of Barcelona, Spain

REVIEWED BY
Chris Neale,
University of Huddersfield, United Kingdom
Kun Lyu,
Swiss Federal Institute of Technology Lausanne, Switzerland

*CORRESPONDENCE
Marilia K. S. Lopes,
marilia.soares@inrs.ca

RECEIVED 03 July 2023
ACCEPTED 08 April 2024
PUBLISHED 24 April 2024

CITATION
Lopes MKS and Falk TH (2024), Audio-visual-olfactory immersive digital nature exposure for stress and anxiety reduction: a systematic review on systems, outcomes, and challenges. *Front. Virtual Real.* 5:1252539.
doi: 10.3389/frvir.2024.1252539

Audio-visual-olfactory immersive digital nature exposure for stress and anxiety reduction: a systematic review on systems, outcomes, and challenges

Marilia K. S. Lopes* and Tiago H. Falk

Institut National de la Recherche Scientifique (INRS-EMT), University of Quebec, Montreal, QC, Canada

Evidence supporting the benefits of immersive virtual reality (VR) and exposure to nature for the wellbeing of individuals is steadily growing. So-called digital forest bathing experiences take advantage of the immersiveness of VR to make individuals feel like they are immersed in nature, which has led to documented improvements in mental health. The majority of existing studies have relied on conventional VR experiences, which stimulate only two senses: auditory and visual. However, the principle behind forest bathing is to have one stimulate all of their senses to be completely immersed in nature. As recent advances in olfactory technologies have emerged, multisensory immersive experiences that stimulate more than two senses may provide additional benefits. In this systematic literature review, we investigate the multisensory digital nature setups used and their psychological and psychophysiological outcomes; particular focus is placed on the inclusion of smells as the third sensory modality. We searched papers published between 2016 and April 2023 on PubMed, Science Direct, Web of Science, Scopus, Google Scholar, and IEEE Xplore. Results from our quality assessment revealed that the majority of studies (twelve) were of medium or high quality, while two were classified as low quality. Overall, the findings from the reviewed studies indicate a positive effect of including smells to digital nature experiences, with outcomes often comparable to conventional exposure to natural environments. The review concludes with a discussion of limitations observed in the examined studies and proposes recommendations for future research in this domain.

KEYWORDS

virtual reality, natural environment, olfactory stimuli, multisensory virtual reality, psychological outcome, psychophysiological outcome, stress/anxiety, relaxation

1 Introduction

In today's world, stress and anxiety have emerged as pervasive global problems, particularly in the wake of the COVID-19 pandemic, which further contributed to the crisis with employment instability, reduced socialization, and increased financial pressures. Several studies have reported an unprecedented surge in mental health disorders (Smith B. M. et al., 2020; Pietrabissa and Simpson, 2020; Millroth and Frey, 2021). For example, relative to the pre-COVID-19 era, there has been a global rise of 25.6% in anxiety and 27.6% in depression, highlighting the urgency and significance of addressing mental health issues

(Hawes et al., 2021; Santomauro et al., 2021). The detrimental impact of these psychological challenges on mental wellbeing and overall quality of life has prompted a growing need for effective interventions, especially non-pharmacological ones.

One non-pharmacological intervention that has shown great promise is to immerse oneself in nature, which has been shown to reduce stress/anxiety states and to provide different mental health benefits (Giannico et al., 2021; Kotera et al., 2021; Yao et al., 2021), as well as changes in physiological signals (Corazon et al., 2019; Fu et al., 2022). In Japan, for example, the practice of engaging in multisensory nature immersion has been embraced for many years and is known as 'forest bathing' (Shinrin-yoku) (Hansen et al., 2017), where one spends time in a forest to enhance health, wellness, and happiness. Despite these well-documented benefits, exposure in natural environments may not be accessible to many, such as individuals with limited mobility or hospitalized patients. The World Health Organization estimates that about 15% of the world's population faces disability, including limited mobility. Additionally, it is known that approximately 55% (in 2017) of the global population resides in urban areas, where access to natural surroundings can be challenging. Projections suggest that this percentage will rise to 68% by the year 2050 (Ritchie and Roser, 2018).

One promising solution that has emerged that provides more inclusive access to nature is that of immersive virtual reality (VR). Immersive VR experiences can simulate real-life environments, allowing users to feel present, and replacing their real surroundings with virtual reproductions. Numerous studies have explored the effects of VR in medical and healthcare (e.g. (Birckhead et al., 2019; Hao et al., 2022)). More recently, the concept of virtual nature bathing has also emerged. The approach is commonly referred to as 'digital Shinrin-yoku' (Reese et al., 2022) and aims to recreate the experience of nature exposure but in a virtual setting.

Virtual nature bathing experiences can vary widely and may involve 360-degree videos of nature environments projected onto head-mounted displays, panoramic photos shown on computer screens, or projections on walls, where users can feel immersed in the forest. As can be seen, these experiences primarily focus on visual and auditory stimuli, where nature scenes (e.g., forests, beaches, waterfalls, lavender fields) and sounds (e.g., birds chirping, leaves rustling, and water flowing) are present. Of course, the sense of immersion and presence are usually modulated by manipulating the quality of the visual (e.g., 8k-resolution high-fidelity 360-degree videos *versus* human-generated synthetic environments) and auditory (e.g., stereo *versus* 3D-audio) inputs.

Based on the principles of shinrin-yoku, however, the health benefits are believed to arise from the stimulation of all senses and complete immersion in nature. Since current VR systems offer only audio-visual stimuli, existing experiences may not be fully immersive and may not be providing maximal benefits to users. Indeed, research has revealed remarkable sensitivity of memory and the sense of presence to olfactory stimuli (Herz, 1998; Chu and Downes, 2002; Munyan et al., 2016), as well as to relaxation (Igarashi et al., 2014a; b; Amores et al., 2018). Smells have also shown to improve attention (Raudenbush et al., 2009; Dozio et al., 2021) to elicit basic emotions, such as happiness Vernet-Maury et al. (1999). In VR, it has also been shown to improve relaxation states (Carulli et al., 2019), to positively influence affective and behavioral reactions to the virtual environment (Flavián et al., 2021), to increase user engagement (Brengman et al., 2022) and sense of realism (Khan and Nilsson, 2023).

Given these insights, there is a compelling rationale to integrate olfactory stimuli in immersive nature exposure experiences, thus moving closer to the multi-sensory nature immersion aspect of Shinrin-yoku with an increased sense of presence, realism, and immersion. While recent studies have documented the impact of VR-based digital nature experiences on various psychological and physiological factors (e.g. (Jo et al., 2019; Syed Abdullah et al., 2021; Lee et al., 2022)), to the best of the author's knowledge, a review on multisensory digital nature exposure where smells are also included has yet to be published. Moreover, as digital nature exposure is aimed primarily at promoting relaxation, mindfulness, and connection with nature, existing reviews lack insights on how faithful the experiences are in achieving the immersion needed to reduce stress and anxiety. In this paper, we aim to fill these gaps. In particular, we explore studies that have investigated audio-visual-olfactory digital nature exposure for stress/anxiety reduction, as well as the use of objective measures to quantify the reduction. The review presents findings on the systems used as interventions, the psychological and physiological effects associated with these interventions, and the limitations that they bring.

2 Methodology

The present systematic review was conducted according to the preferred reporting items for systematic reviews and meta-analyses (PRISMA) Page et al. (2021) guidelines, as detailed in the sections below.

2.1 Search strategy

The research was conducted across six databases, including PubMed, Science Direct, Web of Science, Scopus, Google Scholar, and IEEE Xplore, between January 2016 to April 2023. The following keywords were used: (stress OR anxiety OR relaxation) AND ("virtual reality") AND ("nature scene" OR "digital nature" OR "digital forest" OR "virtual forest" OR "virtual nature" OR "forest bathing" OR "forest therapy" OR "nature therapy" OR "shinrin-yoku" OR "natural environment") AND (olfactory OR aromatherapy OR olfaction) AND (EEG OR biosignal OR physiological OR multisensory). Keywords were searched throughout the title, abstract, and full text of the articles. Although we have not included the term "psychological" as a keyword, we have kept articles that only used subjective measurements through questionnaires, so that we can analyze the systems that were used and quantify any limitations.

2.2 Inclusion and exclusion criteria

We included articles that 1) created a VR experience with or without head-mounted displays (HMD), i.e., such as screens or projectors; 2) included olfactory stimuli in the study; 3) included natural scenes as the virtual environment, e.g., forests, parks, and beaches; 4) included experiments with healthy patients; and 5) provided qualitative and/or quantitative outcome measures. In turn, exclusion criteria included: 1) review papers, books, or book chapters; 2) studies that did not use natural environments with a focus on relieving stress/anxiety and promoting relaxation; 3) studies with fewer than ten participants; 4) studies without experiments with human participants; and 5) studies that used games instead of nature scenes.

2.3 Screening and data extraction

After collecting articles from each database, the Zotero tool was used to remove duplicate articles. Titles and abstracts were then screened to exclude unrelated articles. For full-text screening, two steps were performed: first, the keywords related to olfactory were searched in the text, to check if the article would be kept or not. After removing studies that did not include the olfactory stimulus, the second step was then performed on the remaining papers using the inclusion/exclusion criteria.

To collect detailed information from each study, we extracted the following data from each article: study information (first author, year of publication), demographics (number of subjects, gender distribution, age), study design (control group, exclusion criteria, target population, target problem, experiment description, stimulated sense), intervention (VR device, collected physiological (when applied), olfactory equipment, duration), and outcomes (psychological measurement indicators, results extracted from physiological measurements and follow-up findings).

2.4 Quality assessment checklist

To assess the risk of bias (ROB) in the included studies, we employed the National Institutes of Health—National Heart, Lung, and Blood Institute (NIH-NHLBI) quality assessment tool NHLBI (2022). However, since this tool is primarily designed for clinical studies, which does not perfectly align with the nature of the studies included in this review, we also used a modified checklist based on (1) the Joanna Briggs Institute Critical Appraisal tool for use in JBI Systematic Reviews Tufanaru et al. (2020), namely, the Checklist for Quasi-Experimental Studies (Non-Randomized Experimental Studies), and (2) the Checklist for Randomized Controlled Trials Tufanaru et al. (2020).

The modified checklist comprised nine items, namely,: (1) Random assignment of conditions, (2) Existence of a control group, (3) Similar demographics across conditions, (4) Reliability of outcomes measurement, (5) Statistical analysis, (6) Multiple outcome measurement points, (7) Sample size, (8) Results comparison pre and post intervention/exposure, and (9) Clear study objective (for more details, see Supplementary Table S2.

Each item was assigned a score of one if it was clearly mentioned or addressed, 0.5 if it was partially mentioned or partially clear, and 0 if it was not mentioned or unclear. From the modified checklist, the maximum score achievable is 9. Therefore, studies with scores equal to or below half (4.5) were considered low quality, those between five and 6.5 were considered medium quality, and those with a score greater than or equal to seven were considered high quality.

3 Results and discussion

3.1 Included studies

A total of 1,143 articles were identified, 1,019 were from Google Scholar, 14 from IEEEXplore, 68 from Scopus, 41 from ScienceDirect, one from Web of Science, and none from PubMed. Figure 1 depicts a flow chart of the study selection process. After eliminating duplicates, 814 studies were left to be screened based on title and abstract. From this first pass, 660 papers were eliminated as they did not meet the inclusion criteria, resulting in 154 articles included for full-length analysis. Of these, only 12 articles were considered relevant for the current review. Additionally, two articles were included based on the analysis of the references cited in the originally selected 12 articles. In the end, 14 articles focusing on multisensory VR using audio-visual-olfactory stimuli, nature-based digital environments, and focused on relaxation or restoring stress/anxiety were included in this review.

3.2 Study characteristics

The overview of the selected studies is divided into five different tables, which focus on the provided information about the study population (Table 1), general characteristics of the reviewed studies (Table 2), experiment design (Table 3), main outcomes (Table 4), and the reported restorative effects (Table 5).

3.3 Quality assessment

The detailed responses obtained from both assessment tools can be found in the Supplementary Material. Supplementary Table S1 displays the responses for the NIH-NHLBI quality assessment tool, specifically for before-after (pre-post) studies without a control group. Additionally, Supplementary Table S3 presents the results of the quality evaluation using the modified checklist described in Section 2.4.

The findings suggest that of the 14 studies assessed, six were rated as high quality, six were rated as medium quality, and two were rated as low quality. Most of the studies categorized as medium quality demonstrated elements such as a counterbalanced study design, measurement of both objective and subjective outcomes, and similarity in terms of demographics. However, certain aspects, such as the inclusion of a control group/condition, were not consistently addressed. Overall, we found the quality criteria to be satisfactory. For a more comprehensive understanding of the quality criteria, please refer to Supplementary Table S2.

As can be seen from the Tables, all studies had virtual nature exposure conditions including olfactory stimuli. The virtual nature environment included mostly forests, urban green spaces, beaches, and parks. The range of exposure was characterized by a wide variability. For example, some studies only cited the total experience

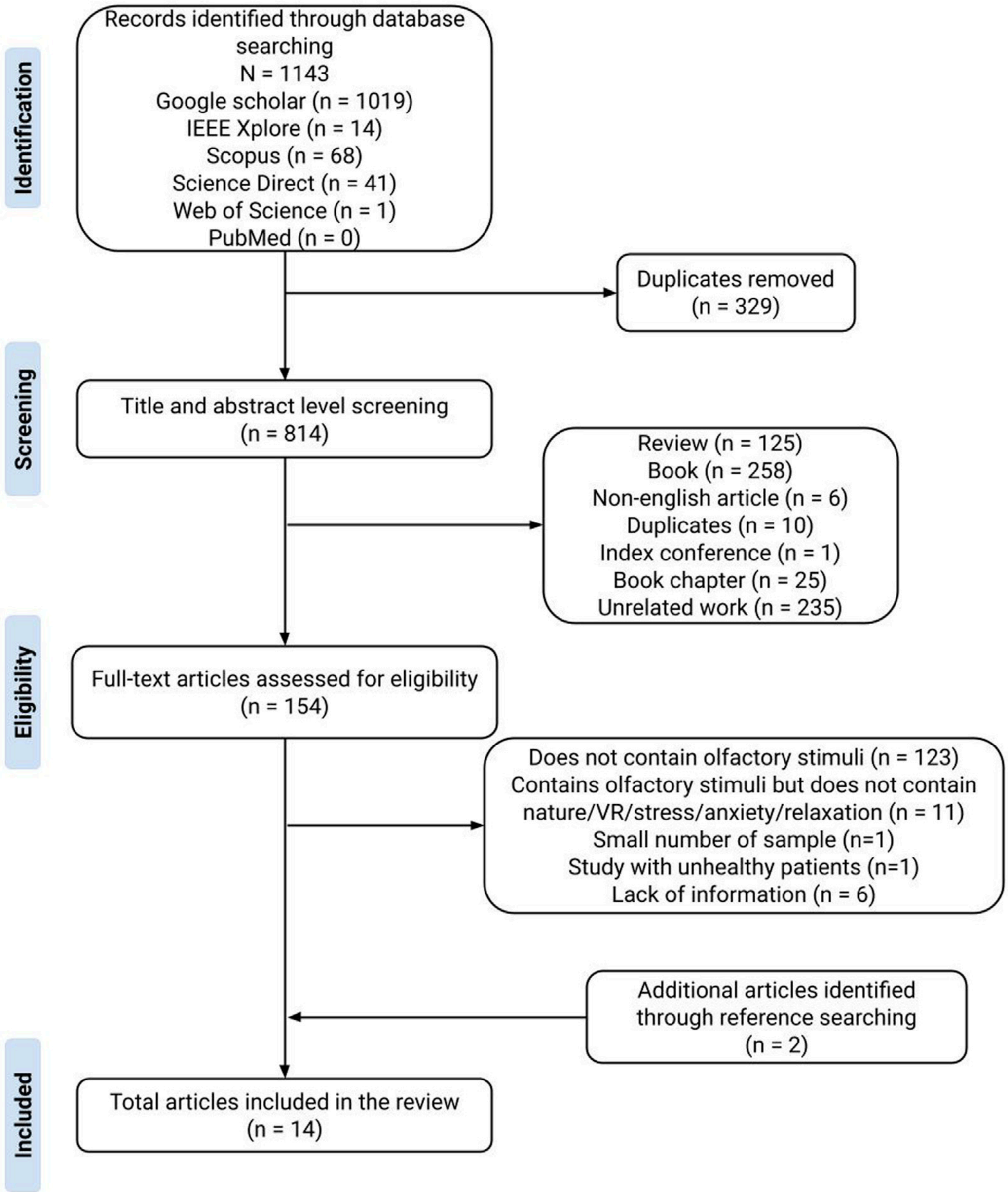


FIGURE 1
Flow chart of paper selection steps.

time, including the time to answer the questionnaires, and times ranged from a minimum of 5 min to a maximum of 1 h 30 min. Other studies, in turn, reported only the time of each condition, ranging from a minimum of 1 min to a maximum 60 min. Lastly, some reported both. In the next sections, more details are provided about the studies.

3.4 Multi-sensory stimuli and systems

3.4.1 Visual stimuli

Studies relied on different nature scenes for relaxation. In Zhang et al. (2023), for example, a panorama photo of Osmanthus trees (a common Chinese garden plant) was used for the intervention and a white wall without any smells as a group control. In Qi et al. (2022), four different environments were used, each presented as a three-dimensional 360-degree panoramic photo. The nature scenes included a lawn, a rose garden, an Osmanthus tree garden, and a pine forest. As in the other study, a white wall with no smell or audio was used for the control group. The work by Lopes et al. (2022) used a 360-degree computer-generated video of a forest environment; the study did not have a control group. The environment used in Abbott and Diaz-Artiles (2022) was a forested area including a campsite, a river, and a wood cabin; for the control condition, subjects were exposed to a stressful stimulus. In Schebella et al. (2020), a 360-degree video of low biodiversity, moderate biodiversity, and high

TABLE 1 Description of study population from the examined papers.

Study	Participants	N	Age range and/or mean ± SD	Female/Male
Zhang et al. (2023)	University students	95	18-26 years old	71/24
Shin et al. (2022)	University students	88	18-24 years old 20.0 ± 1.3	entire sample (F:54; M: 34); no window (F: 21; M: 13); closed window (F: 14; M: 13); open window (F: 19; M: 8)
Qi et al. (2022)	Undergraduate and graduate students	308	18-31 years old 22.92 ± 2.20	268/40
Lopes et al. (2022)	University students	16	27 ± 7.46	6/10
Takayama et al. (2022)	Members of the public	25	36.1 ± 8.13	13/12
Zhao et al. (2022)	University students and staff	38	27.7 ± 11.1	13/25
Abbott and Diaz-Artiles (2022)	University students	10	23.1 ± 2.8	3/7
Schebella et al. (2020)	University staff, students and members of the public	52	37.6 ± 10.6	28/24
Putrino et al. (2020)	Frontline healthcare workers	496	Not mentioned	Not mentioned
Song et al. (2019)	University students	21 female	21.1 ± 1.0	Only F
Hedblom et al. (2019)	Members of the public	154	city (mean age 27); park (mean age 28); forest (mean age 27)	city (F: 28; M: 22); park (F: 26; M: 26); forest (F: 28; M: 24)
Sona et al. (2019)	Students	122	22.69 ± 2.23	64/58
Amores et al. (2018)	Not mentioned	12	28.2 ± 4.4	4/8
Serrano et al. (2016)	Members of the public	136	27.05 ± 8.01	84/52

biodiversity nature scenes was used, along with a virtual urban environment as a control condition. In Hedblom et al. (2019), 360-degree photos of a park and forest were used, while for the control condition, an urban environment was used. Lastly, in Amores et al. (2018), a beach scene was presented in 360-degree video format. For the control condition, it was mentioned that no stimulus was presented and no specific instructions were given to the participants, other than to not close their eyes.

While the studies above used HMDs as the display modality, other studies relied on projectors and display screens/TVs to provide the virtual experience. In Takayama et al. (2022), a video image of a forest environment was projected in a room. In Shin et al. (2022), a 270-degree video of a busy cafe scene was presented under three conditions, two of which included a nature scene. In Zhao et al. (2022), an LED display screen was used to transmit a video of a forest in autumn colors; the control condition had no stimulus. In Putrino et al. (2020), a high-definition projector was used to show a soothing natural landscape, comprised of either ocean scenes or forest scenes. In Song et al. (2019), a 4K-compatible display television was used and the visual stimulus was a photograph view of a forest landscape. The work in Sona et al. (2019), utilized an artificial window comprised of three high-resolution 4K LED screens with speakers to present a park scene and a lounge scenario. In Serrano et al. (2016), an HD video projector was used where participants could watch relaxing scenes, including a beach and a landscape covered in snow. Lastly, the studies in Putrino et al. (2020); Sona et al. (2019) relied on a hybrid setting, where real plants were also used. More details about the virtual environment presentation modality can be seen in Section 3.5.

3.4.2 Auditory stimuli

Audio stimuli were used in most of the studies, with only two studies focusing on the visual-olfactory senses Zhang et al. (2023); Song et al. (2019). The audio was related to nature and often included bird and water sounds. Among the selected articles, only three studies used complementary sounds, in addition to nature ones: Lopes et al. (2022) included a guided self-reflection message presented in a young woman's voice; Putrino et al. (2020) and Serrano et al. (2016) also included relaxing music. For the studies that used the control condition with urban environments Hedblom et al. (2019); Schebella et al. (2020), it was common for traffic sounds to match the environment. In Sona et al. (2019), only instrumental music was played. Finally, the work in Shin et al. (2022) compared an open window (with virtual nature) condition *versus* a closed window (with/without nature) in a coffee shop. In the open window condition, sounds of nature could be heard. In all three conditions, background babble noise, typical of a coffee shop, was also presented at a moderate volume level.

3.4.3 Olfactory stimuli

All scents used in the studies were related to nature for the intervention conditions with the natural environments. Some relied on natural scents, such as the work in Zhang et al. (2023), which used real osmanthus flowers to produce the olfactory

TABLE 2 Description of the characteristics of the reviewed studies.

Study	Pre- and post-physiological measurement	Control group/ Control condition	Study aim
Zhang et al. (2023)	yes	yes	Expound the relationship between garden plant smell scapes and human health
Shin et al. (2022)	no	yes	To measure the restorative qualities of windows with nature views in a busy setting
Qi et al. (2022)	no	yes	Examine effects on physio-psychological restoration on birdsong and multi-sensory combination
Lopes et al. (2022)	no	no	Explored the effects of ultra-reality multisensory digital nature walks on user relaxation and compared subjective and objective findings relative to a conventional audio-visual VR experience
Takayama et al. (2022)	yes	no	Investigated the physiological and psychological therapeutic effects of a digital shinrin-yoku environment constructed indoors
Zhao et al. (2022)	no	no	Present a multisensory workstation to improve productivity, restorative quality, and wellbeing in the open-plan office
Abbott and Diaz-Artiles (2022)	no	no	A multisensory virtual reality environment as a potential tool to maintain the long-term behavioral health of astronauts
Schebella et al. (2020)	yes	yes	Multisenrory immersive natural virtual environment comparing three levels of biodiversity as well as one urban environment
Putrino et al. (2020)	no	no	A multisensory nature-inspired relaxation space to assist healthcare workers
Song et al. (2019)	yes	no	Investigate the physiological effects on brain activity and autonomic nervous activity of forest-related visual, olfactory, and combined visual and olfactory stimuli
Hedblom et al. (2019)	yes	yes	Understanding the link between physiological mechanisms and qualities of urban green spaces. Compare the effects of visual stimuli to the effects of olfactory stimuli and auditory stimuli on physiological stress recovery
Sona et al. (2019)	no	yes	Test the restorative potential of sensory-enriched break environments in a between-subjects with repeated measures design, focusing on the type of environment (natural outdoor vs. built indoor environment) and sensory input (no sensory input vs. audio-visual input vs. audiovisual and olfactory input)
Amores et al. (2018)	yes	no	A virtual reality experience that integrates a wearable, low-cost EEG headband and an olfactory necklace that passively promotes relaxation
Serrano et al. (2016)	no	no	To test the efficacy of a mood-induction procedure in a Virtual Reality environment for inducing relaxation and generating a sense of presence, and to test whether the stimulation of the senses of touch and smell improves the efficacy of this experience

stimulus, and the work in Qi et al. (2022) which used real samples of lawn, rose flowers, osmanthus flowers, and pine. In Song et al. (2019) they used Hinoki cypress leaf oil (Kiseitec Co., Wakayama, Japan) extracted by steam distillation from leaves and twigs of Hinoki cypress trees. On the other hand, eight studies used different types of diffusers to emit smells. In Takayama et al. (2022), the essential oil of Abies sachalinensis (Sakhalin fir) was used. The authors in Shin et al. (2022) used two ultrasonic essential oil diffusers with the aroma of coffee and wet dirt (nature condition). In Zhao et al. (2022), the Aroma Shooter device was used to provide a forest scent. In Abbott and Diaz-Artiles (2022), scents were produced with the Olorama Scent Generator (Olorama Technology Ltd, Spain) and smells of sausage, coffee, wet ground, pine, woods, roses, lavender, and honey were used. In Putrino et al. (2020), yuzu, hinoki, roman chamomile, and lavender essential oils were used. In Sona et al. (2019), scents of rosewood, geranium, ylang-ylang, olibanum (frankincense), and hyssop were used in the natural outdoor condition and composition of rosewood and cardamom in the indoor condition. In Serrano et al. (2016), a lavender scent was distributed with a ceramic diffuser. In Lopes et al. (2022) a SENSIKS (SENSIKS, Netherlands) multisensory pod was used with scents of dirt, forest, and grass.

Lastly, three studies used their methods of aroma emission. In Schebella et al. (2020), scents were applied to cotton pads and attached to the HMD, a range of 16 nature scents (e.g., several types of grass, soil, leaves, and flowers) and four urban scents (e.g., dust and turpentine) were initially tested on a convenience sample of 11 people. In the end, they used only one scent in the urban and low biodiversity scenes, and three different natural scents in the high

TABLE 3 Description of experimental design of reviewed papers.

Study	Conditions	Type of VR (device)	Virtual nature environment	Stimuli	Scent device	Scent used	Duration
Zhang et al. (2023)	Only olfactory vs. Only visual vs. Visual and Olfactory	Pico (but not mentioned)	Panorama photo of Osmanthus trees	visual olfactory	Small container	The flowers of Osmanthus fragrans var	34 min total. 3 min VR environment
Shin et al. (2022)	Brick wall view vs. closed window nature view vs. no open window nature view	Projector	VR simulation of a busy university cafe along with indoor sounds playing in the background and the scent of coffee. Birdsongs and dirt smells were added to the open window condition	visual auditory olfactory	Two ultrasonic essential oil diffusers	Cafe and the aroma of wet dirt	6 min VR environment
Qi et al. (2022)	Only birdsong vs. birdsong + photo (4 types) vs. birdsong + odor (4 types) vs. and birdsong + photo + odor (4 types)	Pico Goblin VR	3D 360-degree panoramic photo: Lawn, rose garden, osmanthus tree garden, and pine forest	visual auditory olfactory	Small container	Leaves from the lawn, flowers of rose bushes, flowers of osmanthus trees and leaves (pine needles) of pine trees	25 min total. 2 min VR environment
Lopes et al. (2022)	Audio-visual only vs. ultra-sensory	Oculus Quest	360-degree video, forest environment	visual auditory olfactory tactile	SENSIKS	Dirt, forest and grass	1min15seg VR environment
Takayama et al. (2022)	Digital forest	Projector	Video images taken in Urahoro-cho, Hokkaido, on the three walls and ceiling of the experimental room	visual auditory olfactory	Diffuser	Essential oil of Abies sachalinensis	1hr30min total. 20 min VR environment
Zhao et al. (2022)	Neutral vs. Forest without scent vs. Forest with scent	LED display (screen)	Video of a forest in colorful autumn color	visual auditory olfactory tactile	Aroma Shooter from aromajoin	Forest scent mixture	1hr5min total. Tutorial of an example session and several cognitive tasks (approx. 20 min). Experiment session (approx. 45 min)
Abbott and Diaz-Artiles (2022)	VR condition without scents vs. VR condition with scents	HTC Vive Pro	Nature inspired VR environment with localized scents and sounds featuring a rushing river, campground, and small cottage	visual auditory olfactory	Olorama Scent Generator (OSG)	Sausage, coffee, wet ground, pine, woods, roses, lavender, honey	Stress stimuli (8 min each). Nature VR environment (15 min each condition)
Schebella et al. (2020)	Urban vs. low biodiversity vs. moderate biodiversity vs. high biodiversity vs. high biodiversity (visual-only)	Oculus Rift	360-degree video of urban, low biodiversity, moderate biodiversity, and high biodiversity	visual auditory olfactory	Cotton pads attached to the head-mounted display	Not specified. The urban and low biodiversity: single scent; high biodiversity: three different natural scents	15 min baseline; TSST total 15 min; 5 min VR environment
Putrino et al. (2020)	Recharge Room	HD projector	Natural scenes (not specified in the paper)	visual auditory olfactory	Diffuser	Yuzu, hinoki, roman chamomile, and lavender essential oils calming scents	15 min experiences
Song et al. (2019)	Visual stimulus vs. olfactory stimulus vs. combined stimulus	4K display television	Forest landscape of Hinoki cypress trees	visual olfactory	Unspecified device	Hinoki cypress leaf essential oil	90s each condition
Hedblom et al. (2019)	Urban vs. park vs. forest	Oculus Rift	2D 360° Virtual Reality photos of urban area in Stockholm; park in Uppsala; city forest in Uppsala	visual auditory olfactory	Computer-controlled olfactometer	Urban: diesel, tar, and gunpowder; Park: grass; Forest: two evergreen species and mushroom	5 min 30s total

(Continued on following page)

TABLE 3 (*Continued*) Description of experimental design of reviewed papers.

Study	Conditions	Type of VR (device)	Virtual nature environment	Stimuli	Scent device	Scent used	Duration
Sona et al. (2019)	No sensory input vs. audiovisual input vs. audiovisual and olfactory input	An artificial window LED screens with speakers	Park and lounge scenery	visual auditory olfactory	Diffuser	Rosewood, geranium, ylang-ylang, olibanum (frankincense) and hyssop in the natural outdoor condition. Rosewood and cardamom in the built indoor condition	65 min
Amores et al. (2018)	No stimulus vs. VR + Scent + Audio	Samsung Gear VR	360° beach video	visual auditory olfactory	Essence necklace	Lavender essential oil	5 min each condition
Serrano et al. (2016)	VR vs. VR + Smell vs. VR + Touch vs. VR + Touch + Smell	HD video-projector	House of Relaxation: affective pictures, a video-clip with a relaxing scene, relaxing music and narratives, sounds of nature, and autobiographical recalls	visual auditory olfactory tactile	Ceramic diffuser	Lavender essential oil	60 min for each condition

biodiversity scene, however, they did not specify the exact name of the aroma. All scents were selected from the Demeter Fragrance Library[1]. In Hedblom et al. (2019), the authors used their own computer-controlled olfactometer (Lundström et al., 2010), where city odors included diesel, tar, and gunpowder and for park odors they used grass and different forest odors, including two evergreen species and mushroom. In Amores et al. (2018), an essence necklace was used to emit a lavender essential oil scent.

The methodology for selecting scents in multisensory VR studies varies, often reflecting the unique objectives and contexts of each research project. For instance, Zhang et al. (2023) selected the scent of Osmanthus, a flower renowned as one of the ten most famous flowers in China, where the study took place, thus highlighting the cultural relevance for scent choice. Takayama et al. (2022); Song et al. (2019) focused on scents from trees that complemented the visual aspects of the VR environment, thus enhancing the sensory congruence for users. Other studies, however, such as that conducted by Putrino et al. (2020); Amores et al. (2018); Serrano et al. (2016), opted for scents such as lavender and Roman chamomile essential oils, chosen for their documented effects in reducing anxiety and stress in the scientific literature.

Qi et al. (2022) took it one step further and used a portable odor sensor (COSMOS XP-329 III R; Rank Value, range 0–2000) to measure and select odor concentrations that closely mimicked the landscapes used in the virtual environments. Schebella et al. (2020) engaged participants in the scent selection process, where 11 subjects identified which scents they found to more closely match the natural or urban settings in the experiment. Seven scents were accurately recognized and subsequently utilized in corresponding VR environments, ensuring the olfactory experience matched the visual setting. Moreover, Sona et al. (2019) not only referenced literature for scent selection but also conducted preliminary tests with 12 subjects to determine the intensity of ambient scents and identify perception thresholds. Other studies Shin et al. (2022); Lopes et al. (2022); Zhao et al. (2022); Abbott and Diaz-Artiles (2022); Hedblom et al. (2019), chose scents based on the landscapes in their experiments.

3.4.4 Tactile and somatosensory stimuli

In addition to the audio-visual-olfactory stimuli detailed above, three studies also investigated tactile/somatosensory stimuli. In Lopes et al. (2022), heating elements and four fans surrounding the sides, back, and front of the users, as well as acoustic vibrations on the user's seat were used. In Zhao et al. (2022), fans were mounted on the top of the table pointing toward the user's head and table surface, as well as infrared lights underneath the table were used for temperature control. Lastly, in Serrano et al. (2016), artificial grass was installed on the ground, allowing the participant to touch and feel it during the immersive VR experience.

3.5 Presentation modality and virtual nature experiences

3.5.1 Display/projector-based experiences

Among the selected studies, seven presented the virtual reality experience using display/projectors in an indoor setting (Serrano et al., 2016; Hedblom et al., 2019; Sona et al., 2019; Putrino et al., 2020; Shin et al., 2022; Takayama et al., 2022; Zhao et al., 2022). In Shin et al. (2022), three conditions were tested: no window (brick wall), closed window (nature view), and open window (nature view) with all conditions including the scene of a coffee shop. Participants were randomized into one of three groups. Participants were immersed in a 270-degree VR simulation within the laboratory. The participant's viewpoint featured a lifelike projection of the student union coffee shop, spanning the front and left walls. Additionally, the right wall showcased three distinct conditions. Throughout all conditions, a coffee fragrance and ambient

1 http://www.demeterfragrance.com/

TABLE 4 Description of reported outcomes of reviewed studies.

Study	Outcome assessment (psychological)	Outcome assessment (physiological)	Main results
Zhang et al. (2023)	None	Blood pressure; pulse pressure difference; pulse; EDA; EEG	The research suggests that integrating olfactory and visual stimuli from garden plants can enhance physiological relaxation and overall health more effectively than experiencing these stimuli in isolation
Shin et al. (2022)	PRS; and ROS	None	The study found that adding sounds and smells to virtual nature windows did not significantly enhance their restorative effects compared to experiences without these stimuli. However, it was observed that virtual windows showcasing nature views can still offer substantial restorative benefits in busy indoor environments
Qi et al. (2022)	STADI-S	ST; EDA; EEG	Audio-Visual (using birdsong) stimuli enhanced physiological restoration and perceived quality without affecting psychology, while olfactory stimuli adversely impacted physiological restoration but increased landscape attraction and harmony; combining both stimuli improved physiological aspects (noted in beta-EEG) and perceived quality, yet still had no psychological impact
Lopes et al. (2022)	Relaxation state from 1 (“Not relaxed at all”) to 5 (“Very much relaxed”)	ECG; ST; EDA; BVP; breathing signals	Significant changes in relaxation were achieved with the proposed system and changes in physiological parameters were also observed
Takayama et al. (2022)	POMS; PANAS; ROS; PRS	maximum and minimum blood pressure; pulse rate; salivary amylase activity; HRVHR.	During exposure, notable changes in physiological states were observed compared to the resting state, accompanied by significantly positive changes in psychological states
Zhao et al. (2022)	Questionnaire (a set of survey questions including the PRS)	HRV	Forest environments (with scent and without scent), were perceived more positively than a neutral condition, but the unscented forest condition showed slower stress development and better recovery
Abbott and Diaz-Artiles (2022)	PANAS; 6-item (STAI-6)	None	After the VR experience with added scents, participants showed significantly lower stress levels compared to the control condition, with a notable decrease in negative emotions and anxiety from both the stressed state and baseline when olfactory stimuli were introduced
Schebella et al. (2020)	VAS Biodiversity Experience Index (BEI). Kim and Biocca’s telepresence	EDA; HR. But EDA data were not included in the analysis	In low biodiversity environments with multisensory stimulation, there was a notable improvement in all wellbeing measures and reduced anxiety, with multisensory experiences leading to better overall wellbeing recovery compared to visual-only experiences
Putrino et al. (2020)	Single-item Likert-style measure of perceived stress	None	After the experience the stress level was significantly reduced
Song et al. (2019)	Questionnaire comprising 4-opposite adjectives	Oxyhemoglobin concentration using near-infrared time-resolved spectroscopy; HRV; HR	The study revealed that combined visual and olfactory stimuli significantly altered physiological responses and enhanced feelings of comfort, relaxation, and a sense of naturalness and realism among participants
Hedblom et al. (2019)	Subjective stress sensibility (4-point Likert); Pleasantness of the landscape (from 1 to 100)	EDA	The study revealed that park and forest environments significantly lowered stress levels, in contrast to urban areas which induced higher stress. Furthermore, it was observed that high pleasantness ratings were associated with reduced physiological stress responses, notably in response to olfactory and, to a lesser extent, auditory stimuli, but not in response to visual stimuli

(Continued on following page)

TABLE 4 (*Continued*) Description of reported outcomes of reviewed studies.

Study	Outcome assessment (psychological)	Outcome assessment (physiological)	Main results
Sona et al. (2019)	Rating (1: pleasant to 7: unpleasant); PRS; 6-point Likert scale (1 = little to 6 = extremely). fatigue, mood, and arousal	None	The study found that both natural and lounge environments were perceived as more pleasant and restorative than a standard break room, indirectly facilitating the recovery of personal resources like mood, fatigue, and arousal, especially when a scent was added to an audiovisual simulation
Amores et al. (2018)	RRS 1 (“Not relaxed at all”) to 7 (Totally relaxed)	EEG (Relax score based on the Renyi entropy)	Subjective perception of relaxation increased when using a VR headset with the olfactory necklace, compared to not being exposed to any stimulus
Serrano et al. (2016)	Clinical Assessment Questionnaire; Beck Depression Inventory II; STADI-S; VAS; SAM; Presence SAM	None	The stimulation of the senses of touch and smell did not show a significant improvement of the mood- induction or the sense of presence

Note: Electrodermal activity (EDA); Electroencephalography (EEG); Electrocardiogram (ECG); skin temperature (ST); blood volume pulse (BVP); Heart Rate Variability (HRV); Heart rate (HR).

conversation noises were used at a moderate volume level. In the open window condition, the immersive experience was further enhanced by the sounds of birds chirping/singing in the background, accompanied by the aroma of wet dirt. In Takayama et al. (2022), participants were exposed to a 20-min digital shinrin-yoku environment that reproduced visual, auditory, and olfactory elements, each session accommodated a maximum of five subjects, and an adequate distance between individuals was maintained. Video images of a forest were projected onto three walls and the ceiling of the experimental room.

The work in Zhao et al. (2022) focused on multimodal augmentation via lighting, audio, video, airflow, heating, and scent elements to improve the productivity and wellbeing of workers in an open-plan office. In this work, a comparison was made between three conditions (neutral vs. unscented forest vs. scented forest), and a stress induction phase was also performed. For the neutral condition, the LED panels displayed no visuals, whereas for the other two conditions (with scent and without scent) a scene forest was displayed. Another study created a *recharge room* with multisensory (visual, auditory, and olfactory) nature-inspired experiences where healthcare workers could relax and reduce stress Putrino et al. (2020). The environment was carefully designed to evoke a serene natural setting. Silk imitation plants were strategically placed to simulate lush greenery, while scenes of natural landscapes were projected onto a blank wall. The ambiance was further enhanced by soft, color-matched lighting, which complemented the projected landscapes. High-definition audio recordings of nature sounds were combined with relaxing music. To complete the multisensory immersion, essential oils, and calming scents were diffused using an essential oil diffuser, heightening the impression of being fully surrounded by a natural environment.

The study in Song et al. (2019) employed four different conditions, each lasting 90 s, namely,: a forest landscape image featuring Hinoki cypress trees without any accompanying odor (visual-only), a gray image infused with Hinoki cypress leaf essential oil (olfactory-only), a forest landscape image of Hinoki cypress trees combined with the scent of Hinoki cypress leaf essential oil (combined stimulus), and a gray image without any odor (control condition). The study conducted by Sona et al. (2019) examined five distinct conditions, including a control condition with no window, sound, or scent. Additionally, two audio-visual conditions were assessed: a nature condition with a window displaying nature scenes, accompanied by bird sounds and neutralizing air/no scent, and a lounge condition featuring a window displaying a lounge scene, instrumental music, and neutralizing air/no scent. Furthermore, the study explored two audio-visual-olfactory conditions where scents were introduced as an additional element.

In Serrano et al. (2016), the experiment was performed in a VR-room, in which a virtual reality environment was created where participants were immersed inside a two-story house called *House of Relaxation*; the goal was to induce relaxation and a sense of presence and to analyze whether adding touch and smell altered the experience outcomes. Inside the environment, users could perform different actions, such as modifying the house's decoration, watching a video clip through a projector, changing the intensity of the lights and the color of the walls, opening or closing doors and windows, moving between the two floors of the house, moving or adding furniture, changing the ground to grass or sand, or change the exterior landscape to reflect another natural context (e.g., beach or snow). The lavender scent was used to stimulate the olfactory sense, and for the tactile stimulus, they used artificial grass. They compared four conditions: audio-visual, audio-visual-olfactory, audio-visual-tactile, and audio-visual-olfactory-tactile.

3.5.2 HMD-based experiences

The other seven studies relied on VR HMDs to present the visual stimuli (Amores et al., 2018; Hedblom et al., 2019; Schebella et al., 2020; Abbott and Diaz-Artiles, 2022; Lopes et al., 2022; Qi et al., 2022; Zhang et al., 2023). The most recent study compared three conditions, i.e., olfactory-only, visual-only, and visual-olfactory) to a control condition Zhang et al. (2023). The experiment took place in a room with white walls and ceilings, ensuring a consistent indoor environment to prevent the subjects' moods from being influenced by external factors, such as weather and light. In the control group, an odorless condition was maintained, with participants exposed to an image of a white wall, devoid of any stimuli associated with plant smells or garden landscapes. In the olfactory condition, participants experienced a natural concentration of fragrance emitted from Osmanthus fragrans. For the visual stimulus, a panoramic photo of Osmanthus trees was displayed. The visual-olfactory condition involved the combination of both visual and olfactory stimuli.

In Qi et al. (2022), a total of 308 volunteers participated in the study and were divided into 14 independent groups. These groups were designed to explore three different stimulus scenarios: audio-visual, audio-olfactory, and audio-visual-olfactory. Four plant landscapes (lawn, rose garden, osmanthus tree garden, and pine forest), with each landscape accompanied by its respective scent when olfactory stimuli were integrated, plus a single birdsong group (1 group) that visualized a white wall and received no odor stimulus, and a control group (1 group) of participants who viewed only a white wall (no birdsong and no odor). Each condition comprised 22 participants, ensuring an equal distribution across the study.

In Lopes et al. (2022), two conditions were compared: audio-visual and audio-visual-olfactory-tactile. Participants were seated inside a SENSIKS multisensory booth, which provided a fully immersive experience. In the audio-visual condition, participants were immersed in a forest environment for approximately 1 minute, accompanied by the soothing sounds of nature. Additionally, a soft voice of a young woman whispered in the background, prompting participants to engage in self-reflection on existence and consciousness. In the audio-visual-olfactory-tactile condition, scents of nature, such as dirt, forest, and grass were included, while fans and heating elements located on the sides, back, and front of the participant were used to simulate wind and sunlight matching the visual content.

Another study that showed a positive impact of introducing olfactory stimuli to reduce negative affect and state anxiety levels was the work in Abbott and Diaz-Artiles (2022). There, they analyzed an audio-visual and an audio-visual-olfactory condition with a task to induce stress between conditions. They built two VR environments using Unreal Engine 4. The first VR environment was designed as a nature-inspired forested area, complete with a campsite, a river, and a wood cabin. To enhance the immersion, realistic sounds of nature such as a rushing river and chirping birds were incorporated. Additionally, when olfactory stimuli were required, scents of nature were introduced. The second VR environment was specifically tailored for the stress induction portion of the experiment, known as the Trier Social Stress Test (TSST). This environment was modeled after a small auditorium, featuring approximately 30 static audience members. These virtual audience members remained passive throughout the session, providing no feedback to the subjects, whether positive or negative. Lastly, in Amores et al. (2018), a baseline condition (no stimulus) was compared to an audio-visual-olfactory condition. The participant was comfortably seated while wearing the Essence necklace, which emitted a fragrance of lavender. Participants were presented with a 360-degree video of a beach setting and the audio accompanying the video.

3.5.3 Nature vs. urban virtual environments

Two studies compared between a natural and an urban environment (Hedblom et al., 2019; Schebella et al., 2020). In Schebella et al. (2020), three natural environments, with different levels of biodiversity (low, moderate, and high) were compared to an urban environment control condition. A significant effect of multisensory biodiversity was observed on stress recovery related to the urban environment. Moreover, in Hedblom et al. (2019) an urban area was compared to an urban park and an urban forest. Each condition included audio-visual-olfactory stimuli appropriate for the scenes. The two virtual nature conditions resulted in greater stress reduction relative to the urban environment. The perceived pleasantness was associated with the olfactory and auditory stimuli, but not with the visual stimulus. The study suggested that olfactory stimuli may be better at facilitating stress reduction than visual stimuli and highlights the relevance of multisensory approaches for stress reduction.

3.6 Outcome assessment (psychological)

3.6.1 Subjective measures

In total, 13 studies performed subjective measurements to analyze the mood changes associated with the multisensory VR exposure. Four studies administered the Perceived Restorativeness Scale (PRS) to measure how restorative an environment was (Sona et al., 2019; Shin et al., 2022; Takayama et al., 2022; Zhao et al., 2022). In Sona et al. (2019), questions about mood, fatigue, and arousal (before and after exposure) were also measured. Two studies used the Restoration Outcomes Scale (ROS) to analyze the subjective restorativeness, where Shin et al. (2022) requested the questionnaire only after VR exposure, while in Takayama et al. (2022) the measurement was requested before and after exposure. Two other studies used a semantic differential (SD) method. In Qi et al. (2022), a survey concerning the overall quality evaluation of the environment was used, while in Song et al. (2019) the participant's subjective spatial impressions were collected via a questionnaire with four opposite adjectives, each evaluated on 13 scales.

Two studies measured mood with the Positive and Negative Affect Schedule (PANAS) (Abbott and Diaz-Artiles, 2022; Takayama et al., 2022). Two studies used the Visual Analogue Scale (VAS) before and after exposure. While in Schebella et al. (2020) VAS was used to measure the perceived stress, anxiety, insecurity, calmness, and happiness levels of the participants, in Serrano et al. (2016) it was used to evaluate levels of sadness, joy, anxiety, and relaxation moods. One study used the Profile of Mood States (Takayama et al., 2022). The State-Trait Anxiety Inventory (STAI) was used in two studies. In Qi et al. (2022), the full 20-item version was used, whereas in Abbott and Diaz-Artiles (2022) a modified version including only six-items was used. The work in Serrano et al. (2016) mentioned using the state scale of the STAI but did not report the results; instead, it reported results using the Self-Assessment Manikin (SAM) (before and after exposure) and the Presence Self-Assessment Manikin (after exposure).

Single-item measures were also used across different studies, including the Relaxation Rating Scale (RRS) which was collected in Amores et al. (2018) before and after the exposure, and in Lopes et al. (2022) just after exposure. Overall, five studies acquired some type of subjective measure only after the virtual experience (Hedblom et al., 2019; Song et al., 2019; Lopes et al., 2022; Shin et al., 2022; Zhao et al., 2022). Moreover, five studies used tasks to induce stress prior to exposure. In Zhao et al. (2022), participants were given a 3-min

TABLE 5 Summary of restorative effects reported in the reviewed studies.

Study	Restorative effects	Outcomes
Zhang et al. (2023)	Not measured	-
Shin et al. (2022)	Yes	Results of one-way ANOVA showed significant restorativeness/restoration across conditions for the ROS subscales 'Being Away' and 'Fascination' ($p < 0.05$)
Qi et al. (2022)	Yes	No significant effect on psychological restoration
Lopes et al. (2022)	Not measured	-
Takayama et al. (2022)	Yes	Psychological restorative effects were confirmed, with a significant decrease ($p < 0.01$) in "negative affect" (measured using PANAS) and a significant increase ($p < 0.01$) in the sense of restoration (measured using ROS) after the experience
Zhao et al. (2022)	Yes	In virtual forest conditions (with and without nature smell), ratings of perceived restoration and focus were significantly higher than the neutral condition ($p < 0.05$, Wilcoxon signed-rank test). In addition, there was a significant difference in restoration perception among the three conditions ($p < 0.05$, Friedman's test)
Abbott and Diaz-Artiles (2022)	Yes	Participants reported significantly lower STAI-6 scores after the audio-visual-olfactory than in the audio-visual condition ($p = 0.03$). Compared to the stressed state, PANAS Negative Affect ($p = 0.003$) and STAI-6 ($p = 0.001$) scores decreased after the audio-visual-olfactory condition. STAI-6 scores ($p = 0.013$) also decreased from baseline in audio-visual-olfactory condition
Schebella et al. (2020)	Yes	Low biodiversity natural environment was the most restorative during recovery from induced stress (TSST-IVE). Natural environments were more restorative than urban environments. Among the natural environments, moderate biodiversity immersive virtual environment was the least restorative
Putrino et al. (2020)	Not measured	-
Song et al. (2019)	Not measured	-
Hedblom et al. (2019)	Not measured	-
Sona et al. (2019)	Yes	The experience with nature view was perceived as more restorative (r = 39, $p < 0.01$) than the control group (no stimulus) and the lounge simulations (r = 0.22, $p < 0.05$)
Amores et al. (2018)	Not measured	-
Serrano et al. (2016)	Not measured	-

reading comprehension task similar to those given on a graduate standardized test. In Abbott and Diaz-Artiles (2022) and Schebella et al. (2020), the modified trier social stress test (TSST) was used. In Hedblom et al. (2019), mild electric shocks were used as a stress trigger. Finally, in Sona et al. (2019), a sequence of tasks was used to induce stress, including a single N-back task, a Stroop task, and an Attention Network Task.

3.6.2 Emotional restoration, mood, and perceived stress

The addition of olfactory stimuli in virtual nature experiences has resulted in inconsistent outcomes concerning effects on emotional restoration and positive/negative emotion. Table 5 shows a summary concerning the restorative effects mentioned in the studies. While some studies have reported an improvement in mood and stress reduction, others have reported no significant differences. For example, in Shin et al. (2022), the conditions where the participants looked out the window at nature showed greater restorative qualities than the absence of nature, but comparisons between the closed window condition (only view of nature) and the open window condition (sight, smell, and sounds of nature) showed no significant differences. Several participants did, however, report the open window condition to be more captivating. Results of one-way ANOVAs showed significant restorativeness/restoration across conditions for the ROS subscales 'Being Away' and 'Fascination' ($p < 0.05$).

In Qi et al. (2022), the audio-visual-olfactory stimuli led to an increased physiological restoration and overall perceived quality, whereas the audio-olfactory stimuli, despite having no significant effect on psychological restoration, showed an increase in perceived overall feelings of attraction to the landscape and a sense of overall harmony. For the audio-visual and audio-visual-olfactory conditions, no significant differences in psychological effects were found. Only for the auditory stimulus was the STAI-S value significantly lower than the control group ($p = 0.04$). In another study Zhao et al. (2022), in turn, after the two virtual forest conditions (with and without nature smell), ratings of perceived restoration and focus were significantly $p < 0.05$ higher than the neutral condition (i.e., post-hoc comparisons, using a Wilcoxon signed-rank test, audio-visual-olfactory, Z = 17.50, $p = 0.000$ and audio-visual Z = 13.0, $p = 0.000$). However, stress development was slower and recovery was greater in the forest without scent condition compared to the other conditions. In addition, Friedman's test showed a significant ($p < 0.05$) difference in restoration perception among the three conditions ($X^2(36) = 39.78$, $p = 0.000$). Also, in Takayama et al. (2022) negative mood states (measured using POMS), significantly decreased ($p < 0.01$) after the experience (i.e., tension-anxiety, depression, anger-hostility, fatigue, and confusion). In addition, psychological restorative effects were confirmed, with a significant decrease ($p < 0.01$) in "negative affect" (measured using PANAS) and a significant increase ($p < 0.01$) in the sense of restoration (measured using ROS) after the experience.

Regarding negative emotions, in Abbott and Diaz-Artiles (2022), participants experienced a notable reduction in STAI-6 scores following the audio-visual-olfactory condition compared to the audio-visual condition ($p = 0.03$). Furthermore, when contrasted with their stressed state, there was a significant decrease in both PANAS Negative Affect ($p = 0.003$) and STAI-6 scores ($p = 0.001$) after the audio-visual-olfactory. Additionally, STAI-6 scores showed a significant decline from the baseline in the audio-visual-olfactory condition ($p = 0.013$). Results suggested that the addition of olfactory stimuli to the VR environment aided in reducing negative affect and state anxiety levels.

For positive emotions, significant improvements in relaxation were reported in Lopes et al. (2022); Amores et al. (2018). In Amores et al. (2018), the results showed that the multisensory VR experience was able to promote increased relaxation relative to the baseline condition, although not a significant difference. The findings corroborate previous studies that showed forest-related stimuli significantly increasing the participants' feelings of comfort and relaxation relative to a control condition ($p < 0.05$) Song et al. (2019). Similar findings were shown in Lopes et al. (2022) where greater relaxation was achieved as more senses were stimulated ($p < 0.05$). Moreover, the work in Sona et al. (2019) the correlation between the two multisensory environment experiences showed that nature views were perceived as more pleasant (r = 0.74, $p < 0.01$) and restorative (r = 39, $p < 0.01$) than the control group (no stimulus). Adding nature smells facilitated the recovery of personal resources (mood, fatigue, arousal) via greater scent pleasantness and fascination (r = 0.18, $p < 0.10$). Moreover, correlation analyses between experiences showed that the view was perceived as more pleasant (r = 0.53, $p < 0.01$) and the environment as more restorative (r = 0.22, $p < 0.05$) in the nature simulations than in the lounge simulations. In Serrano et al. (2016), the results showed that the VR experience was effective in inducing relaxation F(1,132) = 90.31, $p < 0.001$, where positive moods and the sense of presence increased. Moreover, a significant decrease in arousal (F(1,132) = 92.04, $p < 0.001$) was observed after all four VR experiences. However, no statistical differences were found between the four groups on emotions and sense of presence. In Hedblom et al. (2019), the perceived pleasantness, average over environments and sensory stimuli rated on a 1–100 scale, was significantly higher in the park environment (Mean = 69.21, SD = ±11.1) compared to both the forest (Mean = 62.7, SD = ±12.74; $p = 0.01$) and the urban area ($p < 0.001$). Perceived pleasantness was also higher for the forest relative to the urban area (Mean = 37.19, SD = ±14.22; $p < 0.001$). There were significant differences in the ratings of perceived pleasantness among the three environments, as indicated by the sensory stimuli (t = 90.01; $p < 0.001$; pairwise comparisons).

Regarding perceived stress, Schebella et al. (2020) showed that a multisensory experience (visual, auditory, and olfactory) was associated with better recovery in all measures of wellbeing relative to a visual-only experience. Median anxiety recovery scores were significantly greater in the audio-visual-olfactory high biodiversity environment (33.00) than in the visual-only (17.50) using a Mann–Whitney test ($U = 90.50$, $z = 1.971$, $p = 0.047$). When compared to the virtual urban environment, stress recovery was most effective in a low-biodiversity environment ($p < 0.05$). In Putrino et al. (2020), a greater reduction in perceived stress after the multisensory immersive experience was reported ($p < 0.001$), where the level of stress had an average reduction of 59.6% in the self-reported ratings. Similarly, in Hedblom et al. (2019), the park and forest conditions resulted in significant stress reduction, whereas the urban area condition did not.

3.7 Outcome assessment (physiological)

In total, nine articles performed some type of physiological measurement and assessment (Amores et al., 2018; Hedblom et al., 2019; Song et al., 2019; Schebella et al., 2020; Lopes et al., 2022; Qi et al., 2022; Takayama et al., 2022; Zhao et al., 2022; Zhang et al., 2023). Only three of them, however, measured physiological stress through stress induction (Hedblom et al., 2019; Schebella et al., 2020; Zhao et al., 2022). In particular, five studies measured cardiovascular data (Song et al., 2019; Schebella et al., 2020; Lopes et al., 2022; Takayama et al., 2022; Zhao et al., 2022), two measured blood pressure and pulse Zhang et al. (2023); Takayama et al. (2022), one measured salivary amylase activity (Takayama et al., 2022), four assessed electrodermal activity (EDA) (Hedblom et al., 2019; Lopes et al., 2022; Qi et al., 2022; Zhang et al., 2023), one measured skin temperature (Qi et al., 2022), one measured oxy-hemoglobin concentration in left and right prefrontal cortices as an indicator of brain activity using near-infrared spectroscopy (Song et al., 2019), and three used electroencephalography (EEG) to measure neural activity (Amores et al., 2018; Qi et al., 2022; Zhang et al., 2023). More details about each modality are presented next.

3.7.1 Blood pressure and pulse

The study described in Zhang et al. (2023) investigated the effects of olfactory and olfactory-visual stimulation on blood pressure and pulse. During the stimulation, it was observed that systolic blood pressure (SBP) values remained relatively unchanged. However, diastolic blood pressure (DBP) values exhibited a significant increase ($\Delta DBP = 4.37 \pm 1.69$ mmHg, $p < 0.05$). Additionally, the pulse pressure difference (PP) values demonstrated a notable decrease during both stimulations ($\Delta PP = -4.56 \pm 1.24$ mmHg, $p < 0.05$). Furthermore, during olfactory stimulation, pulse values also showed a significant decrease ($\Delta P = -2.34 \pm 1.16$ bmp, $p < 0.05$). In contrast, visual stimulation did not lead to significant changes in SBP, DBP, or PP. These findings suggest that the physiological effects were greater when the olfactory system was stimulated. Conversely, in Takayama et al. (2022) no significant differences in DBP, SBP, or pulse rate were seen before and after experiencing the digital forest bathing environment.

3.7.2 Skin temperature

Only one study looked at skin temperature (ST) changes (Qi et al., 2022). While the results did not show significance, the restorative tendency for auditory stimuli was greater than in the control group.

3.7.3 Salivary amylase

Similarly, only one study measured salivary amylase (Takayama et al., 2022). No significant effect was observed when comparing the data before and after the multisensory digital forest experience.

3.7.4 Near-infrared spectroscopy (NIRS)

The study described in Song et al. (2019) demonstrated that the combined visual and olfactory stimuli yielded significantly reduced oxygenated hemoglobin (oxy-Hb) concentrations in both the left (control, −0.08 ± 0.10 μM; visual-olfactory, −0.48 ± 0.09 μM; $p < 0.05$) and right (control, 0.02 ± 0.10 μM; visual-olfactory, −0.46 ± 0.08 μM; $p < 0.05$) prefrontal cortices, compared to the control condition. Additionally, when exposed to the olfactory stimulus alone, a significant decrease in oxy-Hb concentration was observed in the right prefrontal cortex compared to the control condition (olfactory, −0.32 ± 0.12 μM). This decrease in oxy-Hb concentration aligns with previous literature suggesting its association with physiological calming (Hoshi et al., 2011). Overall, these findings suggest greater relaxation effects resulting from the combined visual and olfactory stimuli.

3.7.5 Cardiovascular data

Regarding the cardiovascular system, there were five studies that show results related to relaxation and stress reduction in multisensory experience (Song et al., 2019; Schebella et al., 2020; Lopes et al., 2022; Takayama et al., 2022; Zhao et al., 2022). In Lopes et al. (2022), despite not showing a statistically significant difference, the mean of the high frequency (HF) component of the electrocardiogram (ECG) signal (across all participants) for the multisensory condition was higher compared to the audiovisual-only condition. These findings are corroborated by other studies that showed increases in HF during deep relaxation situations, such as meditation, or a decrease in HF during stressful situations. In addition, the standard deviation of the absolute first difference of the ECG feature showed a significant negative correlation (r = −0.46, $p < 0.05$) with the subjective relaxation rating for the multisensory condition, thus suggesting that the users felt more relaxed in the multisensory experience. Furthermore, the results also showed a decrease in the maximum blood volume pulse (BVP) signal between the audio-visual and multisensory (audio-visual-olfactory-haptic) conditions, suggesting greater relaxation potential in the latter condition (Parent et al., 2020).

In Takayama et al. (2022), the heart rate significantly decreased ($p < 0.01$) and HF was significantly higher ($p = 0.014$) during the exposure to multisensory forest bathing compared with that during the resting state. On the other hand, the study in Zhao et al. (2022) used heart rate variability (HRV) to compare three conditions during stress-inducing and recovery tasks. They found significant differences ($p < 0.05$) in HRV among these conditions. Specifically, the audio-visual condition showed a notably higher HRV than the neutral condition across all metrics. Additionally, HRV in the audio-visual condition was generally higher than in the audio-visual-olfactory condition. The results suggest that the nature experience without scent was the most effective in promoting relaxation.

Regarding the comparison between urban and natural environments, in Schebella et al. (2020), results showed that there were no significant differences in recovery from stress induction in terms of heart rate and HRV. However, regarding the stimulus in the experience, the multisensory experience showed that the recovery from stress induction relative to a visual-only condition was higher in terms of reduced heart rate and HRV. Moreover, the greatest recovery from stress induction based on median HR recovery scores ($X^2(2) = 9.234$, $p = 0.007$) was found in the low biodiversity condition, relative to the moderate biodiversity natural environment. Lastly, in Song et al. (2019), findings derived from HRV (i.e., ln(LF/HF)) showed that only the visual stimulus resulted in significantly decreased sympathetic nervous activity compared to the control condition (control, −0.26 ± 0.17; visual, −0.67 ± 0.17; $p < 0.05$), which indicates a decrease in stress. When comparing the four conditions (control, visual, olfactory, and visual-olfactory) there was no significant difference in heart rate.

3.7.6 Electrodermal activity

Four studies reported changes in electrodermal activity (EDA) when including multisensory experiences. In Zhang et al. (2023), in the visual condition, the amplitudes of skin conductance (SC) (ΔSC = 0.19 ± 0.01 $\mu\Omega$, $p < 0.05$) increased significantly relative to the control group. In addition, SC values significantly increased under the olfactory–visual stimulus method (ΔSC = 0.45 ± 0.34 $\mu\Omega$, $p < 0.05$), and this increase was significantly higher than that of the olfactory-only condition. These findings may indicate that the effect of smelling and seeing garden plants provides more excitement than just smelling or seeing. On the other hand, the work in Qi et al. (2022) showed that in terms of the physiological restoration, the skin conductance level (SCL) value revealed a negative effect in the audio-olfactory stimulation, only the audio stimulus resulted in higher levels of restoration than the control group, illustrated by the SCL ($p = 0.04$).

According to the findings in Lopes et al. (2022), the participants in the multisensory condition exhibited a lower average of the low-frequency component in the EDA signal compared to the audio-visual condition. An elevation in the low-frequency component of EDA has been linked to stress (Posada-Quintero et al., 2016). Moreover, a Pearson correlation analysis revealed significant correlations ($p < 0.05$) between EDA features and relaxation ratings. Notably, the mean of the negative first difference of the EDA signal showed a correlation of r = 0.42, and the standard deviation of the EDA signal had a correlation of r = −0.43. Lastly, in the study by Hedblom et al. (2019), a greater decrease in EDA was observed in natural (park) environments compared to urban environments, with a significant difference ($\beta = 2.02$, (t(2471) = 2.45; $p < 0.02$)). A similar trend, though not statistically significant, was also noted between park and forest environments ($\beta = 0.56$, (t(2471) = 0.67, $p = 0.25$)). The study used a regression model to explore the relationship between perceived environmental pleasantness and SCL values during stress and recovery periods. It was found that only odor pleasantness significantly predicted SCL during both periods (stress period: $\beta = -0.164$; $p = 0.04$; recovery period: $\beta = -0.188$; $p = 0.03$), with auditory pleasantness showing marginal significance in the recovery period ($\beta = -0.45$; $p = 0.05$).

3.7.7 Electroencephalography

Most studies measuring electroencephalographic activity focused on alpha waves, which represent a state of calm and feelings of quietness and relaxation. Two studies, in addition to alpha waves, also showed results related to beta waves, which are related to effortful thinking and active attention. In Zhang et al. (2023), there was a significant increase in alpha and beta wave power during visual and combined olfactory-visual stimulation compared to before the stimulation. Specifically, alpha waves increased from 13.90 ± 7.19 μV to 20.10 ± 9.45 μV ($p < 0.001$) in the visual condition and from 16.03 ± 5.03 μV to 18.31 ± 6.77 μV ($p < 0.05$) in the olfactory-visual condition. Beta waves also rose from 8.27 ± 5.09 μV to 13.78 ± 6.79 μV ($p < 0.001$) in the visual condition and from 10.12 ± 4.35 μV to 11.52 ± 4.87 μV ($p < 0.05$) in the olfactory-visual condition. No significant changes were noted with olfactory stimulation alone. However, when compared to a control group, all three conditions showed a significant increase in both alpha and beta wave powers.

In Qi et al. (2022), audio-visual-olfactory stimuli led to greater restoration, but only for beta wave power (compared with the only-audio condition). As for the alpha wave power, correlations were seen with a lifting effect in the audio-visual condition and a negative effect in the audio-olfactory condition ($p = 0.00$) with a relatively equal mean difference when compared to the audio-only condition. Combining audio (birdsong) with olfactory stimuli from environments such as rose gardens and pine forests led to a significant decrease in α power ($p \leq 0.01$). An osmanthus garden also notably lowered α power ($p = 0.00$), indicating that certain plant landscapes may induce negative physical arousal. Regarding β band power, adding visual and visual-olfactory stimuli to audio-only (birdsong) scenarios generally improved restoration ($p = 0.00$). However, in lawn settings, the addition of visual-olfactory stimuli had a variable impact on β values ($p = 0.00$). The work in Amores et al. (2018) also investigated changes in theta wave, which can be observed during meditation and drowsiness. A relaxation metric based on the Renyi entropy of different EEG frequency bands was proposed and showed that the relaxation score increased significantly ($p < 0.05$) between the audio-visual-olfactory experience and the control condition (no stimulus).

3.8 Limitations and future study recommendations

Of the 154 studies looking at some experiments with immersive nature environments involving VR, only 14 covered audio-visual-olfactory aspects, suggesting this is an emerging topic with room for innovation and development. In the subsections to follow, several limitations are addressed and recommendations for future studies are provided.

3.8.1 Population

Several studies have faced limitations due to demographic constraints (e.g. (Sona et al., 2019; Qi et al., 2022; Shin et al., 2022; Zhang et al., 2023)). These studies often do not account for the full spectrum of diversity in age, gender, ethnicity, or socioeconomic status. This lack of diversity can hinder the generalizability of the findings, potentially leading to conclusions that may not apply broadly across different demographic groups. For example, the majority of the participants in the studies were comprised of university students, which could not fully represent the target population for multisensory immersive experience interventions. Moreover, very few studies had balanced gender representation in the participant pool. Some previous studies have reported increased chances of motion sickness (known as cybersickness) for females relative to males (e.g. (Kelly et al., 2023)). As such, future studies should strive to include a more diverse sample, including from a patient population, to enhance the generalizability of the findings.

Moreover, the reliability of the reported findings is often compromised by limited sample sizes (e.g. (Amores et al., 2018; Song et al., 2019; Schebella et al., 2020; Abbott and Diaz-Artiles, 2022; Lopes et al., 2022; Takayama et al., 2022; Zhao et al., 2022)). Small sample sizes can result in poor statistical power, making it challenging to detect true effects or differences, and may lead to results that are not replicable in larger, more diverse samples. Researchers should aim to increase sample sizes in future studies. Larger sample sizes would improve the statistical power and the robustness of the findings, contributing to more reliable and valid research outcomes.

Another significant limitation, as noted by Song et al. (2019), is the lack of studies with populations with heightened daily stress (e.g., first responders). Overlooking high-stress groups can lead to a gap in understanding their specific needs and responses. Future studies should concentrate on these populations, as targeted research can provide important insights into stress management and resilience, aiding in the development of effective interventions and therapies for those heavily impacted by stress.

Lastly, cultural background plays a pivotal role in shaping an individual's preference for specific natural odors, deeply influencing the design and customization of the olfactory element in a virtual experience. For example, the Japanese practice of Shinrin-Yoku highlights the cultural value placed on the immersive experience of forest scents, such as those from pines, cedars, and spruces. Furthermore, the type of vegetation and biodiversity characteristic of a region can significantly influence the natural odors that individuals find comforting. For instance, individuals from regions rich in coniferous forests may have a preference for the scents of pine or fir, while those from tropical areas might be drawn to the fragrances of flora like banyan or kauri trees. Acknowledging these cultural and regional differences is crucial in creating more inclusive and personalized virtual reality experiences. By offering a wide selection of natural scents, from the refreshing aroma of a dense forest to the salty breeze of a beach, virtual environments can cater to a global audience. This approach not only enhances the immersive quality of virtual experiences but also fosters a deeper connection between users and the virtual natural environments they choose to explore, based on their personal experiences with nature.

Further research should explore the impact of olfactory preferences on overall relaxation potential.

3.8.2 Lack of stimuli comparisons

Studies, such as those by (Song et al., 2019; Zhang et al., 2023), focused predominantly on visual and olfactory stimuli, while works, such as (Amores et al., 2018; Putrino et al., 2020; Takayama et al., 2022), emphasized multisensory experiences. The absence of a comprehensive integration strategy and comparison of the three sensory inputs can lead to a fragmented understanding of their synergistic potential and impact. Future research should strive to integrate and compare audio, visual, and olfactory stimuli in a cohesive manner. Studies could explore how these stimuli interact and influence mental health outcomes when used together, as opposed to in isolation or in paired combinations. This integrated approach can provide deeper insights into the multisensory experience, particularly in how olfactory stimuli complement and enhance visual and auditory experiences.

3.8.3 Lack of complementary measures

A critical limitation noted in studies, such as that by Zhang et al. (2023), is the exclusive focus on physiological indicators, neglecting psychological aspects. Conversely, research by Shin et al. (2022); Abbott and Diaz-Artiles (2022); Putrino et al. (2020); Sona et al. (2019); Serrano et al. (2016) have primarily relied on subjective analyses. This one-dimensional approach can result in an incomplete understanding of the study subject, as it fails to capture the holistic impact of the interventions or phenomena being studied. Future studies should aim to incorporate both objective (physiological) and subjective (psychological) measures. This dual approach would provide a more comprehensive understanding of the effects being studied. For instance, while physiological data can offer concrete evidence of bodily responses, psychological data can provide insights into personal experiences and perceptions. The integration of both types of data can lead to a more nuanced and complete understanding of the research findings.

In fact, several studies relied on a single objective measurement (e.g. (Amores et al., 2018; Hedblom et al., 2019; Zhao et al., 2022)), thus may have missed some important interactions between the central and nervous systems, which have been linked to different emotional states and behaviors. In turn, the work by Qi et al. (2022) relied on a single subjective measurement, which can result in a narrow understanding of participant experiences. Additionally, the studies by Lopes et al. (2022); Putrino et al. (2020); Amores et al. (2018) relied on single-item self-reports. Multi-item scales and the incorporation of a wider array of subjective tools should be used in future studies to capture a more comprehensive view of participant experiences and perspectives.

3.8.4 Smell perception

3.8.4.1 Narrow focus on odor elements in olfactory environments

Research, such as in Zhang et al. (2023), focused mainly on odor elements in garden green spaces, overlooking the broader spectrum of sensory experiences in olfactory environments. This limited perspective may not fully capture how various sensory inputs together enhance environmental quality and health benefits. Future research should broaden its scope to include how visual, auditory (and tactile) elements interact with olfactory experiences, offering a more comprehensive understanding of olfactory environments and their role in wellbeing and environmental enhancement.

3.8.4.2 Limited strategies in stimulating the sense of smell

Two studies did not compare the multisensory condition with a conventional audio-visual condition, thus making it difficult to fully characterize the impact of including olfaction and its role in anxiety/stress reduction (Amores et al., 2018; Takayama et al., 2022). It is recommended that future studies should include appropriate control conditions to allow direct comparisons of outcomes from conventional to multisensory interventions.

Moreover, the study by Zhao et al. (2022) highlighted a limitation in using a single method for olfactory stimulation (using the Aroma Shooter device), leading to uncertain conclusions about other techniques, such as other diffusion devices, incenses, or the use of natural plant aromas. Despite the fact that only this particular study acknowledged such a limitation, it is important to note that all the studies included in this review uniformly employed a singular method for stimulating the olfactory sense. This limits the understanding of the impact of different smell stimulation methods on stress recovery and health. Future studies should explore diverse smell stimulation methods, including varying scent types, intensities, and combinations, possibly alongside other senses. Comparing these strategies can offer insights into their effectiveness, particularly for stress relief.

3.8.4.3 Interactive olfactory technology

In acknowledging the significance of individual preferences in VR experiences, particularly within the olfactory dimension, it is important to emphasize that personalization and user interaction in virtual olfactory environments also have great importance in the overall experience. The ability for users to tailor their sensory experiences to their preferences can enhance the immersion and realism of VR.

Technologies such as the OVR ION Technology (2017) wearable scent diffusion device allow for up to nine different scents to be diffused, thus offering users a broad spectrum of olfactory experiences. The technology not only allows for the continuous diffusion of "background" scents but also enables interactive experiences where the emission of scents is tied to user actions and specific objects within the virtual environment. For instance, interacting with a virtual rose can trigger the emission of rose petal smells, thus closely mimicking real-world experiences.

Moreover, the Olorama Scent Generator (OSG) Olorama (2013) exemplifies another way for users to interact with virtual objects with up to 10 different smells. This device adapts the olfactory output based on the virtual scene, enhancing the immersive quality of the VR experience. Similarly, the SENSIKS SENSIKS (2017) sensory reality pod, while not portable, supports six distinct scents and modifies scent emissions in response to the virtual content, demonstrating the versatility and adaptability of current olfactory technology.

Additionally, with microcontrollers such as Arduino, developers can design their own VR-integrated smell devices, further democratizing the field and opening up new possibilities for personalized sensory experiences. Examples include the use of computer-controlled olfactometers Lundström et al. (2010), Essence Amores and Maes (2017), and olfactory wearables Amores Fernandez et al. (2023), highlighting the field of virtual olfactory experiences. These examples not only offer a wide range of scents for user selection but also enable dynamic and personalized VR experiences that cater to the nuanced preferences of users. Future studies should consider allowing the user to validate and calibrate the olfactory elements based on their own preferences (e.g., adjust smell "gains" based on proximity to the nose).

3.8.5 Individualized analyses

Studies, such as (Song et al., 2019; Zhang et al., 2023), were the only ones to analyze the stimuli and their effects at the individual level. As smell is very subjective (pleasant for some, unpleasant for others), performing analyses at the group level may average out some interesting individualized outcomes. Future research should highlight both individual and group analyses to provide deeper insights into how each stimulus uniquely contributes to the overall outcome, essential in sensory and psychological research.

3.8.6 Virtual environment choices

Most studies offered a limited range of nature scenes, often focusing on a single type of environment (e.g., a forest). This limited range fails to consider the diverse preferences and responses individuals might have to different natural settings. Only a few works (e.g. (Serrano et al., 2016; Hedblom et al., 2019; Putrino et al., 2020; Qi et al., 2022)) offered different types of nature scenes. Future research should expand the range of environmental options presented to participants. Offering a diverse array of nature scenes, each with distinct characteristics (e.g., beach, mountain, urban green space), can cater to varied individual preferences and potentially yield richer data on environment-person interactions.

Personalization can play a crucial role in how individuals interact with and respond to different settings, impacting their psychological and physiological wellbeing. For example, Reid et al. (2015) demonstrated that odors can evoke emotional and nostalgic responses, enhancing the immersive quality of digital experiences. Incorporating options for personalizing the environmental experience can be a significant step forward in future research. Allowing participants to choose or even modify their environment according to their preferences can provide insights into how personalization affects their experience and response.

Moreover, as highlighted by Lopes et al. (2022), virtual experiments are often short in duration, in the order of a few minutes. Such short durations may limit the effectiveness of heart rate variability, which can often rely on tens of minutes. To capture more detailed biological features, extending experiment times to at least 10 min is recommended for more reliable and comprehensive physiological assessments.

3.8.7 Modality mismatch

Studies such as Amores et al. (2018) and Serrano et al. (2016) highlight a crucial limitation: the mismatch of sensory modalities. For instance, using lavender essential oil in a beach scenario can create sensory dissonance. This inconsistency can disrupt the coherence of the user experience, potentially diminishing the intended effect of the environment, whether it be for relaxation, stress reduction, or immersion. Future research should focus on creating sensory congruence in multisensory environments. This means ensuring that all sensory inputs are harmoniously aligned to reinforce each other.

3.8.8 Limited objective measurements

Of the studies analyzed, only nine utilized objective measurements to quantitatively assess the effects of sensory stimuli. Objective measurement via wearables, for example, may allow for close to real-time monitoring of user affective states, as well as the development of biomarkers that could be used to quantitatively monitor intervention outcomes in the long term. Wearables have been shown to be a useful ally for stress management interventions (e.g. (Smith E. N. et al., 2020; Gomes et al., 2023)), thus should be considered in future studies. Identifying the most sensitive and informative wearable devices or measurements will be crucial for the understanding of the role of olfactory stimuli in digital nature bathing.

Moreover, only four of the nine studies relied on monitoring neural data while the user was immersed in the virtual environment (Amores et al., 2018; Song et al., 2019; Qi et al., 2022; Zhang et al., 2023). Unlike other wearable devices that are typically placed on wrists, chests, or arms, the collection of neural data, while the user is wearing an HMD, may be challenging. There are recent innovations in sensor-embedded HMDs that may help overcome this limitation (e.g. (Cassani et al., 2020; Bernal et al., 2022; Moinnereau et al., 2022)). Physiological measures derived from EEG can provide new insights into the immersive experience and its mental health impact. For instance, the works by Abbasi et al. (2019); Wu et al. (2023) showed the use of EEG to study emotional responses to olfactory stimuli. Future studies should consider the use of such devices to allow for real-time monitoring of mental/cognitive states, as well as the development of new neuromarkers of VR nature exposure outcomes.

3.8.9 Beyond audio-visual-olfactory

While this review did not specifically focus on multisensory experiences relying on only haptic/tactile/somatosensory stimuli, three of the fourteen studies did include audio-visual-olfactory-tactile stimuli. Recent studies are showing that as more senses are stimulated, a greater sense of presence, immersion, and engagement can be achieved (De Jesus Jr et al., 2022; Gougeh et al., 2022), as well as improved neural plasticity (Amini Gougeh and Falk, 2023). Vibroacoustic therapy has shown to be useful for stress management (Boyd-Brewer, 2003). Additionally, the effects of thermal perception have been shown to have restorative benefits Lyu et al. (2022); Song et al. (2024). Future studies should explore the inclusion of tactile/somatosensory stimuli to further enhance the relaxation potential of the intervention.

3.8.10 Quality analyses

Lastly, we attempted to quantify the quality of the reviewed studies via different quality checklists. While the majority was rated as moderate to high quality, two were deemed as low quality, which could have introduced biases and affected the overall reliability of the findings reported herein. Future works should aim to follow quality guidelines in (Tufanaru et al., 2020) to ensure outcomes are reliable.

3.9 Recommendations for future research

Of the 14 studies reviewed, those incorporating olfactory stimuli alongside visual and/or auditory elements showed a marked improvement in stress reduction and relaxation. Here, we base our recommendations on those with a greater number of participants, which are likely to have increased statistical power and can have outcomes better reproduced in future studies. In the examined studies, this corresponded to experiments with the number of participants greater than 50. The study by Zhang et al. (2023), for example, demonstrated that combining visual and olfactory senses significantly enhanced biomarkers of relaxation, with a sample of 95 participants and a statistical significance of $p < 0.05$. The study showed increases in DBP, SC amplitudes, alpha, and beta brainwaves, while PP decreased. They utilized panoramic photos of Osmanthus trees and the natural scent of Osmanthus flowers, suggesting that smells from real objects might be just as effective in inducing relaxation as the ones based on essential oils, as in other studies (e.g., Lopes et al., 2022; Takayama et al., 2022).

Interestingly, the work by Takayama et al. (2022) relied on projected nature scenes, as opposed to immersing the user via an HMD. This highlights the importance of the olfactory channel, regardless of the visual modality used to immerse the user. Future works should explore more closely the differences between audio-visual-olfactory immersion using photos, projected videos, and immersive videos presented via HMDs to gauge the benefits of improved presence on overall relaxation. This aspect was not compared in any of the investigated studies.

Moreover, the works by Schebella et al. (2020) and Hedblom et al. (2019), with 52 and 154 participants respectively, found that nature settings even with low biodiversity, such as an Eucalyptus forest or a park, were significantly more effective in eliciting positive emotions, such as happiness and pleasantness, compared to urban environments ($p < 0.05$). These studies also measured physiological responses through EDA and HR. Key to these studies was also the synchronized presentation of olfactory stimuli with the visual environment and, when applicable, auditory elements, such as birdsong and nature sounds. Immersion of all senses in a natural setting is at the heart of Shinrin-yoku (forest bathing) principles, thus future works should aim to stimulate as many senses as possible while simulating a nature setting.

4 Conclusion

This systematic review has presented a comprehensive synthesis of the psychological and physiological effects of multisensory (audio-visual-olfactory) virtual nature exposure on mental health, as well as existing systems. A total of 14 studies were examined, involving virtual forests, beaches, and parks, where corresponding odors were utilized in the multisensory conditions. The majority of the studies reported enhanced restoration outcomes following exposure to multisensory conditions. We conclude by discussing several limitations and propose some recommendations for future studies. The reviewed literature suggests that, overall, multisensory immersive digital nature exposure can be an intervention that holds promise for mental health and wellbeing.

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

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

The author(s) declare(s) that no financial support was received for the research, authorship, and/or publication of this article. The authors acknowledge funding from the Natural Sciences and Engineering Council (NSERC) of Canada (RGPIN-2021-03246).

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

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2024.1252539/full#supplementary-material

References

Abbasi, N. I., Bose, R., Bezerianos, A., Thakor, N. V., and Dragomir, A. (2019). "Eeg-based classification of olfactory response to pleasant stimuli," in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (IEEE), 5160–5163.

Abbott, R. W., and Diaz-Artiles, A. (2022). The impact of digital scents on behavioral health in a restorative virtual reality environment. *Acta Astronaut.* 197, 145–153. doi:10.1016/j.actaastro.2022.05.025

Amini Gougeh, R., and Falk, T. H. (2023). Enhancing motor imagery detection efficacy using multisensory virtual reality priming. *Front. Neuroergonomics* 4, 1080200. doi:10.3389/fnrgo.2023.1080200

Amores, J., and Maes, P. (2017). "Essence: olfactory interfaces for unconscious influence of mood and cognitive performance," in Proceedings of the 2017 CHI conference on human factors in computing systems, 28–34.

Amores, J., Richer, R., Zhao, N., Maes, P., and Eskofier, B. M. (2018). "Promoting relaxation using virtual reality, olfactory interfaces and wearable eeg," in 2018 IEEE 15th international conference on wearable and implantable body sensor networks (BSN) (IEEE), 98–101.

Amores Fernandez, J., Mehra, N., Rasch, B., and Maes, P. (2023). "Olfactory wearables for mobile targeted memory reactivation," in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 1–20.

Bernal, G., Hidalgo, N., Russomanno, C., and Maes, P. (2022). "Galea: a physiological sensing system for behavioral research in virtual environments," in 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (IEEE), 66–76.

Birckhead, B., Khalil, C., Liu, X., Conovitz, S., Rizzo, A., Danovitch, I., et al. (2019). Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. *JMIR Ment. health* 6, e11973. doi:10.2196/11973

Boyd-Brewer, C. (2003). Vibroacoustic therapy: sound vibrations in medicine. *Altern. Complementary Ther.* 9, 257–263. doi:10.1089/107628003322490706

Brengman, M., Willems, K., and De Gauquier, L. (2022). Customer engagement in multi-sensory virtual reality advertising: the effect of sound and scent congruence. *Front. Psychol.* 13, 747456. doi:10.3389/fpsyg.2022.747456

Carulli, M., Bordegoni, M., Bernecich, F., Spadoni, E., and Bolzan, P. (2019). "A multisensory virtual reality system for astronauts' entertainment and relaxation," in International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (American Society of Mechanical Engineers), 59179, V001T02A081.

Cassani, R., Moinnereau, M.-A., Ivanescu, L., Rosanne, O., and Falk, T. H. (2020). Neural interface instrumented virtual reality headsets: toward next-generation immersive applications. *IEEE Syst. Man, Cybern. Mag.* 6, 20–28. doi:10.1109/msmc.2019.2953627

Chu, S., and Downes, J. J. (2002). Proust nose best: odors are better cues of autobiographical memory. *Mem. cognition* 30, 511–518. doi:10.3758/bf03194952

Corazon, S. S., Sidenius, U., Poulsen, D. V., Gramkow, M. C., and Stigsdotter, U. K. (2019). Psycho-physiological stress recovery in outdoor nature-based interventions: a systematic review of the past eight years of research. *Int. J. Environ. Res. Public Health* 16, 1711. doi:10.3390/ijerph16101711

De Jesus, B., Jr, Lopes, M., Moinnereau, M.-A., Gougeh, R. A., Rosanne, O. M., Schubert, W., et al. (2022). "Quantifying multisensory immersive experiences using wearables: is (stimulating) more (senses) always merrier?," in *Proceedings of the 2nd workshop on multisensory experiences-SensoryX'22 (SBC).*

Dozio, N., Maggioni, E., Pittera, D., Gallace, A., and Obrist, M. (2021). May i smell your attention: exploration of smell and sound for visuospatial attention in virtual reality. *Front. Psychol.* 12, 671470. doi:10.3389/fpsyg.2021.671470

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

Fu, D., Serra, N. I., Mansion, H., Mansion, E. T., and Blain-Moraes, S. (2022). Assessing the effects of nature on physiological states using wearable technologies. *Int. J. Environ. Res. Public Health* 19, 1231. doi:10.3390/ijerph19031231

Giannico, V., Spano, G., Elia, M., D'Este, M., Sanesi, G., and Lafortezza, R. (2021). Green spaces, quality of life, and citizen perception in european cities. *Environ. Res.* 196, 110922. doi:10.1016/j.envres.2021.110922

Gomes, N., Pato, M., Lourenço, A. R., and Datia, N. (2023). A survey on wearable sensors for mental health monitoring. *Sensors* 23, 1330. doi:10.3390/s23031330

Gougeh, R. A., De Jesus, B. J., Lopes, M. K., Moinnereau, M.-A., Schubert, W., and Falk, T. H. (2022). "Quantifying user behaviour in multisensory immersive experiences," in 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE) (IEEE), 64–68.

Hansen, M. M., Jones, R., and Tocchini, K. (2017). Shinrin-yoku (forest bathing) and nature therapy: a state-of-the-art review. *Int. J. Environ. Res. public health* 14, 851. doi:10.3390/ijerph14080851

Hao, J., Xie, H., Harp, K., Chen, Z., and Siu, K.-C. (2022). Effects of virtual reality intervention on neural plasticity in stroke rehabilitation: a systematic review. *Archives Phys. Med. Rehabilitation* 103, 523–541. doi:10.1016/j.apmr.2021.06.024

Hawes, M. T., Szenczy, A. K., Klein, D. N., Hajcak, G., and Nelson, B. D. (2021). Increases in depression and anxiety symptoms in adolescents and young adults during the covid-19 pandemic. *Psychol. Med.* 52, 3222–3230. doi:10.1017/s0033291720005358

Hedblom, M., Gunnarsson, B., Iravani, B., Knez, I., Schaefer, M., Thorsson, P., et al. (2019). Reduction of physiological stress by urban green space in a multisensory virtual experiment. *Sci. Rep.* 9, 10113. doi:10.1038/s41598-019-46099-7

Herz, R. S. (1998). Are odors the best cues to memory? a cross-modal comparison of associative memory stimuli a. *Ann. N. Y. Acad. Sci.* 855, 670–674. doi:10.1111/j.1749-6632.1998.tb10643.x

Hoshi, Y., Huang, J., Kohri, S., Iguchi, Y., Naya, M., Okamoto, T., et al. (2011). Recognition of human emotions from cerebral blood flow changes in the frontal region: a study with event-related near-infrared spectroscopy. *J. Neuroimaging* 21, e94–e101. doi:10.1111/j.1552-6569.2009.00454.x

Igarashi, M., Ikei, H., Song, C., and Miyazaki, Y. (2014a). Effects of olfactory stimulation with rose and orange oil on prefrontal cortex activity. *Complementary Ther. Med.* 22, 1027–1031. doi:10.1016/j.ctim.2014.09.003

Igarashi, M., Song, C., Ikei, H., Ohira, T., and Miyazaki, Y. (2014b). Effect of olfactory stimulation by fresh rose flowers on autonomic nervous activity. *J. Altern. complementary Med.* 20, 727–731. doi:10.1089/acm.2014.0029

Jo, H., Song, C., and Miyazaki, Y. (2019). Physiological benefits of viewing nature: a systematic review of indoor experiments. *Int. J. Environ. Res. Public Health* 16, 4739. doi:10.3390/ijerph16234739

Kelly, J. W., Gilbert, S. B., Dorneich, M. C., and Costabile, K. A. (2023). "Gender differences in cybersickness: clarifying confusion and identifying paths forward," in 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (IEEE), 283–288.

Khan, H., and Nilsson, D. (2023). "Smell of fire increases behavioural realism in virtual reality: a case study on a recreated mgm grand hotel fire," in 2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (IEEE), 820–828.

Kotera, Y., Lyons, M., Vione, K. C., and Norton, B. (2021). Effect of nature walks on depression and anxiety: a systematic review. *Sustainability* 13, 4015. doi:10.3390/su13074015

Lee, M., Kim, E., Choe, J., Choi, S., Ha, S., and Kim, G. (2022). Psychological effects of green experiences in a virtual environment: a systematic review. *Forests* 13, 1625. doi:10.3390/f13101625

Lopes, M. K., de Jesus, B. J., Moinnereau, M.-A., Gougeh, R. A., Rosanne, O. M., Schubert, W., et al. (2022). "Nat (ur) e: Quantifying the relaxation potential of ultra-reality multisensory nature walk experiences," in 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE) (IEEE), 459–464.

Lundström, J. N., Gordon, A. R., Alden, E. C., Boesveldt, S., and Albrecht, J. (2010). Methods for building an inexpensive computer-controlled olfactometer for temporally-precise experiments. *Int. J. Psychophysiol.* 78, 179–189. doi:10.1016/j.ijpsycho.2010.07.007

Lyu, K., de Dear, R., Brambilla, A., and Globa, A. (2022). Restorative benefits of semi-outdoor environments at the workplace: does the thermal realm matter? *Build. Environ.* 222, 109355. doi:10.1016/j.buildenv.2022.109355

Millroth, P., and Frey, R. (2021). Fear and anxiety in the face of covid-19: negative dispositions towards risk and uncertainty as vulnerability factors. *J. Anxiety Disord.* 83, 102454. doi:10.1016/j.janxdis.2021.102454

Moinnereau, M.-A., Oliveira, A. A., and Falk, T. H. (2022). Instrumenting a virtual reality headset for at-home gamer experience monitoring and behavioural assessment. *Front. Virtual Real.* 3, 971054. doi:10.3389/frvir.2022.971054

Munyan, B. G., III, Neer, S. M., Beidel, D. C., and Jentsch, F. (2016). Olfactory stimuli increase presence in virtual environments. *PloS one* 11, e0157568. doi:10.1371/journal.pone.0157568

Nhlbi, N. (2022). *National institutes of health. study quality assessment tools: quality assessment tool*. Available at: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools (Accessed May).

Olorama, O. T. (2013). *Olorama technology olorama*. Available at: https://www.olorama.com/(Accessed February 19, 2024).

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The prisma 2020 statement: an updated guideline for reporting systematic reviews. *Int. J. Surg.* 88, 105906. doi:10.1016/j.ijsu.2021.105906

Parent, M., Albuquerque, I., Tiwari, A., Cassani, R., Gagnon, J.-F., Lafond, D., et al. (2020). PASS: a multimodal database of physical activity and stress for mobile passive body/brain-computer interface research. *Front. Neurosci.* 14, 542934. doi:10.3389/fnins.2020.542934

Pietrabissa, G., and Simpson, S. G. (2020). Psychological consequences of social isolation during covid-19 outbreak. *Front. Psychol.* 11, 2201. doi:10.3389/fpsyg.2020.02201

Posada-Quintero, H. F., Florian, J. P., Orjuela-Cañón, A. D., Aljama-Corrales, T., Charleston-Villalobos, S., and Chon, K. H. (2016). Power spectral density analysis of electrodermal activity for sympathetic function assessment. *Ann. Biomed. Eng.* 44, 3124–3135. doi:10.1007/s10439-016-1606-6

Putrino, D., Ripp, J., Herrera, J. E., Cortes, M., Kellner, C., Rizk, D., et al. (2020). Multisensory, nature-inspired recharge rooms yield short-term reductions in perceived stress among frontline healthcare workers. *Front. Psychol.* 11, 560833. doi:10.3389/fpsyg.2020.560833

Qi, Y., Chen, Q., Lin, F., Liu, Q., Zhang, X., Guo, J., et al. (2022). Comparative study on birdsong and its multi-sensory combinational effects on physio-psychological restoration. *J. Environ. Psychol.* 83, 101879. doi:10.1016/j.jenvp.2022.101879

Raudenbush, B., Grayhem, R., Sears, T., and Wilson, I. (2009). Effects of peppermint and cinnamon odor administration on simulated driving alertness, mood and workload. *North Am. J. Psychol.* 11, 245.

Reese, G., Stahlberg, J., and Menzel, C. (2022). Digital shinrin-yoku: do nature experiences in virtual reality reduce stress and increase well-being as strongly as similar experiences in a physical forest? *Virtual Real.* 26, 1245–1255. doi:10.1007/s10055-022-00631-9

Reid, C. A., Green, J. D., Wildschut, T., and Sedikides, C. (2015). Scent-evoked nostalgia. *Memory* 23, 157–166. doi:10.1080/09658211.2013.876048

Ritchie, H., and Roser, M. (2018). *Urbanization. Our world in data*.

Santomauro, D. F., Herrera, A. M. M., Shadid, J., Zheng, P., Ashbaugh, C., Pigott, D. M., et al. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the covid-19 pandemic. *Lancet* 398, 1700–1712. doi:10.1016/s0140-6736(21)02143-7

Schebella, M. F., Weber, D., Schultz, L., and Weinstein, P. (2020). The nature of reality: human stress recovery during exposure to biodiverse, multisensory virtual environments. *Int. J. Environ. Res. Public Health* 17, 56. doi:10.3390/ijerph17010056

SENSIKS (2017). *Sensiks*. Available at: https://www.sensiks.com/(Accessed February 19, 2024).

Serrano, B., Baños, R. M., and Botella, C. (2016). Virtual reality and stimulation of touch and smell for inducing relaxation: a randomized controlled trial. *Comput. Hum. Behav.* 55, 1–8. doi:10.1016/j.chb.2015.08.007

Shin, S., Browning, M. H., and Dzhambov, A. M. (2022). Window access to nature restores: a virtual reality experiment with greenspace views, sounds, and smells. *Ecopsychology* 14, 253–265. doi:10.1089/eco.2021.0032

Smith, B. M., Twohy, A. J., and Smith, G. S. (2020a). Psychological inflexibility and intolerance of uncertainty moderate the relationship between social isolation and mental health outcomes during covid-19. *J. Contextual Behav. Sci.* 18, 162–174. doi:10.1016/j.jcbs.2020.09.005

Smith, E. N., Santoro, E., Moraveji, N., Susi, M., and Crum, A. J. (2020b). Integrating wearables in stress management interventions: promising evidence from a randomized trial. *Int. J. Stress Manag.* 27, 172–182. doi:10.1037/str0000137

Sona, B., Dietl, E., and Steidle, A. (2019). Recovery in sensory-enriched break environments: integrating vision, sound and scent into simulated indoor and outdoor environments. *Ergonomics* 62, 521–536. doi:10.1080/00140139.2018.1491643

Song, C., Ikei, H., and Miyazaki, Y. (2019). Physiological effects of forest-related visual, olfactory, and combined stimuli on humans: an additive combined effect. *Urban For. Urban Green.* 44, 126437. doi:10.1016/j.ufug.2019.126437

Song, S., Xiao, Y., Tu, R., and Yin, S. (2024). Effects of thermal perception on restorative benefits by green space exposure: a pilot study in hot-humid China. *Urban Clim.* 53, 101767. doi:10.1016/j.uclim.2023.101767

Syed Abdullah, S. S., Awang Rambli, D. R., Sulaiman, S., Alyan, E., Merienne, F., and Diyana, N. (2021). The impact of virtual nature therapy on stress responses: a systematic qualitative review. *Forests* 12, 1776. doi:10.3390/f12121776

Takayama, N., Morikawa, T., Koga, K., Miyazaki, Y., Harada, K., Fukumoto, K., et al. (2022). Exploring the physiological and psychological effects of digital shinrin-yoku and its characteristics as a restorative environment. *Int. J. Environ. Res. Public Health* 19, 1202. doi:10.3390/ijerph19031202

Technology, O. (2017). *Ovr technology presentation, research and description*. Available at: https://ovrtechnology.com/(Accessed February 19, 2024).

Tufanaru, C., Munn, Z., Aromataris, E., Campbell, J., and Hopp, L. (2020). "Systematic reviews of effectiveness," in *JBI Manual for Evidence Synthesis*. Editors E. Aromataris, C. Lockwood, K. Porritt, B. Pilla, and Z. Jordan (Australia: The Joanna Briggs Institute Adelaide). doi:10.46658/JBIMES-24-03

Vernet-Maury, E., Alaoui-Ismaïli, O., Dittmar, A., Delhomme, G., and Chanel, J. (1999). Basic emotions induced by odorants: a new approach based on autonomic pattern results. *J. Aut. Nerv. Syst.* 75, 176–183. doi:10.1016/s0165-1838(98)00168-4

Wu, M., Teng, W., Fan, C., Pei, S., Li, P., and Lv, Z. (2023). An investigation of olfactory-enhanced video on eeg-based emotion recognition. *IEEE Trans. Neural Syst. Rehabilitation Eng.* 31, 1602–1613. doi:10.1109/tnsre.2023.3253866

Yao, W., Zhang, X., and Gong, Q. (2021). The effect of exposure to the natural environment on stress reduction: a meta-analysis. *Urban For. urban Green.* 57, 126932. doi:10.1016/j.ufug.2020.126932

Zhang, X., Guo, J., Zhang, X., and Zhang, Q. (2023). Physiological effects of a garden plant smellscape from the perspective of perceptual interaction. *Int. J. Environ. Res. Public Health* 20, 5004. doi:10.3390/ijerph20065004

Zhao, N., Kodama, E. C., and Paradiso, J. A. (2022). Mediated atmosphere table (mat): adaptive multimodal media system for stress restoration. *IEEE Internet Things J.* 9, 23614–23625. doi:10.1109/jiot.2022.3190929

frontiers | Frontiers in Virtual Reality

TYPE Original Research
PUBLISHED 04 June 2024
DOI 10.3389/frvir.2024.1379351

Check for updates

OPEN ACCESS

EDITED BY
Justyna Świdrak,
August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Spain

REVIEWED BY
Alejandro Beacco,
University of Barcelona, Spain
Pierre Bourdin-Kreitz,
Open University of Catalonia, Spain

*CORRESPONDENCE
Yasuhito Sawahata,
sawahata.y-jq@nhk.or.jp

RECEIVED 31 January 2024
ACCEPTED 14 May 2024
PUBLISHED 04 June 2024

CITATION
Sawahata Y, Harasawa M and Komine K (2024), Synergy and medial effects of multimodal cueing with auditory and electrostatic force stimuli on visual field guidance in 360° VR. *Front. Virtual Real.* 5:1379351. doi: 10.3389/frvir.2024.1379351

Synergy and medial effects of multimodal cueing with auditory and electrostatic force stimuli on visual field guidance in 360° VR

Yasuhito Sawahata*, Masamitsu Harasawa and Kazuteru Komine

Science and Technology Research Laboratories, Japan Broadcasting Corporation, Tokyo, Japan

This study investigates the effects of multimodal cues on visual field guidance in 360° virtual reality (VR). Although this technology provides highly immersive visual experiences through spontaneous viewing, this capability can disrupt the quality of experience and cause users to miss important objects or scenes. Multimodal cueing using non-visual stimuli to guide the users' heading, or their visual field, has the potential to preserve the spontaneous viewing experience without interfering with the original content. In this study, we present a visual field guidance method that imparts auditory and haptic stimulations using an artificial electrostatic force that can induce a subtle "fluffy" sensation on the skin. We conducted a visual search experiment in VR, wherein the participants attempted to find visual target stimuli both with and without multimodal cues, to investigate the behavioral characteristics produced by the guidance method. The results showed that the cues aided the participants in locating the target stimuli. However, the performance with simultaneous auditory and electrostatic cues was situated between those obtained when each cue was presented individually (*medial effect*), and no improvement was observed even when multiple cue stimuli pointed to the same target. In addition, a simulation analysis showed that this intermediate performance can be explained by the integrated perception model; that is, it is caused by an imbalanced perceptual uncertainty in each sensory cue for orienting to the correct view direction. The simulation analysis also showed that an improved performance (*synergy effect*) can be observed depending on the balance of the uncertainty, suggesting that a relative amount of uncertainty for each cue determines the performance. These results suggest that electrostatic force can be used to guide 360° viewing in VR, and that the performance of visual field guidance can be improved by introducing multimodal cues, the uncertainty of which is modulated to be less than or comparable to that of other cues. Our findings on the conditions that modulate multimodal cueing effects contribute to maximizing the quality of spontaneous 360° viewing experiences with multimodal guidance.

KEYWORDS

out-of-view problem, visual field guidance, electrostatic force, haptics, multimodal processing, integrated perception model

1 Introduction

The presentation of multimodal sensory information in virtual reality (VR) can considerably enhance the sense of presence and immersion. In daily life, we perceive the surrounding physical world through multiple senses, such as visual, auditory, and haptic senses, and interact with it based on these perceptions (Gibson, 1979; Flach and Holden, 1998; Dalgarno and Lee, 2010). Therefore, introducing multimodal stimulations into VR can enhance realism and significantly improve the experience. In fact, numerous studies have reported the benefits of multimodal VR (Mikropoulos and Natsis, 2011; Murray et al., 2016; Wang et al., 2016; Martin et al., 2022; Melo et al., 2022).

Head-mounted displays (HMDs) offer a highly immersive visual experience by spontaneously allowing users to view a 360° visual world; however, this feature may disrupt the 360° viewing experience, causing users to miss important objects or scenes that are located outside their visual field, thereby resulting in the "out-of-view" problem (Gruenefeld, El Ali, et al., 2017b). In 360° VR, the visual field of the user is defined by the viewport of the HMD. As nothing is presented outside the viewport, users have no opportunity for perception without changing the head direction. To address this problem, the presentation of arrows (Lin Y.-C et al., 2017; Schmitz et al., 2020; Wallgrun et al., 2020), peripheral flickering (Schmitz et al., 2020; Wallgrun et al., 2020), and picture-in-picture previews and thumbnails (Lin Y. T. et al., 2017; Yamaguchi et al., 2021) have been employed and shown to guide the gaze and visual attention effectively. However, these approaches also exhibit the problem of inevitably interfering with the video content, potentially disrupting the spontaneous viewing experience and mitigating the benefit of 360° video viewing (Sheikh et al., 2016; Pavel et al., 2017; Tong et al., 2019). Addressing this problem will significantly improve the 360° video viewing experience, especially for VR content with fixed time events, such as live scenes, movies, and dramas.

Several studies have explored the potential of multimodal stimuli to guide user behavior in 360° VR while preserving the original content (Rothe et al., 2019; Malpica, Serrano, Allue, et al., 2020b). Diegetic cues based on non-visual sensory stimuli such as directional sound emanating from a VR scene provide natural and intuitive guidance that feels appropriate in VR settings (Nielsen et al., 2016; Sheikh et al., 2016; Rothe et al., 2017; Rothe and Hußmann, 2018; Tong et al., 2019), exhibiting good compatibility with immersive 360° video viewing. At present, audio output is usually supported by any available HMDs and is the most common cue for visual field guidance. Because visual field guidance using non-visual stimuli is expected to provide high-quality VR experiences (Rothe et al., 2019), extensive research on various multimodal stimulation methods, including haptic stimulation, can aid the design of better VR experiences.

This study introduces electrostatic force stimuli to guide user behavior in selecting visual images that are displayed on the HMD (Figure 1). Previous studies have shown that applying an electrostatic force to the human body can induce a "fluffy" haptic sensation (Fukushima and Kajimoto, 2012a; 2012b; Suzuki et al., 2020; Karasawa and Kajimoto, 2021). Unlike some species of fish, amphibians, and mammals, humans do not possess electroreceptive abilities that allow them to perceive electric fields directly (Proske et al., 1998; Newton et al., 2019; Hüttner et al., 2023). However, as discussed in Karasawa and Kajimoto (2021), the haptic sensations that are produced through electrostatic stimulation are strongly related to the hair on the skin. Therefore, humans can indirectly perceive electrostatic stimulation through cutaneous mechanoreceptors (Horch et al., 1977; Johnson, 2001; Zimmerman et al., 2014), which are primarily stimulated by hair movements owing to electrostatic forces. Perceiving the physical world through cutaneous haptic sensations is a common experience in daily life, such as feeling the movement of air, and is expected to be a candidate method to guide user behavior naturally.

Many studies have proposed various methods of providing haptic sensations for visual field guidance, such as vibrations (Matsuda et al., 2020), normal forces on the face (Chang et al., 2018), and muscle stimulation (Tanaka et al., 2022), demonstrating that multimodal stimulation can improve the VR experience. Electrostatic force stimulation also provides haptic sensations, but can stimulate a relatively large area of the human body in a "fluffy" and subtle manner, which differs significantly from stimuli produced by other tactile stimulation methods, such as direct vibration stimulation through actuators. Karasawa and Kajimoto (2021) showed that electrostatic force stimulation can provide a feeling of *presence*. Previously, Slater (2009) and Slater et al. (2022) provided two views of immersive VR experiences, namely, place illusion (PI) and plausibility illusion (Psi), which refer to the

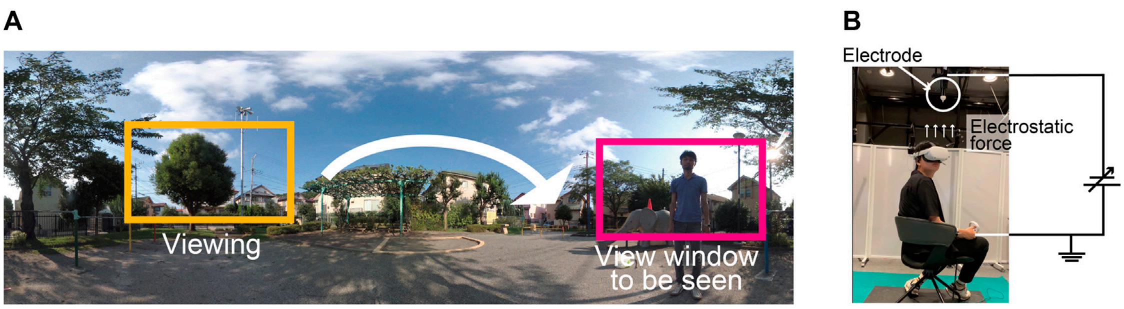


FIGURE 1
Visual field guidance in 360° VR using electrostatic force stimuli to mitigate the out-of-view problem. **(A)** Gentle visual field guidance. A user is viewing the scene depicted in the orange frame, whereas an important situation exists in the scene depicted in the red frame. Guiding the visual field to the proper direction will improve the user experience. **(B)** Haptic stimulus presentation using electrostatic forces. Electrostatic force helps the user to discover the important scene without affecting the original 360° VR content.

sensation of being in a real place and the illusion that the depicted scenario is actually occurring, respectively. In this sense, the effects of haptic stimulation on user experiences in VR belong to Psi. Such fluffy, subtle stimulation of the skin by electrostatic force has the potential to simulate the sensations of airflow, chills, and goosebumps, which are common daily-life experiences. The introduction of such modalities will enhance the plausibility of VR and lead to better VR experiences.

In this study, we presented electrostatic force stimuli using corona discharge, which is a phenomenon wherein ions are continuously emitted from a needle electrode at high voltages, allowing the provision of stimuli from a distance. Specifically, we placed the electrode above the user's head to stimulate a large area, from the head to the body (Figure 1B). Previous studies have employed plate- or pole-shaped electrodes to present such stimuli (Fukushima and Kajimoto, 2012a; 2012b; Karasawa and Kajimoto, 2021) and required the user to place their forearm close to the electrodes of the stimulation device, thereby limiting their body movement. The force becomes imperceptible even if the body parts are located 10 cm from the electrode (Karasawa and Kajimoto, 2021). As a typical VR user moves more than this distance, these conventional methods are not suitable for some VR applications that require physical movement. In addition, these devices are too bulky to be worn on the body. The proposed method can potentially overcome this limitation of distance and provide haptic sensations to VR users from a distance, thereby enabling the use of electrostatic force stimulation for visual field guidance in VR.

We evaluated the proposed visual field guidance method using multimodal cues in a psychophysical experiment. Previous studies have systematically evaluated visual field guidance using visual cues (Gruenefeld, Ennenga, et al., 2017a; Gruenefeld, El Ali, et al., 2017b; Danieau et al., 2017; Gruenefeld et al., 2018; 2019; Harada and Ohyama, 2022) in VR versions of visual search experiments (Treisman and Gelade, 1980; McElree and Carrasco, 1999). This study similarly investigated the effects of multimodal cues on visual searching.

Although numerous studies have shown that multiple modalities in VR can significantly improve the immersive experience (Ranasinghe et al., 2017; 2018; Cooper et al., 2018), it is unclear whether visual field guidance can also be improved by introducing multiple non-overt cues. We believe that multiple overt cues, such as visual arrows and halos, would help users to perform search tasks. However, this is not necessarily true for non-overt, subtle, and vague cues. Although guidance through subtle cues can minimize content intrusion (Bailey et al., 2009; Nielsen et al., 2016; Sheikh et al., 2016; Bala et al., 2019), it is not always guaranteed to be effective (Rothe et al., 2018). However, employing multiple subtle cues and integrating them into a coherent cue may provide effective overall guidance. In this study, in addition to electrostatic forces, we introduced weak auditory stimuli as subtle environmental cues to investigate the interaction effects of electrostatic and auditory cues on the guidance performance in VR as well as whether they improve, worsen, or have no effect on the guidance performance.

The nature of multimodal perception, which involves the integration of various sensory inputs to produce a coherent perception, has been understood using statistical models, such as maximum likelihood estimation and integration based on Bayes' theorem (Ernst and Banks, 2002; Ernst, 2006; 2007; Spence, 2011). Although such computational modeling approaches are also expected to aid in comprehending the underlying mechanisms of multimodal cueing effects on visual field guidance, to the best of our knowledge, this aspect remains unexplored. Therefore, we adopted a similar approach using computational models and investigated the effects of various cueing conditions on visual field guidance. Thus, this study offers a detailed understanding of multimodal visual field guidance and knowledge for predicting user behavior under various cue conditions.

We first introduce electrostatic force and auditory stimuli as multimodal cues in a visual search task and then show that electrostatic force can potentially address the out-of-view problem. Because auditory stimuli have been commonly used in previous studies to guide user behavior (Walker and Lindsay, 2003; Rothe et al., 2017; Bala et al., 2018; Malpica, Serrano, Allue, et al., 2020a; Malpica, Serrano, Gutierrez, et al., 2020b; Chao et al., 2020; Masia et al., 2021), a baseline is provided for comparisons. In the visual search task, the participants were instructed to find a specific visual target as quickly as possible in 360° VR, both with and without sensory cues. We anticipated that the cueing would reduce the cumulative travel angles associated with updating the head direction during the search. Therefore, a comparison of the task performances in each condition revealed the effect of multimodal cueing on the visual field guidance.

In this study, we hypothesized that performance with multimodal cueing in the visual search task in VR would show one of the following three effects: 1) a performance improvement compared to that with electrostatic force or auditory cues (*synergy effect*); 2) the same performance as that with the better cue, not considering the performance worth the other cue (*masking effect*); and 3) performance between the individual performances with each cue (*medial effect*). We conducted a psychophysical experiment to investigate which of these effects were observed with multimodal cues. Subsequently, through the psychophysical experiment and an additional simulation analysis, we demonstrated that both the *synergy* and *medial effects* can be observed depending on the balance of perceptual uncertainties for each cue and the variance in the selection of the head direction. Finally, we investigated the conditions for effective multimodal visual field guidance.

2 Materials and methods

2.1 Visual search experiment with multimodal cues

This subsection describes the experiment that was conducted to investigate the effects of visual field guidance on visual search performance in 360° VR using haptic and auditory cues. In addition, the multimodal effects of simultaneous cueing using haptic and auditory stimuli were investigated. The search performance was measured based on the travel angles, which are the cumulative rotation angles of the head direction, as described in detail in Section 2.2.1.1. Finally, we determined which of the effects, namely, *synergy*, *medial*, or *masking*, were likely by comparing the travel angles obtained in each cue condition.

2.1.1 Participants

Fifteen participants (seven male, eight female; aged 21–33 years, mean: 24.4) were recruited for this experiment. All participants had normal or corrected-to-normal vision. Two participants were excluded because their psychological thresholds for the electrostatic force stimuli were too high and exceeded the intensity range that our apparatus could present. Informed consent was obtained from all participants, and the study design was approved by the ethics committee of the Science and Technology Research Laboratories, Japan Broadcasting Corporation.

2.1.2 Apparatus

A corona charging gun (GC90N, Green Techno, Japan) was used to present electrostatic force stimuli. This device comprises a needle-shaped electrode (ion-emitting gun) and a high-voltage power supply unit (rated voltage range: 0 to −90 [kV]). The electrostatic force intensity was modulated by adjusting the applied voltage. The gun was hung from the ceiling and placed approximately 50 cm above the participant's head, as shown in Figure 1B. In addition, the participant wore a wristband attached to the ground to avoid accidental shocks owing to unintentional charging.

A standalone HMD, Meta Quest 2 (Meta, United States), was used to present the 360° visual images and auditory stimuli, and the controller joystick (for the right hand) was used to collect the responses. The HMD communicated with the corona charging gun via an Arduino-based microcomputer (M5Stick-C PLUS, M5Stack Technology, China) to control the analog inputs for the gun. The delay between the auditory and electrostatic force stimuli was a maximum of 20 ms, which was sufficiently small to perform the task. The participants viewed the 360° images while sitting in a swivel chair to facilitate viewing. They wore wired earphones (SE215, SURE, United States), which were connected to the HMD and used to present auditory stimuli using functions provided in Unity (Unity Technologies, United States) throughout the experiment, even when no auditory stimuli were presented. The experimental room was soundproof. Participant safety was duly considered; the floor was covered with an electrically grounded conductive mat, which collected ions that were not meant for the participant, thereby preventing unintentional charging of other objects in the room.

2.1.3 Stimuli

2.1.3.1 Visual stimuli

The target and distractor stimuli were presented in a VR environment implemented in Unity (2021.3.2 f1). The target stimulus included a randomly selected white symbol among "├", "┤", "┬", and "┴," whereas the distractor stimuli included white "┼" symbols. These stimuli were displayed on a gray background and distributed within a range of -10°–10° relative to each intersection of the latitudes and longitudes of a sphere with a 5-m radius that was centered at the origin. The referential latitudes and longitudes were placed at each 36° position of the horizontal 360° view and 22.5° positions between the elevation angles of −45° and 45°. Thus, 1 target and 39 distractor stimuli were presented at 10 × 4 locations. The stimuli sizes were randomly selected from visual angles ranging from 2.86° ± 1.43°, both horizontally and vertically. The difficulty of the task was modulated by varying the stimulus size and placement and the parameter values were selected based on our preliminary experiments.

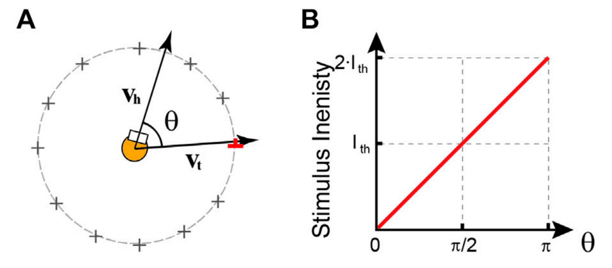


FIGURE 2
Stimulus intensity modulation. The stimulus intensity was linearly modulated in response to the inner angle, θ, between the head direction vector $\mathbf{v}_h$ and vector to the target stimulus $\mathbf{v}_t$. **(A)** Schematic view of θ, $\mathbf{v}_h$, and $\mathbf{v}_t$, which are defined in the *xyz* space. **(B)** Linear relationship between stimulus intensity and θ. Both the electrostatic force and auditory cues were modulated.

2.1.3.2 Electrostatic force stimuli

In this study, the electrostatic force stimuli are referred to as haptic stimuli induced by the corona charging gun. The electrostatic force intensity was determined based on the gun voltage. We selected the physical intensity of the electrostatic force for each participant based on their psychological threshold; the intensity ranged from zero to twice the threshold. Thus, we ensured that the stimulus intensity was psychologically equivalent among all participants. The threshold I_{th}, which was largely dependent on each participant, was measured before the experiments using the method of staircase, and it typically ranged from −10 to −30 kV. We linearly modulated the stimulus intensity in response to the inner angle θ between the head-direction vector $\mathbf{v}_h$ and target stimulus vector $\mathbf{v}_t$, as shown in Figure 2A. When the target was in front, i.e., $\theta = 0$, no electrostatic force was presented, whereas when it was behind, i.e., $\theta = \pi$, the strongest electrostatic force of $2I_{th}$ was presented. Therefore, the electrostatic force was regarded as a cue stimulus because participants could potentially find the target stimulus by updating their head direction to avoid the subtle haptic sensations. That is, when the haptic sensations were sufficiently weak, the target stimulus was likely to be within the participant's visual field. This is the natural behavior of most people because a strong electrostatic stimulus is typically considered unpleasant.

2.1.3.3 Auditory stimuli

Monaural white noise was used as the auditory stimulus. We used the same modulation method for the auditory stimuli as that for the electrostatic force stimuli, as shown in Figure 2. Specifically, we linearly modulated the stimulus intensity in response to the inner angle θ between the head-direction vector $\mathbf{v}_h$ and target stimulus vector $\mathbf{v}_t$. When the target was in front, i.e., $\theta = 0$, no sound was presented, whereas when it was behind, i.e., $\theta = \pi$, the maximum amplitude (volume) of the stimulus of $2I_{th}$ was presented. As with the electrostatic force stimuli, the threshold I_{th} for auditory stimuli was measured for each participant before the experiments using the method of staircase.

2.1.4 Task and conditions

We designed a within-participant experiment to compare the effects of haptic and auditory guidance in a visual search task. The participants were instructed to find the target stimulus and indicate its direction using the joystick on the VR controller. For example, when they discovered a target stimulus "⊥," they tilted the joystick upward as quickly as possible. The trial was terminated once the joystick was manipulated. Feedback was provided between sessions, showing the success rate of the previous session, to encourage participants to complete the task. The task was conducted both with and without sensory cues, resulting in four conditions based on the combinations of cue stimuli: visual only (V), vision with auditory (A), vision with electrostatic force (E), and vision with auditory and electrostatic force (AE) cues.

2.1.5 Procedure

The experiment included 12 sessions comprising 12 visual search trials, for a total of 144 trials per participant. Therefore, each condition (V, A, E, and AE) was presented 36 times in one experiment. In three of the 12 sessions, only condition V was presented, whereas in the other sessions, conditions A, E, and AE were presented in a pseudo-random order. Before each session, we informed the participants whether the next session would be a V-only session or a session with the non-visual-cued conditions. This prevented participants from waiting for non-visual cues during condition V and inadvertently wasting search time.

Each trial comprised a rest period of variable-length (3–6 s) and a 10-s search period. In the rest period, 40 randomly generated distractors were presented, whereas in the following search period, one of the distractors was replaced with a target stimulus. The trials progressed as soon as the target stimulus was found or when the 10-s time limit was reached. Note that the participants underwent two practice sessions to understand the task and response methods prior to these sessions.

2.2 Analysis

2.2.1 Behavioral data analysis

2.2.1.1 Modeling

We recorded the participants' responses and extents of their head movements during the search period. The trials with a correct response were labeled as successful, whereas those with an incorrect or no response were labeled as failed. The travel angle was defined as the accumulated rotational changes in the head direction during the target search. If guidance by electrostatic forces and auditory cues is effective, the travel angles should be shorter than those with no cues. Therefore, we investigated the modulation efficiency of the target discovery according to cue type.

The travel angle allowed us to model the participants' behavior in the visual search experiment with non-overt multimodal cues appropriately. In the original visual search experiment (Treisman and Gelade, 1980; McElree and Carrasco, 1999), wherein participants had to find the target stimuli with specified visual features as quickly as possible, the performance was measured by the reaction time required for identification. These experimental paradigms have recently been extended to investigate user behavior in VR. Cue-based visual search experiments in VR involve the analysis of reaction times and/or movement angles towards a target object (Gruenefeld, Ennenga, et al., 2017a; Gruenefeld, El Ali, et al., 2017b; Danieau et al., 2017; Gruenefeld et al., 2018; 2019; Schmitz et al., 2020; Harada and Ohyama, 2022). In addition, previous studies employed overt cues that directly indicated the target location, whereas we employed non-overt cues that weakly indicated them, without interfering with the visuals. This difference could have affected the behavior of participants, depending on their individual traits. For example, some participants may have adopted a scanning strategy wherein they sequentially scanned the surrounding visual world, ignoring the cues because they considered subtle cues to be unreliable. Participants with better physical ability could have completed the task faster using this strategy. In such cases, the reaction time would not accurately reflect the effects of cueing on the visual search performance and the effects would differ significantly from those we were investigating. Because behaviors including scanning that are not based on presented cues would result in larger travel angles, the effects of cues would likely be better reflected in the travel angle than in the reaction time. Therefore, we employed travel angles instead of reaction times to evaluate the performance.

We employed Bayesian modeling to evaluate the efficacy of each cue, as follows:

$$p(k|\lambda, \Phi) = \frac{(\lambda\Phi)^k}{k!} e^{-\lambda\Phi}, \tag{1}$$

where k is the number of discoveries (successful trials), λ is the expected target discovery rate, and Φ is the total travel angle. The probability of k given λ and Φ was calculated using the Poisson process (see 2.2.1.2). Note that $\Phi = \sum_{i=1}^{n} \phi_i$, where n is the number of trials and ϕ_i is the travel angle during the i-th trial. By applying the Bayes theorem to Eq. 1, the posterior distribution of λ can be expressed as $p(\lambda|k, \Phi) \propto p(k|\lambda, \Phi)p(\lambda)$. By assuming a noninformative prior on $p(\lambda)$, $p(\lambda|k, \Phi)$ is proportional to the right side of Eq. 1. Therefore, the expectation of $p(\lambda|k, \Phi)$ represents the target discovery rate, as follows:

$$\lambda = E[p(\lambda|k, \Phi)] = \frac{k}{\Phi} \tag{2}$$

Thus, λ was interpreted as the number of discoveries per travel angle.

2.2.1.2 Poisson process model derivation

The total travel angle Φ was divided into N bins of width $\Delta\phi = \Phi/N$. The probability of finding a target stimulus in a bin with the expected λ is $\Delta\phi\lambda$. Therefore, the probability of finding targets in k from N bins is represented by the following binomial distribution:

$$p(k|\lambda) = \frac{N!}{(N-k)!k!}(1 - \Delta\phi\lambda)^{N-k}(\Delta\phi\lambda)^k. \tag{3}$$

By minimizing the bin width using $N \to \infty$, we obtain $\lim_{N\to\infty}(1 - \Delta\phi\lambda)^{N-k} = \lim_{\epsilon\to 0}(1+\epsilon)^{-1/\epsilon} = e^{-\lambda\Phi}$ using the following relationships: $\epsilon = -\Delta\phi\lambda$ and $N - k \approx N = \Phi/\Delta\phi$. In addition, a relationship exists between $N!/(N-k)! \approx N^k = (\Phi/\Delta\phi)^k$ and $N \to \infty$. Finally, we obtain the following Poisson process:

$$p(k|\lambda) = \frac{1}{k!}\left(\frac{\Phi}{\Delta\phi}\right)^k e^{-\lambda\Phi}(\Delta\phi\lambda)^k = \frac{(\lambda\Phi)^k}{k!}e^{-\lambda\Phi}. \tag{4}$$

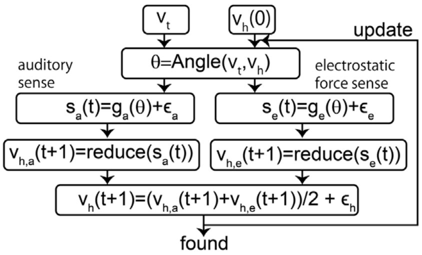


FIGURE 3
Perceptual model of visual search with multimodal cues. The possible head directions were estimated separately based on the synthesized auditory and electrostatic force sensations generated by $g_a(\cdot)$ and $g_e(\cdot)$. The final head direction in each iteration was determined by averaging the estimated directions.

2.2.1.3 Statistics

We created a dataset by pooling all observations that were obtained from the participants. Thereafter, we obtained the posterior distributions of the target discovery rate, $p(\lambda|k, \Phi)$, for each condition. Subsequently, the significance of the visual field guidance was assessed by comparing the distribution shapes. For example, when λ was larger for condition E than that for condition V and their distributions overlapped slightly, we concluded that the electrostatic force-based guidance significantly affected the visual field guidance. The overlap was quantified by the area under the curve (AUC) metric, the value of which ranged from 0 to 1; a smaller overlap resulted in an AUC value closer to 1. We compared the posterior distribution of λ in condition AE with those in conditions A and E to identify the multimodal effect.

2.3 Simulation analysis

2.3.1 Overview

To better comprehend how participants processed the multimodal inputs in the experiment, we conducted a simulation analysis assuming a perceptual model wherein a participant determined the head direction by simply averaging two vectors directed towards the target induced through auditory and haptic sensations, as shown in Figure 3, constituting the most typical explanation of the multimodal effect (Ernst and Banks, 2002; Ernst, 2006; 2007). We manipulated the noise levels ϵ_a, ϵ_e, and $\boldsymbol{\epsilon}_h$, assumed for the auditory sensations, haptic sensations, and orienting head directions, respectively, as shown in Figure 3. Thereafter, we examined the relationship between the noise levels and target discovery rates for each stimulus condition.

We implemented a computational model to determine the target stimulus direction based on the synthesized sensations. The head direction vector at time t is represented as $\mathbf{v}_h(t)$ ($\|\mathbf{v}_h(t)\| = 1$), and the auditory and electrostatic force sensory inputs for the model are denoted as $s_a(t)$ and $s_e(t)$, respectively. The model estimated the next head direction $\mathbf{v}_h(t+1)$ such that the sensory inputs were reduced. By iterating these procedures, $s_a(\cdot)$ and/or $s_e(\cdot)$ were minimized and the target stimulus in the direction of $\mathbf{v}_t$ could be identified. The detailed procedure is presented in Section 2.3.2.

The simulation was initially conducted using randomly generated $\mathbf{v}_h(t)$ and $\mathbf{v}_t$ values. The search was iterated according to the synthesized sensations $s_a(t)$ and $s_e(t)$, using different noise levels ϵ_a and ϵ_e, as shown in Figure 3. To simulate multimodal processing, the model estimated $\mathbf{v}_h(t+1)$ by averaging the $\mathbf{v}_{h,a}(t+1)$ and $\mathbf{v}_{h,e}(t+1)$ estimations. The term $\boldsymbol{\epsilon}_h \in \mathbb{R}^3$ was introduced to represent orienting errors between the estimated and actual directions owing to the physical constraints and other factors during the real experiment. Note that in the unimodal conditions, $\mathbf{v}_h(t+1) = \mathbf{v}_{h,a}(t+1)$ or $\mathbf{v}_h(t+1) = \mathbf{v}_{h,e}(t+1)$. An inner angle of $< \pm 30^\circ$ between $\mathbf{v}_h(t+1)$ and $\mathbf{v}_t$ indicated that the target stimulus was found and the iteration was terminated.

We ran the simulation using the parameter settings that were closest to those used in the real experiment; for example, the maximum amount and speed of head rotation and the number of trials were appropriately selected. The simulation was performed 468 times for each condition, corresponding to the setup in the real experiment (36 trials × 13 participants). The travel angle and target discovery rate were computed using the methods described in Section 2.2.1.1. To examine the effects of ϵ_a, ϵ_e, and $\boldsymbol{\epsilon}_h$ on λ for conditions A, E, and AE, we generated them using the following parameters: $\epsilon_a \sim \mathcal{N}(0, \sigma_a^2)$ with $0.05^2 < \sigma_a^2 < 0.50^2$, $\epsilon_e \sim \mathcal{N}(0, \sigma_e^2)$ with $0.05^2 < \sigma_e^2 < 0.50^2$, and $\boldsymbol{\epsilon}_h \sim \mathcal{N}(0, \sigma_h^2 \mathrm{I})$ with $\sigma_h^2 = 0.01^2$ and 0.1^2.

2.3.2 Procedure

In this section, we describe the details of the simulation, as summarized in Section 2.3.1 and Figure 3.

The simulation model iteratively updated the head direction vector $\mathbf{v}_h(\cdot)$ based on synthetic sensory inputs. The next head direction $\mathbf{v}_h(t+1)$ was determined using two steps: first, the possible head directions that minimized the target vector ($\mathbf{v}_t$) error were estimated independently for each modality, and thereafter, $\mathbf{v}_h(t+1)$ was obtained by averaging the estimated directions. In reality, because $\mathbf{v}_t$ was unknown, it was substituted with its estimate, which was obtained using an auditory or electrostatic force sensation, i.e., $\hat{\mathbf{v}}_{t,a}$ or $\hat{\mathbf{v}}_{t,e}$, respectively. Thus, the model determined the next head direction using a gradient descent search, as follows:

$$\mathbf{v}_{h,a}(t+1) = \mathbf{v}_h(t) - \alpha \nabla \left(\|\hat{\mathbf{v}}_{t,a}(t) - \mathbf{v}_h\|^2 \right)\Big|_{\mathbf{v}_h = \mathbf{v}_h(t)}, \tag{5}$$

$$\mathbf{v}_{h,e}(t+1) = \mathbf{v}_h(t) - \alpha \nabla \left(\|\hat{\mathbf{v}}_{t,e}(t) - \mathbf{v}_h\|^2 \right)\Big|_{\mathbf{v}_h = \mathbf{v}_h(t)}, \tag{6}$$

where $\alpha\,(>0)$ is a step-size parameter. The value of α corresponds to the head rotation speed during the experiment. Thus, Eqs 5, 6 can be rewritten as:

$$\mathbf{v}_{h,a}(t+1) = \mathbf{v}_h(t) - 2\alpha\left(\hat{\mathbf{v}}_{t,a}(t) - \mathbf{v}_h(t)\right), \tag{7}$$

$$\mathbf{v}_{h,e}(t+1) = \mathbf{v}_h(t) - 2\alpha\left(\hat{\mathbf{v}}_{t,e}(t) - \mathbf{v}_h(t)\right). \tag{8}$$

Finally, the next head direction vector was obtained as follows:

$$\mathbf{v}_h(t+1) = \frac{1}{2}\left(\mathbf{v}_{h,a}(t+1) + \mathbf{v}_{h,e}(t+1)\right) + \boldsymbol{\epsilon}_h, \tag{9}$$

where $\boldsymbol{\epsilon}_h$ follows $\mathcal{N}(0, \sigma_h^2 \mathrm{I})$ and represents the fluctuations associated with the head motion. Note that Eq. 9 is normalized before the next iteration. In unimodal simulations, Eq. 9 can be substituted with $\mathbf{v}_h(t+1) = \mathbf{v}_{h,a}(t+1) + \boldsymbol{\epsilon}_h$ or $\mathbf{v}_h(t+1) = \mathbf{v}_{h,e}(t+1) + \boldsymbol{\epsilon}_h$.

$\mathbf{v}_t$ was estimated using past auditory and somatosensory observations $s_a(t), ..., s_a(t-N)$ and $s_e(t), ..., s_e(t-N)$, respectively, where N is the number of observations used for the estimation. Because the stimulus intensities are given by the inner angle between the head and target directions, the simulated sensory inputs $s_a(t)$ and $s_e(t)$ can be expressed as

$$s_a(t) = \cos^{-1}(\mathbf{v}_t \cdot \mathbf{v}_h(t)) + \epsilon_a, \quad (10)$$

$$s_e(t) = \cos^{-1}(\mathbf{v}_t \cdot \mathbf{v}_h(t)) + \epsilon_e, \quad (11)$$

where ϵ_a and ϵ_e are the noise terms that follow normal distributions with $\mathcal{N}(0, \sigma_a^2)$ and $\mathcal{N}(0, \sigma_e^2)$, respectively, and σ_a^2 and σ_e^2 indicate the amount of noise generated.

We define a head-direction matrix $\mathrm{V}_h = [\,\mathbf{v}_h(t) \ \cdots \ \mathbf{v}_h(t-T)\,]^{\mathrm{t}}$ and observation vectors $\mathbf{s}_a = [\,s_a(t) \ \cdots \ s_a(t-T)\,]^{\mathrm{t}}$ and $\mathbf{s}_e = [\,s_e(t) \ \cdots \ s_e(t-T)\,]^{\mathrm{t}}$. If we assume that $\mathbf{s}_a$ and $\mathbf{s}_e$ are negatively correlated with $\mathrm{V}_h\mathbf{v}_t$, we can estimate $\mathbf{v}_t$ such that $\mathbf{s}_a^{\mathrm{t}}\mathrm{V}_h\mathbf{v}_t$ and $\mathbf{s}_e^{\mathrm{t}}\mathrm{V}_h\mathbf{v}_t$ are minimized under the constraint of $\|\mathbf{v}_t\| = 1$, assuming that V_h, $\mathbf{s}_a$, and $\mathbf{s}_e$ are centered in advance. Letting $\mathbf{v}_{t,a}$ and $\mathbf{v}_{t,e}$ be the target vectors obtained through auditory and haptic signals, the estimation is then tractable using the method of the Lagrange multiplier method, as follows:

$$L_a = \mathbf{s}_a^{\mathrm{t}}\mathrm{V}_h\mathbf{v}_{t,a} + \lambda_a\left(\mathbf{v}_{t,a}^{\mathrm{t}}\mathbf{v}_{t,a} - 1\right), \quad (12)$$

$$L_e = \mathbf{s}_e^{\mathrm{t}}\mathrm{V}_h\mathbf{v}_{t,e} + \lambda_e\left(\mathbf{v}_{t,e}^{\mathrm{t}}\mathbf{v}_{t,e} - 1\right), \quad (13)$$

where L_a and L_e are the Lagrangian functions for each modality, and λ_a and λ_e are the Lagrange multipliers. By considering $\hat{\mathbf{v}}_{t,a}$ and $\hat{\mathbf{v}}_{t,e}$ as the estimators of $\mathbf{v}_{t,a}$ and $\mathbf{v}_{t,e}$, respectively, in Eqs 12, 13, we obtain their values by considering $\partial L_a/\partial \mathbf{v}_{t,a} = \mathbf{0}$ and $\partial L_e/\partial \mathbf{v}_{t,e} = \mathbf{0}$ with $\|\mathbf{v}_{t,a}\| = 1$ and $\|\mathbf{v}_{t,e}\| = 1$, respectively:

$$\hat{\mathbf{v}}_{t,a} = -\frac{\mathrm{V}_h^{\mathrm{t}}\mathbf{s}_a}{\sqrt{\mathbf{s}_a^{\mathrm{t}}\mathrm{V}_h\mathrm{V}_h^{\mathrm{t}}\mathbf{s}_a}}, \quad (14)$$

$$\hat{\mathbf{v}}_{t,e} = -\frac{\mathrm{V}_h^{\mathrm{t}}\mathbf{s}_e}{\sqrt{\mathbf{s}_e^{\mathrm{t}}\mathrm{V}_h\mathrm{V}_h^{\mathrm{t}}\mathbf{s}_e}}. \quad (15)$$

Therefore, by substituting $\hat{\mathbf{v}}_{t,a}(t)$ and $\hat{\mathbf{v}}_{t,e}(t)$ in Eqs 7, 8 with Eqs 14, 15, the next head direction vectors could be estimated.

Initially, $\mathbf{v}_t$ and $\mathbf{v}_h(0)$ were randomly selected from the 360° omnidirectional candidates. In addition, $\mathbf{v}_h(t)$ and $1 < t < N$ were also generated around $\mathbf{v}_h(0)$. Based on the sampled $\mathbf{v}_t$ and $\mathbf{v}_h(t)$, synthetic sensations were generated using Eqs 9–11.

The target search was conducted using a maximum of 1000 steps. The simulation parameter values of $N = 10$ and $\alpha = 0.01$ were selected to ensure that the target discovery rates were similar to those observed in the real experiments.

3 Results

3.1 Behavioral results

We pooled all data obtained from the 13 participants. The number of successful trials k and accumulated travel distances Φ for each condition were $k = 352$ and $\Phi = 3.12 \times 10^3$ for V; $k = 422$ and $\Phi = 2.44 \times 10^3$ for A; $k = 391$ and $\Phi = 2.80 \times 10^3$ for E; and $k = 417$ and $\Phi = 2.66 \times 10^3$ for AE. Note that the numbers of trials with response errors, which appeared to be owing to manipulation errors, were: 6, 7, 10, and 4 for the V, A, E, and AE, respectively. Figure 4 shows histograms of the successful and failed trials plotted against the travel angles, wherein the blue and red plots denote successful and failed trials, respectively. The target stimuli were identified in all conditions even if the travel angles were short or close to zero because they could appear in the participant's visual field at the beginning of the trial, as the target locations were determined randomly. The failed trials featured longer travel angles, suggesting that the participants looked around but were unable to complete the task within the time limit.

Figure 5 shows the posterior distributions of λ for each condition. The expected discovery rates λ_V, λ_A, λ_E, and λ_{AE} for each condition were 0.113, 0.173, 0.140, and 0.157, respectively. As predicted, guidance with the auditory cues significantly improved the target discovery rate compared to condition V, and no overlap was observed between the distributions. In addition, the target discovery rate improved significantly in condition E compared with condition V, although not as much as that in condition A. The AUC between the distributions of conditions V and E was 0.998, suggesting that λ for condition E was significantly higher than that for condition V. In addition, there was no overlap between the distributions of condition V and the other conditions, A and AE, indicating that the AUCs were 1. This indicates that visual field guidance using electrostatic force is effective even in a VR environment wherein users view a 360° world using both head and body movements.

We observed that the performance in condition AE was situated between those in conditions A and E, thereby rejecting the possibilities of *synergy* and *masking effects* because the search using both cues did not enhance the performance and the participants could not ignore the other cue. This result supports the *medial effect*, which was one of the anticipated candidates.

3.2 Simulation results

Figure 6 shows λ for each cue condition, plotted against the uncertainty ratio for the AE cues for varying electrostatic force uncertainty under constant auditory cue uncertainty. σ_a^2, σ_e^2, and σ_h^2 represent the perceptual uncertainty, i.e., the variances for ϵ_a, ϵ_e, and ϵ_h, respectively, as shown in Figure 3. Summaries of the parameters and observations shown in Figures 6A, D are presented in Tables 1, 2, respectively.

As shown in Figures 6A, D, the expected λ values for unimodal cueing in condition E and multimodal cueing in condition AE asymptotically decreased as σ_e^2/σ_a^2 increased in both cases with $\sigma_h^2 = 0.01^2$ and 0.1^2. The *medial effect*, which was elicited in our behavioral data, was particularly observed for larger ratios of σ_e^2/σ_a^2 and σ_h^2, as shown in Figures 6D, F. Notably, when the σ_e^2/σ_a^2 ratio was close to 1 under $\sigma_h^2 = 0.01^2$ (Figures 6A, B), we observed the *synergy effect*, wherein λ with multimodal cueing was better than that with unimodal cueing. The magnitude of the σ_e^2/σ_a^2 ratio indicates the bias level of the uncertainty of the electrostatic force sensation over the auditory sensation. Therefore, these results suggest that both *medial* and *synergy effects* were observable under the assumption of the typical multimodal integration model (Ernst and Banks, 2002; Ernst, 2006; 2007) (Figure 3), depending on the uncertainty bias for each sensation and head motion.

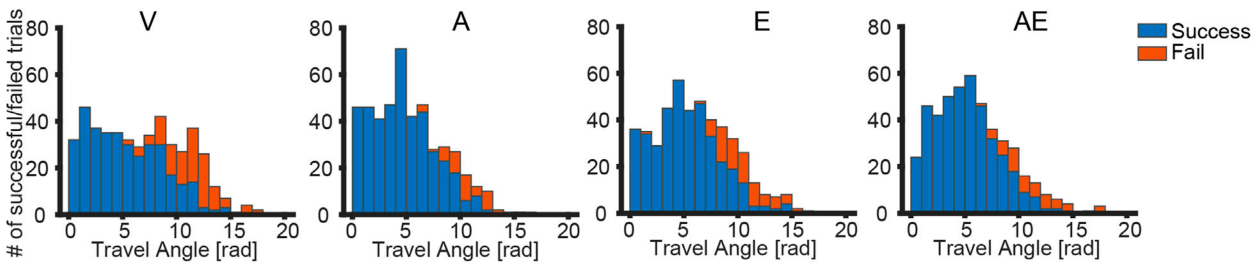


FIGURE 4
Relationships between the number of discoveries and travel angles. The data of 13 participants were pooled. The panels, from left to right, shows the relationships for each condition: vision only (V), vision + auditory cue (A), vision + electrostatic force cue (E), and vision + auditory and electrostatic force cues (AE).

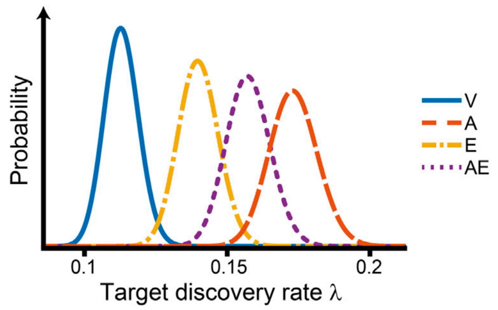


FIGURE 5
λ variations for each condition. A larger λ indicates better performance. Evidently, compared with condition V, λ was improved more under conditions E and A. Each plot was drawn based on Eq. 1, with the number of successful trials k, and the accumulated travel distances Φ observed for each condition.

4 Discussion

We demonstrated the multimodal effects of AE cues on visual field guidance in 360° VR, and found that both *medial* and *synergy effects* were observable depending on the uncertainty of the cue stimuli through the psychophysical experiment and the simulation analysis. Specifically, guidance performance with multimodal cueing is modulated by balancing the perceptual uncertainty elicited by each cue stimulus. We also demonstrated that the applicability of the electrostatic force-based stimulation method in VR applications; electrostatic stimulation through the corona charging gun allowed users to make large body movements. These results suggest that multimodal cueing with electrostatic force has sufficient potential to guide user behavior in 360° VR gently, offering a highly immersive visual experience through spontaneous viewing.

We showed that electrostatic force can be used as a haptic cue to guide the visual field. However, the search performance did not reach that with the auditory cue, even though we selected cue intensities that varied equally in small ranges around the supra- and sub-thresholds, with no significant difference in the perceptual domain. In the informal post-experiment interviews, some participants reported that the sensation induced by the electrostatic force was attenuated, especially while moving. In addition, most participants reported that the auditory cue made it easier to identify the target location. This suggests that the haptic sensation was affected by body motion that inevitably accompanied the updating of the head direction. The increased uncertainty for the haptic sensation was estimated to be approximately five times greater than that for the auditory sensation, as suggested by the simulation results (Figure 6F). Thus, the perception of changes in the stimulus intensity associated with visual field updates acts as a cue for estimating the target direction, which means that increasing the electrostatic field intensity such that it is strong enough to resist the effects of body motion could mitigate this uncertainty. As suggested by the simulation results presented in Section 3.2, reducing the perceptual uncertainty improves the search performance. This finding has been overlooked in previous studies that mainly focused on visual field guidance using overt cue stimuli (Gruenefeld, Ennenga, et al., 2017a; Gruenefeld, El Ali, et al., 2017b; Danieau et al., 2017; Gruenefeld et al., 2018; 2019; Harada and Ohyama, 2022). This results in the requirement for the property of cue stimuli to improve performance in multimodal visual field guidance.

The *medial effect* might have been counterintuitive because participants received more information regarding the target stimulus from multimodal cues than unimodal cues. Because the cues conveyed the same information, the *synergy effect* was more likely if participants used the received information properly. The simulation analysis showed that both effects could be observed under specific noise settings. This can also be explained theoretically: let A and E be random variables for auditory and electrostatic force sensations, respectively. Then, $V(A)$ and $V(E)$ represent the variances for each sensation. According to the integrated perception model (Ernst and Banks, 2002; Ernst, 2006; 2007), the total sensation variance can be expressed as

$$V\left(\frac{A+E}{2}\right) = \frac{1}{4}\left(V(A) + V(E) + 2Cov(A,E)\right), \qquad (16)$$

where $Cov(A,E)$ denotes the covariance between A and E. If A and E are independent, i.e., $Cov(A,E) = 0$, and $V(A)$ and $V(E)$ are equal, according to Eq. 16, $V((A+E)/2)$ is less than both $V(A)$ and $V(E)$, indicating a more efficient search performance than unimodal cueing (*synergy effect*) because smaller variances improve the performance. For example, if $V(E) = 5V(A)$, $V((A+E)/2)$ should be $3/2 \cdot V(A)$, suggesting intermediate performance if $V(A) < V((A+E)/2) < V(E)$ is used (*medial effect*). However, if $V(A)$ and $V(E)$ are not independent and $Cov(A,E)$ has a certain value, $V((A+E)/2)$

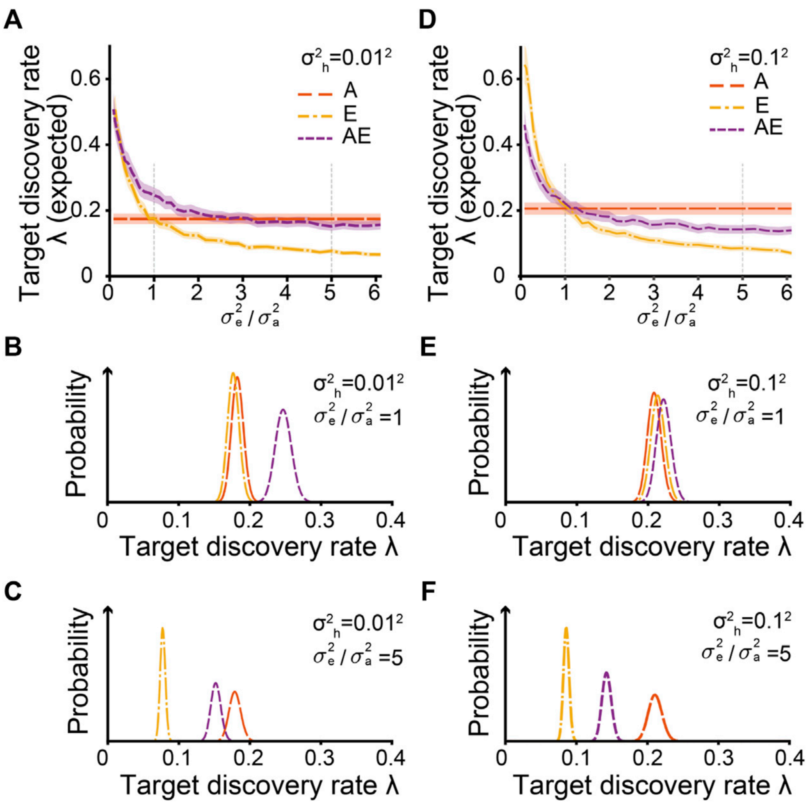


FIGURE 6
Comparisons of λ in the simulation analysis. **(A–C)** and **(D–F)** show the effects of $\sigma_h^2 = 0.01^2$ and 0.1^2 on λ, respectively. The expected λ values under each condition are plotted against the σ_e^2/σ_a^2 ratio using a representative value of $\sigma_a^2 = 0.17^2$. As no electrostatic-force stimuli were presented in condition A, λ could not be technically plotted against σ_e^2/σ_a^2. However, for reference, as λ for condition A was independent of σ_e^2, we plotted the expected λ values for condition A as straight lines through σ_e^2/σ_a^2 using a constant σ_a^2 value. The shaded areas behind the plots denote 95% credible intervals of the posterior distribution of λ. **(B,E)**, and **(C,F)** show the posterior distributions of λ for $\sigma_e^2/\sigma_a^2 = 1$ and 5, respectively. The original values presented in **(A–C)** and **(D–F)** are shown in Tables 1, 2, respectively.

increases and the *synergy effect* fades. In reality, ϵ_h in Figure 3 controlled the dependence between A and E, as the variance of ϵ_h is determined based on the observation of *synergy* or *medial effects.* These results support the validity of the integrated perception model shown in Figure 3 as the underlying mechanism of visual search tasks with multimodal cues.

Addressing the out-of-view problem has been a major challenge in 360° VR video viewing (Lin Y. T. et al., 2017; Schmitz et al., 2020; Wallgrun et al., 2020; Yamaguchi et al., 2021). Gentle and diegetic guidance that does not interfere with the visual content has received substantial attention from VR content providers (Nielsen et al., 2016; Sheikh et al., 2016; Rothe et al., 2017; Rothe and Hußmann, 2018; Bala et al., 2019; Tong et al., 2019). This study showed that subtle cues using artificial electrostatic force can guide the visual field, thereby demonstrating the application potential for 360° VR. Whereas previous studies using static electricity have severely limited the movements of the user (Fukushima and Kajimoto, 2012b; 2012a; Karasawa and Kajimoto, 2021), the use of the corona discharge phenomenon mitigated this limitation. The simulation analysis using the computational model helped to provide an understanding of the mechanisms of multimodal cueing. Similar to the observations in this study, previous studies using non-overt cues with perceptual uncertainty have reported both positive and negative effects of

TABLE 1 Summary of parameters and observations in simulation analysis ($\sigma_h^2 = 0.01^2$).

σ_e^2	σ_e^2/σ_a^2	E			AE		
		k	Φ	λ	k	Φ	λ
0.05^2	0.09	468	952.9	0.491	468	921.3	0.508
0.08^2	0.22	468	1105.4	0.423	468	1133.7	0.413
0.11^2	0.42	468	1583.2	0.296	468	1362.7	0.343
0.14^2	0.68	468	2176.8	0.215	468	1741.6	0.269
0.17^2	1.00	468	2651.1	0.177	468	1897.4	0.247
0.20^2	1.38	468	3106.3	0.151	468	2093.9	0.224
0.23^2	1.83	467	3761.7	0.124	468	2339.1	0.200
0.26^2	2.34	466	4249.6	0.110	468	2577.8	0.182
0.29^2	2.91	464	4691.1	0.099	468	2616.1	0.179
0.32^2	3.54	461	5150.7	0.090	468	2837.3	0.165
0.35^2	4.24	460	5707.8	0.081	467	2818.1	0.166
0.38^2	5.00	457	5920.3	0.077	468	3082.3	0.152
0.41^2	5.82	446	6698.8	0.067	467	3045.0	0.154

TABLE 2 Summary of parameters and observations in simulation analysis ($\sigma_h^2 = 0.1^2$).

σ_e^2	σ_e^2/σ_a^2	E			AE		
		k	Φ	λ	k	Φ	λ
0.05^2	0.09	468	727.3	0.643	468	1016.1	0.461
0.08^2	0.22	468	882.3	0.530	468	1285.8	0.364
0.11^2	0.42	468	1338.2	0.350	468	1595.7	0.293
0.14^2	0.68	468	1753.9	0.267	468	1899.7	0.246
0.17^2	1.00	468	2191.7	0.214	468	2111.3	0.222
0.20^2	1.38	468	2906.4	0.161	468	2386.7	0.196
0.23^2	1.83	468	3337.2	0.140	468	2567.0	0.182
0.26^2	2.34	468	3567.1	0.131	468	2811.7	0.166
0.29^2	2.91	468	4272.9	0.110	468	2990.2	0.157
0.32^2	3.54	467	4723.8	0.099	468	2925.0	0.160
0.35^2	4.24	467	5069.9	0.092	468	3162.1	0.148
0.38^2	5.00	466	5443.8	0.086	468	3288.5	0.142
0.41^2	5.82	462	5884.6	0.079	468	3418.1	0.137

multimodal cueing in 360° VR (Sheikh et al., 2016; Rothe and Hußmann, 2018; Bala et al., 2019; Malpica, Serrano, Gutierrez, et al., 2020a). We believe that our results also provide a rational explanation for these previous findings.

However, this study had some limitations. Some participants exhibited insufficient sensitivity to the electrostatic force stimuli. Although their hair moved when they were exposed to static electricity, they reported low sensations, which may be caused by skin moisture or other factors; however, this phenomenon has not yet been investigated. Furthermore, as humans are incapable of electroreception, it is reasonable to believe that the mechanoreceptors in the skin are involved in providing the sensations (Horch et al., 1977; Johnson, 2001; Zimmerman et al., 2014); however, this must be investigated further. In addition, the wristband used to tether the participants to the ground may have restricted free body movement; this can be addressed by introducing an ionizer that remotely neutralizes the charge level (Ohsawa, 2005), thereby allowing participants to move freely. Finally, the results presented in this study were obtained under reductive conditions. While the results provide insight into stimulus design, further experiments are required to demonstrate the effectiveness in real-world VR applications such as video viewing and gaming, which will be the focus of our future study.

In future work, we will implement electrostatic stimulation in a VR application. We believe that haptic stimulation by electrostatic force could be used not only to guide the visual field, but also to enhance the user's subjective impression. Although this has not been discussed here, we have experimentally implemented a VR game wherein a user shoots zombies charged with static electricity approaching from all sides. The electrostatic force-based stimulus can result in unpleasant sensations. Other haptic stimuli, such as vibrations, could also be used to cue the zombies. However, we believe that these stimuli are too obvious and artificial, and may detract from the subjective quality of experience to a certain extent. The use of static electricity can result in an unsettling experience for users when charged zombies approach them from behind. Thus, by comparing the effects of electrostatic force and other haptic stimuli on subjective impressions, we will be able to demonstrate the availability of electrostatic force-based stimulation to provide a highly immersive experience.

5 Conclusion

We investigated the multimodal effects of auditory and electrostatic force-based haptic cues on visual field guidance in 360° VR, demonstrating the potential for a visual field guidance method that does not interfere with the visual content. We found that modulating the degree of perceptual uncertainty for each cue improves the overall guidance performance under simultaneous multimodal cueing. Moreover, we presented a simple haptic stimulation method using only a single channel of a corona charging gun. In the future, we will increase the number of channels to present more complex stimulations in a larger area by dynamically controlling the electric fields, allowing for remote haptic stimulation under a six-degrees-of-freedom viewing condition. Finally, our results showed that multimodal stimuli have the potential to increase the richness in VR environments.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, upon reasonable request.

Ethics statement

The studies involving humans were approved by Japan Broadcasting Corporation. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YS: Conceptualization, Methodology, Formal analysis, Writing–original draft, Writing–review and editing. MH: Supervision, Writing–review and editing. KK: Project administration, Writing–review and editing.

Funding

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

Conflict of interest

Authors YS, MH, and KK were employed by Japan Broadcasting Corporation.

Publisher's note

References

Bailey, R., McNamara, A., Sudarsanam, N., and Grimm, C. (2009). Subtle gaze direction. *ACM Trans. Graph.* 28 (4), 1–14. doi:10.1145/1559755.1559757

Bala, P., Masu, R., Nisi, V., and Nunes, N. (2018). "Cue control: interactive sound spatialization for 360° videos," in *Interactive storytelling. ICIDS 2018. Lecture notes in computer science, vol 11318* Editors R. Rouse, H. Koenitz, and M. Haahr (Cham: Springer), 333–337. doi:10.1007/978-3-030-04028-4_36

Bala, P., Masu, R., Nisi, V., and Nunes, N. (2019). "When the elephant trumps," in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK, May 4-9, 2019, 1–13.

Chang, H.-Y., Tseng, W.-J., Tsai, C.-E., Chen, H.-Y., Peiris, R. L., and Chan, L. (2018). "FacePush: introducing normal force on face with head-mounted displays," in Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, 927–935.

Chao, F. Y., Ozcinar, C., Wang, C., Zerman, E., Zhang, L., Hamidouche, W., et al. (2020). "Audio-visual perception of omnidirectional video for virtual reality applications," in 2020 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2020, 2–7.

Cooper, N., Milella, F., Pinto, C., Cant, I., White, M., and Meyer, G. (2018). The effects of substitute multisensory feedback on task performance and the sense of presence in a virtual reality environment. *PLOS ONE* 13 (2), e0191846. doi:10.1371/journal.pone.0191846

Dalgarno, B., and Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *Br. J. Educ. Technol.* 41 (1), 10–32. doi:10.1111/j.1467-8535.2009.01038.x

Danieau, F., Guillo, A., and Doré, R. (2017). "Attention guidance for immersive video content in head-mounted displays," in 2017 IEEE Virtual Reality (VR), 205–206.

Ernst, M. O. (2006). "A bayesian view on multimodal cue integration," in *Human body perception from the inside out: advances in visual cognition* (Oxford University Press), 105–131.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *J. Vis.* 7 (5), 7. doi:10.1167/7.5.7

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415 (6870), 429–433. doi:10.1038/415429a

Flach, J. M., and Holden, J. G. (1998). The reality of experience: gibson's way. *Presence* 7 (1), 90–95. doi:10.1162/105474698565550

Fukushima, S., and Kajimoto, H. (2012a). "Chilly chair: facilitating an emotional feeling with artificial piloerection," in *ACM SIGGRAPH 2012 emerging Technologies (SIGGRAPH '12), 1, article*, 5–1. doi:10.1145/2343456.2343461

Fukushima, S., and Kajimoto, H. (2012b). "Facilitating a surprised feeling by artificial control of piloerection on the forearm," in Proceedings of the 3rd Augmented Human International Conference (AH '12), Article 8, 1–4.

Gibson, J. J. (1979) *The ecological approach to visual perception 3.* Houghton Mifflin.

Gruenefeld, U., El Ali, A., Boll, S., and Heuten, W. (2018). "Beyond halo and wedge: visualizing out-of-view objects on head-mounted virtual and augmented reality devices," in Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '18), 1–11.

Gruenefeld, U., El Ali, A., Heuten, W., and Boll, S. (2017a). "Visualizing out-of-view objects in head-mounted augmented reality," in Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17), Article 87, 1–7.

Gruenefeld, U., Ennenga, D., Ali, A.El, Heuten, W., and Boll, S. (2017b). "EyeSee360: designing a visualization technique for out-of-view objects in head-mounted augmented reality," in Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17), 109–118.

Gruenefeld, U., Koethe, I., Lange, D., Weirb, S., and Heuten, W. (2019). "Comparing techniques for visualizing moving out-of-view objects in head-mounted virtual reality," in 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 742–746.

Harada, Y., and Ohyama, J. (2022). Quantitative evaluation of visual guidance effects for 360-degree directions. *Virtual Real.* 26 (2), 759–770. doi:10.1007/s10055-021-00574-7

Horch, K. W., Tuckett, R. P., and Burgess, P. R. (1977). A key to the classification of cutaneous mechanoreceptors. *J. Investigative Dermatology* 69 (1), 75–82. doi:10.1111/1523-1747.ep12497887

Hüttner, T., von Fersen, L., Miersch, L., and Dehnhardt, G. (2023). Passive electroreception in bottlenose dolphins (*Tursiops truncatus*): implication for micro- and large-scale orientation. *J. Exp. Biol.* 226 (22), jeb245845. doi:10.1242/jeb.245845

Johnson, K. (2001). The roles and functions of cutaneous mechanoreceptors. *Curr. Opin. Neurobiol.* 11 (4), 455–461. doi:10.1016/S0959-4388(00)00234-8

Karasawa, M., and Kajimoto, H. (2021). "Presentation of a feeling of presence using an electrostatic field: presence-like sensation presentation using an electrostatic field," in Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21), Article 285, 1–4.

Lin, Y.-C., Chang, Y.-J., Hu, H.-N., Cheng, H.-T., Huang, C.-W., and Sun, M. (2017). "Tell me where to look," in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2535–2545.

Lin, Y. T., Liao, Y. C., Teng, S. Y., Chung, Y. J., Chan, L., and Chen, B. Y. (2017). "Outside-in: visualizing out-of-sight regions-of-interest in a 360 video using spatial picture-in-picture previews," in Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17), 255–265.

Malpica, S., Serrano, A., Allue, M., Bedia, M. G., and Masia, B. (2020a). Crossmodal perception in virtual reality. *Multimedia Tools Appl.* 79 (5–6), 3311–3331. doi:10.1007/s11042-019-7331-z

Malpica, S., Serrano, A., Gutierrez, D., and Masia, B. (2020b). Auditory stimuli degrade visual performance in virtual reality. *Sci. Rep.* 10 (1), 12363–12369. doi:10.1038/s41598-020-69135-3

Martin, D., Malpica, S., Gutierrez, D., Masia, B., and Serrano, A. (2022). Multimodality in VR: a survey. *ACM Comput. Surv.* 54 (10s), 1–36. doi:10.1145/3508361

Masia, B., Camon, J., Gutierrez, D., and Serrano, A. (2021). Influence of directional sound cues on users' exploration across 360° movie cuts. *IEEE Comput. Graph. Appl.* 41 (4), 64–75. doi:10.1109/MCG.2021.3064688

Matsuda, A., Nozawa, K., Takata, K., Izumihara, A., and Rekimoto, J. (2020). "HapticPointer," in Proceedings of the Augmented Humans International Conference (AHs '20), 1–10.

McElree, B., and Carrasco, M. (1999). The temporal dynamics of visual search: evidence for parallel processing in feature and conjunction searches. *J. Exp. Psychol. Hum. Percept. Perform.* 25 (6), 1517–1539. doi:10.1037/0096-1523.25.6.1517

Melo, M., Goncalves, G., Monteiro, P., Coelho, H., Vasconcelos-Raposo, J., and Bessa, M. (2022). Do multisensory stimuli benefit the virtual reality experience? A systematic review. *IEEE Trans. Vis. Comput. Graph.* 28 (2), 1428–1442. doi:10.1109/TVCG.2020.3010088

Mikropoulos, T. A., and Natsis, A. (2011). Educational virtual environments: a ten-year review of empirical research (1999–2009). *Comput. Educ.* 56 (3), 769–780. doi:10.1016/j.compedu.2010.10.020

Murray, N., Lee, B., Qiao, Y., and Muntean, G. M. (2016). Olfaction-enhanced multimedia: a survey of application domains, displays, and research challenges. *ACM Comput. Surv.* 48 (4), 1–34. doi:10.1145/2816454

Newton, K. C., Gill, A. B., and Kajiura, S. M. (2019). Electroreception in marine fishes: chondrichthyans. *J. Fish Biol.* 95 (1), 135–154. doi:10.1111/jfb.14068

Nielsen, L. T., Møller, M. B., Hartmeyer, S. D., Ljung, T. C. M., Nilsson, N. C., Nordahl, R., et al. (2016). "Missing the point," in Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, 229–232.

Ohsawa, A. (2005). Modeling of charge neutralization by ionizer. *J. Electrost.* 63 (6–10), 767–773. doi:10.1016/j.elstat.2005.03.043

Pavel, A., Hartmann, B., and Agrawala, M. (2017). "Shot orientation controls for interactive cinematography with 360° video," in Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17), 289–297.

Proske, U., Gregory, J. E., and Iggo, A. (1998). Sensory receptors in monotremes. *Philosophical Trans. R. Soc. B Biol. Sci.* 353 (1372), 1187–1198. doi:10.1098/rstb.1998.0275

Ranasinghe, N., Jain, P., Karwita, S., Tolley, D., and Do, E. Y.-L. (2017). "Ambiotherm," in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 1731–1742.

Ranasinghe, N., Jain, P., Thi Ngoc Tram, N., Koh, K. C. R., Tolley, D., Karwita, S., et al. (2018). "Season traveller," in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1–13.

Rothe, S., Althammer, F., and Khamis, M. (2018). "GazeRecall: using gaze direction to increase recall of details in cinematic virtual reality," in Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia, 115–119.

Rothe, S., Buschek, D., and Hußmann, H. (2019). Guidance in cinematic virtual reality-taxonomy, research status and challenges. *Multimodal Technol. Interact.* 3 (1), 19. doi:10.3390/mti3010019

Rothe, S., and Hußmann, H. (2018). "Guiding the viewer in cinematic virtual reality by diegetic cues," in *Augmented reality, virtual reality, and computer graphics. AVR 2018. Lecture notes in computer science* Editors L. De Paolis, and P. Bourdot (Cham: Springer), 101–117. doi:10.1007/978-3-319-95270-3_7

Rothe, S., Hußmann, H., and Allary, M. (2017). "Diegetic cues for guiding the viewer in cinematic virtual reality," in Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, 1–2.

Schmitz, A., Macquarrie, A., Julier, S., Binetti, N., and Steed, A. (2020). "Directing versus attracting attention: exploring the effectiveness of central and peripheral cues in panoramic videos," in Proceedings - 2020 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2020, 63–72.

Sheikh, A., Brown, A., Watson, Z., and Evans, M. (2016). Directing attention in 360-degree video. *IBC 2016 Conf.*, 1–9. doi:10.1049/ibc.2016.0029

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Trans. R. Soc. B Biol. Sci.* 364 (1535), 3549–3557. doi:10.1098/rstb.2009.0138

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

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* 73 (4), 971–995. doi:10.3758/s13414-010-0073-7

Suzuki, K., Abe, K., and Sato, H. (2020). Proposal of perception method of existence of objects in 3D space using quasi-electrostatic field. *Int. Conf. Human-Computer Interact.*, 561–571. doi:10.1007/978-3-030-49760-6_40

Tanaka, Y., Nishida, J., and Lopes, P. (2022). Electrical head actuation: enabling interactive systems to directly manipulate head orientation. *Proc. 2022 CHI Conf. Hum. Factors Comput. Syst.* 1, 1–15. doi:10.1145/3491102.3501910

Tong, L., Jung, S., and Lindeman, R. W. (2019). "Action units: directing user attention in 360-degree video based VR," in 25th ACM Symposium on Virtual Reality Software and Technology, 1–2.

Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12 (1), 97–136. doi:10.1016/0010-0285(80)90005-5

Walker, B. N., and Lindsay, J. (2003). "Effect of beacon sounds on navigation performance in a virtual reality environment," in Proceedings of the 9th International Conference on Auditory Display (ICAD2003), July, 204–207.

Wallgrun, J. O., Bagher, M. M., Sajjadi, P., and Klippel, A. (2020). "A comparison of visual attention guiding approaches for 360° image-based VR tours," in 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 83–91.

Wang, D., Li, T., Zhang, Y., and Hou, J. (2016). Survey on multisensory feedback virtual reality dental training systems. *Eur. J. Dent. Educ.* 20 (4), 248–260. doi:10.1111/eje.12173

Yamaguchi, S., Ogawa, N., and Narumi, T. (2021). "Now I'm not afraid: reducing fear of missing out in 360° videos on a head-mounted display using a panoramic thumbnail," in 2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 176–183.

Zimmerman, A., Bai, L., and Ginty, D. D. (2014). The gentle touch receptors of mammalian skin. *Science* 346 (6212), 950–954. doi:10.1126/science.1254229

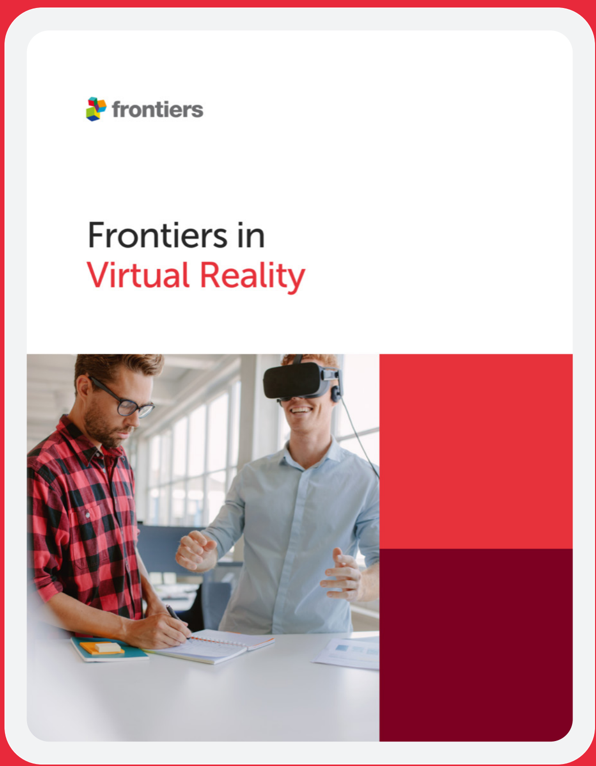
Frontiers in Virtual Reality

Explores the possibilities of virtual and extended reality

An exciting new journal in its field which advances our understanding of extended reality to develop new technologies and find applications for society.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact